

ECE 230L

Textbook

You shall delete it 24 hours
after you download it.

No commercial use permitted.

Semiconductor Physics and Devices

Basic Principles

Fourth Edition

Donald A. Neamen

University of New Mexico



CONTENTS

Preface	x
Prologue—Semiconductors and the Integrated Circuit	xvii

PART I—Semiconductor Material Properties

CHAPTER 1

The Crystal Structure of Solids 1

1.0	Preview	1
1.1	Semiconductor Materials	1
1.2	Types of Solids	2
1.3	Space Lattices	3
1.3.1	Primitive and Unit Cell	3
1.3.2	Basic Crystal Structures	4
1.3.3	Crystal Planes and Miller Indices	6
1.3.4	Directions in Crystals	9
1.4	The Diamond Structure	10
1.5	Atomic Bonding	12
*1.6	Imperfections and Impurities in Solids	14
1.6.1	Imperfections in Solids	14
1.6.2	Impurities in Solids	16
*1.7	Growth of Semiconductor Materials	17
1.7.1	Growth from a Melt	17
1.7.2	Epitaxial Growth	19
1.8	Summary	20
	Problems	21

CHAPTER 2

Introduction to Quantum Mechanics 25

2.0	Preview	25
2.1	Principles of Quantum Mechanics	26
2.1.1	Energy Quanta	26
2.1.2	Wave–Particle Duality	27
2.1.3	The Uncertainty Principle	30

2.2	Schrodinger's Wave Equation	31
2.2.1	<i>The Wave Equation</i>	31
2.2.2	<i>Physical Meaning of the Wave Function</i>	32
2.2.3	<i>Boundary Conditions</i>	33
2.3	Applications of Schrodinger's Wave Equation	34
2.3.1	<i>Electron in Free Space</i>	35
2.3.2	<i>The Infinite Potential Well</i>	36
2.3.3	<i>The Step Potential Function</i>	39
2.3.4	<i>The Potential Barrier and Tunneling</i>	44
2.4	Extensions of the Wave Theory to Atoms	46
2.4.1	<i>The One-Electron Atom</i>	46
2.4.2	<i>The Periodic Table</i>	50
2.5	Summary	51
	Problems	52

CHAPTER 3

Introduction to the Quantum Theory of Solids 58

3.0	Preview	58
3.1	Allowed and Forbidden Energy Bands	59
3.1.1	<i>Formation of Energy Bands</i>	59
*3.1.2	<i>The Kronig–Penney Model</i>	63
3.1.3	<i>The k-Space Diagram</i>	67
3.2	Electrical Conduction in Solids	72
3.2.1	<i>The Energy Band and the Bond Model</i>	72
3.2.2	<i>Drift Current</i>	74
3.2.3	<i>Electron Effective Mass</i>	75
3.2.4	<i>Concept of the Hole</i>	78
3.2.5	<i>Metals, Insulators, and Semiconductors</i>	80
3.3	Extension to Three Dimensions	83
3.3.1	<i>The k-Space Diagrams of Si and GaAs</i>	83
3.3.2	<i>Additional Effective Mass Concepts</i>	85

3.4	Density of States Function	85
3.4.1	<i>Mathematical Derivation</i>	85
3.4.2	<i>Extension to Semiconductors</i>	88
3.5	Statistical Mechanics	91
3.5.1	<i>Statistical Laws</i>	91
3.5.2	<i>The Fermi–Dirac Probability Function</i>	91
3.5.3	<i>The Distribution Function and the Fermi Energy</i>	93
3.6	Summary	98
	Problems	100

C H A P T E R 4**The Semiconductor in Equilibrium 106**

4.0	Preview	106
4.1	Charge Carriers in Semiconductors	107
4.1.1	<i>Equilibrium Distribution of Electrons and Holes</i>	107
4.1.2	<i>The n_0 and p_0 Equations</i>	109
4.1.3	<i>The Intrinsic Carrier Concentration</i>	113
4.1.4	<i>The Intrinsic Fermi-Level Position</i>	116
4.2	Dopant Atoms and Energy Levels	118
4.2.1	<i>Qualitative Description</i>	118
4.2.2	<i>Ionization Energy</i>	120
4.2.3	<i>Group III–V Semiconductors</i>	122
4.3	The Extrinsic Semiconductor	123
4.3.1	<i>Equilibrium Distribution of Electrons and Holes</i>	123
4.3.2	<i>The $n_0 p_0$ Product</i>	127
*4.3.3	<i>The Fermi–Dirac Integral</i>	128
4.3.4	<i>Degenerate and Nondegenerate Semiconductors</i>	130
4.4	Statistics of Donors and Acceptors	131
4.4.1	<i>Probability Function</i>	131
4.4.2	<i>Complete Ionization and Freeze-Out</i>	132
4.5	Charge Neutrality	135
4.5.1	<i>Compensated Semiconductors</i>	135
4.5.2	<i>Equilibrium Electron and Hole Concentrations</i>	136
4.6	Position of Fermi Energy Level	141
4.6.1	<i>Mathematical Derivation</i>	142
4.6.2	<i>Variation of E_F with Doping Concentration and Temperature</i>	144
4.6.3	<i>Relevance of the Fermi Energy</i>	145

4.7 Summary 147**Problems** 149**C H A P T E R 5****Carrier Transport Phenomena 156**

5.0	Preview	156
5.1	Carrier Drift	157
5.1.1	<i>Drift Current Density</i>	157
5.1.2	<i>Mobility Effects</i>	159
5.1.3	<i>Conductivity</i>	164
5.1.4	<i>Velocity Saturation</i>	169
5.2	Carrier Diffusion	172
5.2.1	<i>Diffusion Current Density</i>	172
5.2.2	<i>Total Current Density</i>	175
5.3	Graded Impurity Distribution	176
5.3.1	<i>Induced Electric Field</i>	176
5.3.2	<i>The Einstein Relation</i>	178
*5.4	The Hall Effect	180
5.5	Summary	183
	Problems	184

C H A P T E R 6**Nonequilibrium Excess Carriers in Semiconductors 192**

6.0	Preview	192
6.1	Carrier Generation and Recombination	193
6.1.1	<i>The Semiconductor in Equilibrium</i>	193
6.1.2	<i>Excess Carrier Generation and Recombination</i>	194
6.2	Characteristics of Excess Carriers	198
6.2.1	<i>Continuity Equations</i>	198
6.2.2	<i>Time-Dependent Diffusion Equations</i>	199
6.3	Ambipolar Transport	201
6.3.1	<i>Derivation of the Ambipolar Transport Equation</i>	201
6.3.2	<i>Limits of Extrinsic Doping and Low Injection</i>	203
6.3.3	<i>Applications of the Ambipolar Transport Equation</i>	206
6.3.4	<i>Dielectric Relaxation Time Constant</i>	214
*6.3.5	<i>Haynes–Shockley Experiment</i>	216

6.4	Quasi-Fermi Energy Levels	219	8.1.4	Minority Carrier Distribution	283
*6.5	Excess Carrier Lifetime	221	8.1.5	Ideal pn Junction Current	286
	6.5.1 Shockley–Read–Hall Theory of Recombination	221	8.1.6	Summary of Physics	290
	6.5.2 Limits of Extrinsic Doping and Low Injection	225	8.1.7	Temperature Effects	292
*6.6	Surface Effects	227	8.1.8	The “Short” Diode	293
	6.6.1 Surface States	227	8.2	Generation–Recombination Currents and High-Injection Levels	295
	6.6.2 Surface Recombination Velocity	229	8.2.1	Generation–Recombination Currents	296
6.7	Summary	231	8.2.2	High-Level Injection	302
	Problems	233	8.3	Small-Signal Model of the pn Junction	304

PART II—Fundamental Semiconductor Devices

CHAPTER 7

The pn Junction 241

7.0	Preview	241	7.0	Preview	241
7.1	Basic Structure of the pn Junction	242	7.1	Basic Structure of the pn Junction	242
7.2	Zero Applied Bias	243	7.2.1	Built-in Potential Barrier	243
	7.2.2	Electric Field	7.2.2	Electric Field	246
	7.2.3	Space Charge Width	7.2.3	Space Charge Width	249
7.3	Reverse Applied Bias	251	7.3.1	Space Charge Width and Electric Field	251
	7.3.2	Junction Capacitance	7.3.2	Junction Capacitance	254
	7.3.3	One-Sided Junctions	7.3.3	One-Sided Junctions	256
7.4	Junction Breakdown	258	7.4	Junction Breakdown	258
*7.5	Nonuniformly Doped Junctions	262	*7.5.1	Linearly Graded Junctions	263
	7.5.2	Hyperabrupt Junctions	7.5.2	Hyperabrupt Junctions	265
7.6	Summary	267	7.6	Summary	267
	Problems	269		Problems	269

CHAPTER 8

The pn Junction Diode 276

8.0	Preview	276	8.0	Preview	276
8.1	pn Junction Current	277	8.1.1	Qualitative Description of Charge Flow in a pn Junction	277
	8.1.2	Ideal Current–Voltage Relationship	8.1.2	Ideal Current–Voltage Relationship	278
	8.1.3	Boundary Conditions	8.1.3	Boundary Conditions	279

	8.1.4	Minority Carrier Distribution	8.1.4	Minority Carrier Distribution	283
	8.1.5	Ideal pn Junction Current	8.1.5	Ideal pn Junction Current	286
	8.1.6	Summary of Physics	8.1.6	Summary of Physics	290
	8.1.7	Temperature Effects	8.1.7	Temperature Effects	292
	8.1.8	The “Short” Diode	8.1.8	The “Short” Diode	293
8.2	Generation–Recombination Currents and High-Injection Levels	295	8.2	Generation–Recombination Currents and High-Injection Levels	295
	8.2.1	Generation–Recombination Currents	8.2.1	Generation–Recombination Currents	296
	8.2.2	High-Level Injection	8.2.2	High-Level Injection	302
8.3	Small-Signal Model of the pn Junction	304	8.3	Small-Signal Model of the pn Junction	304
	8.3.1	Diffusion Resistance	8.3.1	Diffusion Resistance	305
	8.3.2	Small-Signal Admittance	8.3.2	Small-Signal Admittance	306
	8.3.3	Equivalent Circuit	8.3.3	Equivalent Circuit	313
*8.4	Charge Storage and Diode Transients	314	*8.4	Charge Storage and Diode Transients	314
	8.4.1	The Turn-off Transient	8.4.1	The Turn-off Transient	315
	8.4.2	The Turn-on Transient	8.4.2	The Turn-on Transient	317
*8.5	The Tunnel Diode	318	*8.5	The Tunnel Diode	318
8.6	Summary	321	8.6	Summary	321
	Problems	323		Problems	323
			CHAPTER 9		
			Metal–Semiconductor and Semiconductor Heterojunctions	331	
9.0	Preview	331	9.0	Preview	331
9.1	The Schottky Barrier Diode	332	9.1	The Schottky Barrier Diode	332
	9.1.1	Qualitative Characteristics	9.1.1	Qualitative Characteristics	332
	9.1.2	Ideal Junction Properties	9.1.2	Ideal Junction Properties	334
	9.1.3	Nonideal Effects on the Barrier Height	9.1.3	Nonideal Effects on the Barrier Height	338
	9.1.4	Current–Voltage Relationship	9.1.4	Current–Voltage Relationship	342
	9.1.5	Comparison of the Schottky Barrier Diode and the pn Junction Diode	9.1.5	Comparison of the Schottky Barrier Diode and the pn Junction Diode	345
9.2	Metal–Semiconductor Ohmic Contacts	349	9.2	Metal–Semiconductor Ohmic Contacts	349
	9.2.1	Ideal Nonrectifying Barrier	9.2.1	Ideal Nonrectifying Barrier	349
	9.2.2	Tunneling Barrier	9.2.2	Tunneling Barrier	351
	9.2.3	Specific Contact Resistance	9.2.3	Specific Contact Resistance	352
9.3	Heterojunctions	354	9.3	Heterojunctions	354
	9.3.1	Heterojunction Materials	9.3.1	Heterojunction Materials	354
	9.3.2	Energy-Band Diagrams	9.3.2	Energy-Band Diagrams	354
	9.3.3	Two-Dimensional Electron Gas	9.3.3	Two-Dimensional Electron Gas	356
	9.3.4	Equilibrium Electrostatics	9.3.4	Equilibrium Electrostatics	358
	9.3.5	Current–Voltage Characteristics	9.3.5	Current–Voltage Characteristics	363

- 9.4** Summary 363
Problems 365

CHAPTER 10

Fundamentals of the Metal–Oxide–Semiconductor Field-Effect Transistor 371

- 10.0** Preview 371
10.1 The Two-Terminal MOS Structure 372
 10.1.1 Energy-Band Diagrams 372
 10.1.2 Depletion Layer Thickness 376
 10.1.3 Surface Charge Density 380
 10.1.4 Work Function Differences 382
 10.1.5 Flat-Band Voltage 385
 10.1.6 Threshold Voltage 388
10.2 Capacitance–Voltage Characteristics 394
 10.2.1 Ideal C–V Characteristics 394
 10.2.2 Frequency Effects 399
 10.2.3 Fixed Oxide and Interface Charge Effects 400
10.3 The Basic MOSFET Operation 403
 10.3.1 MOSFET Structures 403
 10.3.2 Current–Voltage Relationship—Concepts 404
 *10.3.3 Current–Voltage Relationship—Mathematical Derivation 410
 10.3.4 Transconductance 418
 10.3.5 Substrate Bias Effects 419
10.4 Frequency Limitations 422
 10.4.1 Small-Signal Equivalent Circuit 422
 10.4.2 Frequency Limitation Factors and Cutoff Frequency 425
***10.5** The CMOS Technology 427
10.6 Summary 430
Problems 433

CHAPTER 11

Metal–Oxide–Semiconductor Field-Effect Transistor: Additional Concepts 443

- 11.0** Preview 443
11.1 Nonideal Effects 444
 11.1.1 Subthreshold Conduction 444

- 11.1.2 Channel Length Modulation 446
 11.1.3 Mobility Variation 450
 11.1.4 Velocity Saturation 452
 11.1.5 Ballistic Transport 453
11.2 MOSFET Scaling 455
 11.2.1 Constant-Field Scaling 455
 11.2.2 Threshold Voltage—First Approximation 456
 11.2.3 Generalized Scaling 457
11.3 Threshold Voltage Modifications 457
 11.3.1 Short-Channel Effects 457
 11.3.2 Narrow-Channel Effects 461
11.4 Additional Electrical Characteristics 464
 11.4.1 Breakdown Voltage 464
 *11.4.2 The Lightly Doped Drain Transistor 470
 11.4.3 Threshold Adjustment by Ion Implantation 472
***11.5** Radiation and Hot-Electron Effects 475
 11.5.1 Radiation-Induced Oxide Charge 475
 11.5.2 Radiation-Induced Interface States 478
 11.5.3 Hot-Electron Charging Effects 480
11.6 Summary 481
Problems 483

CHAPTER 12

The Bipolar Transistor 491

- 12.0** Preview 491
12.1 The Bipolar Transistor Action 492
 12.1.1 The Basic Principle of Operation 493
 12.1.2 Simplified Transistor Current Relation—Qualitative Discussion 495
 12.1.3 The Modes of Operation 498
 12.1.4 Amplification with Bipolar Transistors 500
12.2 Minority Carrier Distribution 501
 12.2.1 Forward-Active Mode 502
 12.2.2 Other Modes of Operation 508
12.3 Transistor Currents and Low-Frequency Common-Base Current Gain 509
 12.3.1 Current Gain—Contributing Factors 509
 12.3.2 Derivation of Transistor Current Components and Current Gain Factors 512

12.3.3	<i>Summary</i>	517	*13.3	Nonideal Effects	593
12.3.4	<i>Example Calculations of the Gain Factors</i>	517	13.3.1	<i>Channel Length Modulation</i>	594
12.4	Nonideal Effects	522	13.3.2	<i>Velocity Saturation Effects</i>	596
12.4.1	<i>Base Width Modulation</i>	522	13.3.3	<i>Subthreshold and Gate Current Effects</i>	596
12.4.2	<i>High Injection</i>	524	*13.4	Equivalent Circuit and Frequency Limitations	598
12.4.3	<i>Emitter Bandgap Narrowing</i>	526	13.4.1	<i>Small-Signal Equivalent Circuit</i>	598
12.4.4	<i>Current Crowding</i>	528	13.4.2	<i>Frequency Limitation Factors and Cutoff Frequency</i>	600
*12.4.5	<i>Nonuniform Base Doping</i>	530	*13.5	High Electron Mobility Transistor	602
12.4.6	<i>Breakdown Voltage</i>	531	13.5.1	<i>Quantum Well Structures</i>	603
12.5	Equivalent Circuit Models	536	13.5.2	<i>Transistor Performance</i>	604
*12.5.1	<i>Ebers–Moll Model</i>	537	13.6	Summary	609
12.5.2	<i>Gummel–Poon Model</i>	540	Problems	611	
12.5.3	<i>Hybrid-Pi Model</i>	541	PART III—Specialized Semiconductor Devices		
12.6	Frequency Limitations	545	CHAPTER 14		
12.6.1	<i>Time-Delay Factors</i>	545	Optical Devices	618	
12.6.2	<i>Transistor Cutoff Frequency</i>	546			
12.7	Large-Signal Switching	549	14.0	Preview	618
12.7.1	<i>Switching Characteristics</i>	549			
12.7.2	<i>The Schottky-Clamped Transistor</i>	551	14.1	Optical Absorption	619
*12.8	Other Bipolar Transistor Structures	552	14.1.1	<i>Photon Absorption Coefficient</i>	619
12.8.1	<i>Polysilicon Emitter BJT</i>	552	14.1.2	<i>Electron–Hole Pair Generation Rate</i>	622
12.8.2	<i>Silicon–Germanium Base Transistor</i>	554			
12.8.3	<i>Heterojunction Bipolar Transistors</i>	556	14.2	Solar Cells	624
12.9	Summary	558	14.2.1	<i>The pn Junction Solar Cell</i>	624
Problems	560	14.2.2	<i>Conversion Efficiency and Solar Concentration</i>	627	
C H A P T E R 13		14.2.3	<i>Nonuniform Absorption Effects</i>	628	
The Junction Field-Effect Transistor	571	14.2.4	<i>The Heterojunction Solar Cell</i>	629	
13.0	Preview	571	14.2.5	<i>Amorphous Silicon Solar Cells</i>	630
13.1	JFET Concepts	572			
13.1.1	<i>Basic pn JFET Operation</i>	572	14.3	Photodetectors	633
13.1.2	<i>Basic MESFET Operation</i>	576	14.3.1	<i>Photoconductor</i>	633
13.2	The Device Characteristics	578	14.3.2	<i>Photodiode</i>	635
13.2.1	<i>Internal Pinchoff Voltage, Pinchoff Voltage, and Drain-to-Source Saturation Voltage</i>	578	14.3.3	<i>PIN Photodiode</i>	640
13.2.2	<i>Ideal DC Current–Voltage Relationship—Depletion Mode JFET</i>	582	14.3.4	<i>Avalanche Photodiode</i>	641
13.2.3	<i>Transconductance</i>	587	14.3.5	<i>Phototransistor</i>	642
13.2.4	<i>The MESFET</i>	588			
14.4	Photoluminescence and Electroluminescence	643			
14.4.1	<i>Basic Transitions</i>	644			
14.4.2	<i>Luminescent Efficiency</i>	645			
14.4.3	<i>Materials</i>	646			

14.5	Light Emitting Diodes	648
14.5.1	<i>Generation of Light</i>	648
14.5.2	<i>Internal Quantum Efficiency</i>	649
14.5.3	<i>External Quantum Efficiency</i>	650
14.5.4	<i>LED Devices</i>	652
14.6	Laser Diodes	654
14.6.1	<i>Stimulated Emission and Population Inversion</i>	655
14.6.2	<i>Optical Cavity</i>	657
14.6.3	<i>Threshold Current</i>	658
14.6.4	<i>Device Structures and Characteristics</i>	660
14.7	Summary	661
	Problems	664

CHAPTER 15
Semiconductor Microwave and Power Devices 670

15.0	Preview	670
15.1	Tunnel Diode	671
15.2	Gunn Diode	672
15.3	Impatt Diode	675
15.4	Power Bipolar Transistors	677
15.4.1	<i>Vertical Power Transistor Structure</i>	677
15.4.2	<i>Power Transistor Characteristics</i>	678
15.4.3	<i>Darlington Pair Configuration</i>	682
15.5	Power MOSFETs	684
15.5.1	<i>Power Transistor Structures</i>	684
15.5.2	<i>Power MOSFET Characteristics</i>	685
15.5.3	<i>Parasitic BJT</i>	689
15.6	The Thyristor	691
15.6.1	<i>The Basic Characteristics</i>	691
15.6.2	<i>Triggering the SCR</i>	694

15.6.3	<i>SCR Turn-Off</i>	697
15.6.4	<i>Device Structures</i>	697
15.7	Summary	701
	Problems	703

APPENDIX A
Selected List of Symbols 707

APPENDIX B
System of Units, Conversion Factors, and General Constants 715

APPENDIX C
The Periodic Table 719

APPENDIX D
Unit of Energy—The Electron Volt 720

APPENDIX E
“Derivation” of Schrodinger’s Wave Equation 722

APPENDIX F
Effective Mass Concepts 724

APPENDIX G
The Error Function 729

APPENDIX H
Answers to Selected Problems 730

Index 738

1

C H A P T E R

The Crystal Structure of Solids

This text deals with the electrical properties and characteristics of semiconductor materials and devices. The electrical properties of solids are therefore of primary interest. The semiconductor is in general a single-crystal material. The electrical properties of a single-crystal material are determined not only by the chemical composition but also by the arrangement of atoms in the solid; this being true, a brief study of the crystal structure of solids is warranted. The formation, or growth, of the single-crystal material is an important part of semiconductor technology. A short discussion of several growth techniques is included in this chapter to provide the reader with some of the terminology that describes semiconductor device structures. ■

1.0 | PREVIEW

In this chapter, we will:

- Describe three classifications of solids—amorphous, polycrystalline, and single crystal.
- Discuss the concept of a unit cell.
- Describe three simple crystal structures and determine the volume and surface density of atoms in each structure.
- Describe the diamond crystal structure.
- Briefly discuss several methods of forming single-crystal semiconductor materials.

1.1 | SEMICONDUCTOR MATERIALS

Semiconductors are a group of materials having conductivities between those of metals and insulators. Two general classifications of semiconductors are the elemental semiconductor materials, found in group IV of the periodic table, and the compound semiconductor materials, most of which are formed from special combinations of group III and group V elements. Table 1.1 shows a portion of the periodic table in

Table 1.1 | A portion of the periodic table

III	IV	V
5 B Boron	6 C Carbon	
13 Al Aluminum	14 Si Silicon	15 P Phosphorus
31 Ga Gallium	32 Ge Germanium	33 As Arsenic
49 In Indium		51 Sb Antimony

Table 1.2 | A list of some semiconductor materials

Elemental semiconductors	
Si	Silicon
Ge	Germanium
Compound semiconductors	
AlP	Aluminum phosphide
AlAs	Aluminum arsenide
GaP	Gallium phosphide
GaAs	Gallium arsenide
InP	Indium phosphide

which the more common semiconductors are found and Table 1.2 lists a few of the semiconductor materials. (Semiconductors can also be formed from combinations of group II and group VI elements, but in general these will not be considered in this text.)

The elemental materials, those that are composed of single species of atoms, are silicon and germanium. Silicon is by far the most common semiconductor used in integrated circuits and will be emphasized to a great extent.

The two-element, or *binary*, compounds such as gallium arsenide or gallium phosphide are formed by combining one group III and one group V element. Gallium arsenide is one of the more common of the compound semiconductors. Its good optical properties make it useful in optical devices. GaAs is also used in specialized applications in which, for example, high speed is required.

We can also form a three-element, or *ternary*, compound semiconductor. An example is $\text{Al}_x\text{Ga}_{1-x}\text{As}$, in which the subscript x indicates the fraction of the lower atomic number element component. More complex semiconductors can also be formed that provide flexibility when choosing material properties.

1.2 | TYPES OF SOLIDS

Amorphous, polycrystalline, and single crystals are the three general types of solids. Each type is characterized by the size of an ordered region within the material. An ordered region is a spatial volume in which atoms or molecules have a regular geometric arrangement or periodicity. Amorphous materials have order only within a few atomic or molecular dimensions, while polycrystalline materials have a high degree of order over many atomic or molecular dimensions. These ordered regions, or single-crystal regions, vary in size and orientation with respect to one another. The single-crystal regions are called grains and are separated from one another by grain boundaries. Single-crystal materials, ideally, have a high degree of order, or regular geometric periodicity, throughout the entire volume of the material. The advantage of a single-crystal material is that, in general, its electrical properties are superior

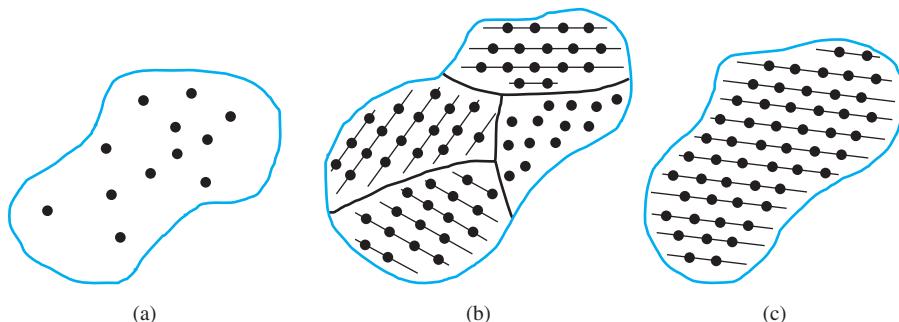


Figure 1.1 | Schematics of three general types of crystals: (a) amorphous, (b) polycrystalline, (c) single.

to those of a nonsingle-crystal material, since grain boundaries tend to degrade the electrical characteristics. Two-dimensional representations of amorphous, polycrystalline, and single-crystal materials are shown in Figure 1.1.

1.3 | SPACE LATTICES

Our primary emphasis in this text will be on the single-crystal material with its regular geometric periodicity in the atomic arrangement. A representative unit, or a group of atoms, is repeated at regular intervals in each of the three dimensions to form the single crystal. The periodic arrangement of atoms in the crystal is called the *lattice*.

1.3.1 Primitive and Unit Cell

We can represent a particular atomic array by a dot that is called a *lattice point*. Figure 1.2 shows an infinite two-dimensional array of lattice points. The simplest means of repeating an atomic array is by translation. Each lattice point in Figure 1.2 can be translated a distance a_1 in one direction and a distance b_1 in a second noncolinear direction to generate the two-dimensional lattice. A third noncolinear translation will produce the three-dimensional lattice. The translation directions need not be perpendicular.

Since the three-dimensional lattice is a periodic repetition of a group of atoms, we do not need to consider the entire lattice, but only a fundamental unit that is being repeated. A *unit cell* is a small volume of the crystal that can be used to reproduce the entire crystal. A unit cell is not a unique entity. Figure 1.3 shows several possible unit cells in a two-dimensional lattice.

The unit cell A can be translated in directions a_2 and b_2 , the unit cell B can be translated in directions a_3 and b_3 , and the entire two-dimensional lattice can be constructed by the translations of either of these unit cells. The unit cells C and D in Figure 1.3 can also be used to construct the entire lattice by using the appropriate translations. This discussion of two-dimensional unit cells can easily be extended to three dimensions to describe a real single-crystal material.

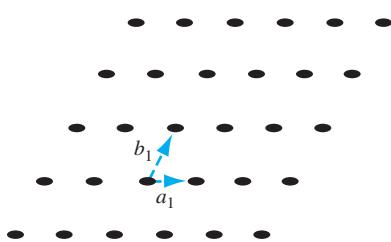


Figure 1.2 | Two-dimensional representation of a single-crystal lattice.

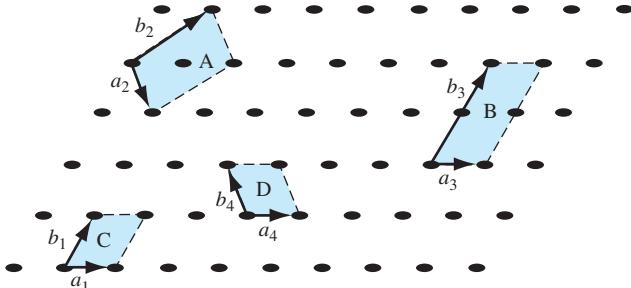


Figure 1.3 | Two-dimensional representation of a single-crystal lattice showing various possible unit cells.

A *primitive cell* is the smallest unit cell that can be repeated to form the lattice. In many cases, it is more convenient to use a unit cell that is not a primitive cell. Unit cells may be chosen that have orthogonal sides, for example, whereas the sides of a primitive cell may be nonorthogonal.

A generalized three-dimensional unit cell is shown in Figure 1.4. The relationship between this cell and the lattice is characterized by three vectors \bar{a} , \bar{b} , and \bar{c} , which need not be perpendicular and which may or may not be equal in length. Every equivalent lattice point in the three-dimensional crystal can be found using the vector

$$\bar{r} = p\bar{a} + q\bar{b} + s\bar{c} \quad (1.1)$$

where p , q , and s are integers. Since the location of the origin is arbitrary, we will let p , q , and s be positive integers for simplicity. The magnitudes of the vectors \bar{a} , \bar{b} , and \bar{c} are the lattice constants of the unit cell.

1.3.2 Basic Crystal Structures

Before we discuss the semiconductor crystal, let us consider three crystal structures and determine some of the basic characteristics of these crystals. Figure 1.5 shows the simple cubic, body-centered cubic, and face-centered cubic structures. For these simple structures, we may choose unit cells such that the general vectors \bar{a} , \bar{b} , and \bar{c}

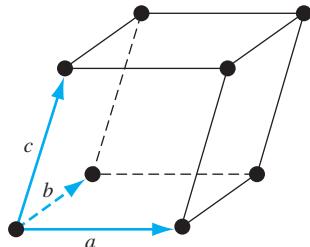


Figure 1.4 | A generalized primitive unit cell.

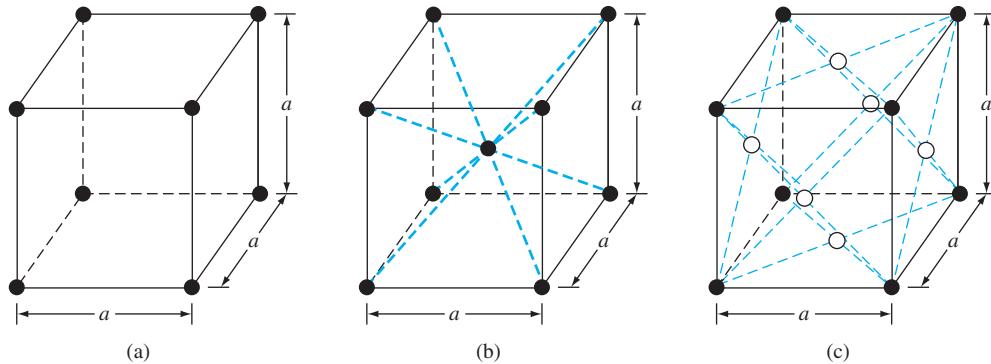


Figure 1.5 | Three lattice types: (a) simple cubic, (b) body-centered cubic, (c) face-centered cubic.

are perpendicular to each other and the lengths are equal. The lattice constant of each unit cell in Figure 1.5 is designated as “ a .” The *simple cubic* (sc) structure has an atom located at each corner; the *body-centered cubic* (bcc) structure has an additional atom at the center of the cube; and the *face-centered cubic* (fcc) structure has additional atoms on each face plane.

By knowing the crystal structure of a material and its lattice dimensions, we can determine several characteristics of the crystal. For example, we can determine the volume density of atoms.

Objective: Find the volume density of atoms in a crystal.

EXAMPLE 1.1

Consider a single-crystal material that is a body-centered cubic, as shown in Figure 1.5b, with a lattice constant $a = 5 \text{ \AA} = 5 \times 10^{-8} \text{ cm}$. A corner atom is shared by eight unit cells that meet at each corner so that each corner atom effectively contributes one-eighth of its volume to each unit cell. The eight corner atoms then contribute an equivalent of one atom to the unit cell. If we add the body-centered atom to the corner atoms, each unit cell contains an equivalent of two atoms.

Solution

The number of atoms per unit cell is $\frac{1}{8} \times 8 + 1 = 2$

The volume density of atoms is then found as

$$\text{Volume Density} = \frac{\# \text{ atoms per unit cell}}{\text{volume of unit cell}}$$

So

$$\text{Volume Density} = \frac{2}{a^3} = \frac{2}{(5 \times 10^{-8})^3} = 1.6 \times 10^{22} \text{ atoms/cm}^3$$

EXERCISE PROBLEM

Ex 1.1 The lattice constant of a face-centered cubic lattice is 4.25 \AA . Determine the
(a) effective number of atoms per unit cell and (b) volume density of atoms.

[Ans. (a) 4; (b) $5.21 \times 10^{22} \text{ cm}^{-3}$]

1.3.3 Crystal Planes and Miller Indices

Since real crystals are not infinitely large, they eventually terminate at a surface. Semiconductor devices are fabricated at or near a surface, so the surface properties may influence the device characteristics. We would like to be able to describe these surfaces in terms of the lattice. Surfaces, or planes through the crystal, can be described by first considering the intercepts of the plane along the \bar{a} , \bar{b} , and \bar{c} axes used to describe the lattice.

EXAMPLE 1.2

Objective: Describe the plane shown in Figure 1.6. (The lattice points in Figure 1.6 are shown along the \bar{a} , \bar{b} , and \bar{c} axes only.)

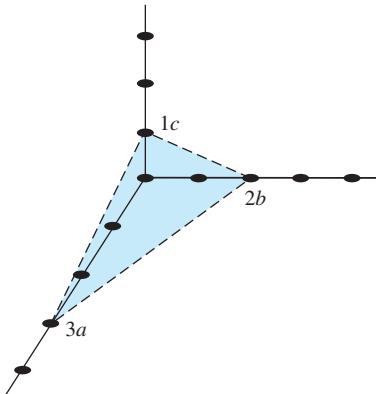


Figure 1.6 | A representative crystal-lattice plane.

■ Solution

From Equation (1.1), the intercepts of the plane correspond to $p = 3$, $q = 2$, and $s = 1$. Now write the reciprocals of the intercepts, which gives

$$\left(\frac{1}{3}, \frac{1}{2}, \frac{1}{1}\right)$$

Multiply by the lowest common denominator, which in this case is 6, to obtain $(2, 3, 6)$. The plane in Figure 1.6 is then referred to as the (236) plane. The integers are referred to as the Miller indices. We will refer to a general plane as the (hkl) plane.

■ Comment

We can show that the same three Miller indices are obtained for any plane that is parallel to the one shown in Figure 1.6. Any parallel plane is entirely equivalent to any other.

■ EXERCISE PROBLEM

Ex 1.2 Describe the lattice plane shown in Figure 1.7. [Ans. (211) plane]

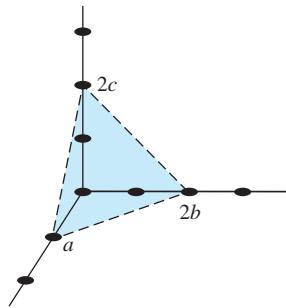


Figure 1.7 | Figure for Exercise Problem Ex 1.2.

Three planes that are commonly considered in a cubic crystal are shown in Figure 1.8. The plane in Figure 1.8a is parallel to the \bar{b} and \bar{c} axes so the intercepts are given as $p = 1$, $q = \infty$, and $s = \infty$. Taking the reciprocal, we obtain the Miller indices as $(1, 0, 0)$, so the plane shown in Figure 1.8a is referred to as the (100) plane. Again, any plane parallel to the one shown in Figure 1.8a and separated by an integral number of lattice constants is equivalent and is referred to as the (100) plane. One advantage to taking the reciprocal of the intercepts to obtain the Miller indices is that the use of infinity is avoided when describing a plane that is parallel to an axis. If we were to describe a plane passing through the origin of our system, we would obtain infinity as one or more of the Miller indices after taking the reciprocal of the intercepts. However, the location of the origin of our system is entirely arbitrary and so, by translating the origin to another equivalent lattice point, we can avoid the use of infinity in the set of Miller indices.

For the simple cubic structure, the body-centered cubic, and the face-centered cubic, there is a high degree of symmetry. The axes can be rotated by 90° in each

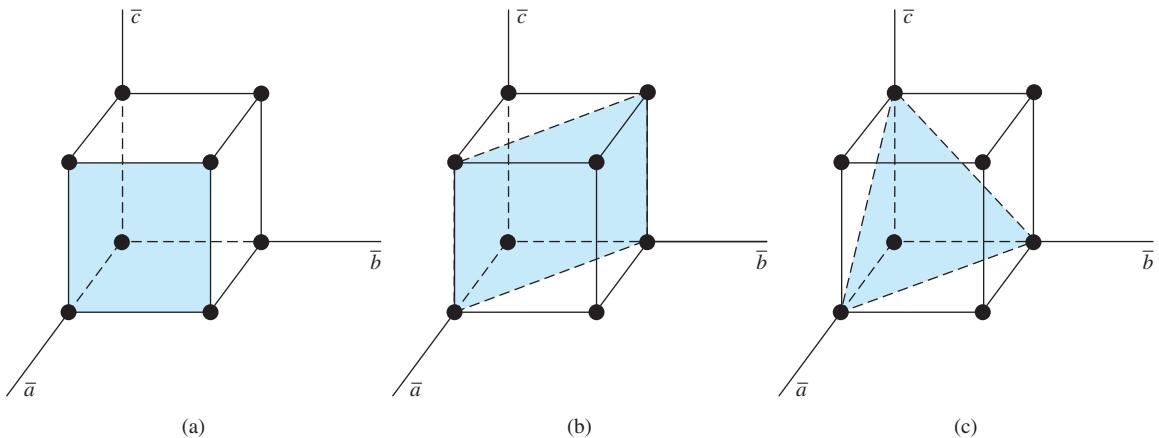


Figure 1.8 | Three lattice planes: (a) (100) plane, (b) (110) plane, (c) (111) plane.

of the three dimensions and each lattice point can again be described by Equation (1.1) as

$$\bar{r} = p\bar{a} + q\bar{b} + s\bar{c} \quad (1.1)$$

Each face plane of the cubic structure shown in Figure 1.8a is entirely equivalent. These planes are grouped together and are referred to as the {100} set of planes.

We may also consider the planes shown in Figures 1.8b and 1.8c. The intercepts of the plane shown in Figure 1.8b are $p = 1$, $q = 1$, and $s = \infty$. The Miller indices are found by taking the reciprocal of these intercepts and, as a result, this plane is referred to as the (110) plane. In a similar way, the plane shown in Figure 1.8c is referred to as the (111) plane.

One characteristic of a crystal that can be determined is the distance between nearest equivalent parallel planes. Another characteristic is the surface concentration of atoms, number per square centimeter ($\#/cm^2$), that are cut by a particular plane. Again, a single-crystal semiconductor is not infinitely large and must terminate at some surface. The surface density of atoms may be important, for example, in determining how another material, such as an insulator, will “fit” on the surface of a semiconductor material.

EXAMPLE 1.3

Objective: Calculate the surface density of atoms on a particular plane in a crystal.

Consider the body-centered cubic structure and the (110) plane shown in Figure 1.9a. Assume the atoms can be represented as hard spheres with the closest atoms touching each other. Assume the lattice constant is $a_1 = 5 \text{ \AA}$. Figure 1.9b shows how the atoms are cut by the (110) plane.

The atom at each corner is shared by four similar equivalent lattice planes, so each corner atom effectively contributes one-fourth of its area to this lattice plane as indicated in the figure. The four corner atoms then effectively contribute one atom to this lattice plane. The atom in the center is completely enclosed in the lattice plane. There is no other equivalent plane that

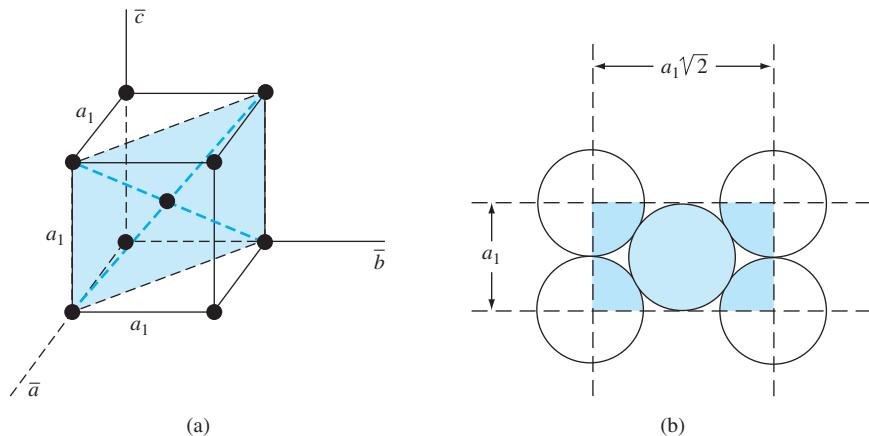


Figure 1.9 | (a) The (110) plane in a body-centered cubic and (b) the atoms cut by the (110) plane in a body-centered cubic.

cuts the center atom and the corner atoms, so the entire center atom is included in the number of atoms in the crystal plane. The lattice plane in Figure 1.9b, then, contains two atoms.

■ Solution

The number of atoms per lattice plane is $\frac{1}{4} \times 4 + 1 = 2$

The surface density of atoms is then found as

$$\text{Surface Density} = \frac{\# \text{ of atoms per lattice plane}}{\text{area of lattice plane}}$$

So

$$\begin{aligned}\text{Surface Density} &= \frac{2}{(a_1)(a_1\sqrt{2})} = \frac{2}{(5 \times 10^{-8})^2\sqrt{2}} \\ &= 5.66 \times 10^{14} \text{ atoms/cm}^2\end{aligned}$$

■ Comment

The surface density of atoms is a function of the particular crystal plane in the lattice and generally varies from one crystal plane to another.

■ EXERCISE PROBLEM

Ex 1.3 The lattice constant of a face-centered-cubic structure is 4.25 Å. Calculate the surface density of atoms for a (a) (100) plane and (b) (110) plane.

[Ans. (a) $1.11 \times 10^{15} \text{ cm}^{-2}$; (b) $7.83 \times 10^{14} \text{ cm}^{-2}$]

1.3.4 Directions in Crystals

In addition to describing crystal planes in a lattice, we may want to describe a particular direction in the crystal. The direction can be expressed as a set of three integers that are the components of a vector in that direction. For example, the body diagonal in a simple cubic lattice is composed of vector components 1, 1, 1. The body diagonal is then described as the [111] direction. The brackets are used to designate direction as distinct from the parentheses used for the crystal planes. The three basic directions and the associated crystal planes for the simple cubic structure are shown in Figure 1.10. Note that in the simple cubic lattices, the $[hkl]$ direction is perpendicular to the (hkl) plane. This perpendicularity may not be true in noncubic lattices.

TEST YOUR UNDERSTANDING

TYU 1.1 The volume density of atoms for a simple cubic lattice is $4 \times 10^{22} \text{ cm}^{-3}$. Assume that the atoms are hard spheres with each atom touching its nearest neighbor. Determine the lattice constant and the radius of the atom. (Ans. $a = 2.92 \text{ \AA}$, $r = 1.46 \text{ \AA}$)

TYU 1.2 Consider a simple cubic structure with a lattice constant of $a = 4.65 \text{ \AA}$. Determine the surface density of atoms in the (a) (100) plane, (b) (110) plane, and (c) (111) plane. (Ans. (a) $4.62 \times 10^{14} \text{ cm}^{-2}$; (b) $3.27 \times 10^{14} \text{ cm}^{-2}$; (c) $2.67 \times 10^{14} \text{ cm}^{-2}$)

TYU 1.3 (a) Determine the distance between nearest (100) planes in a simple cubic lattice with a lattice constant of $a = 4.83 \text{ \AA}$. (b) Repeat part (a) for the (110) plane.

[Ans. (a) 4.83 \AA ; (b) 3.42 \AA]

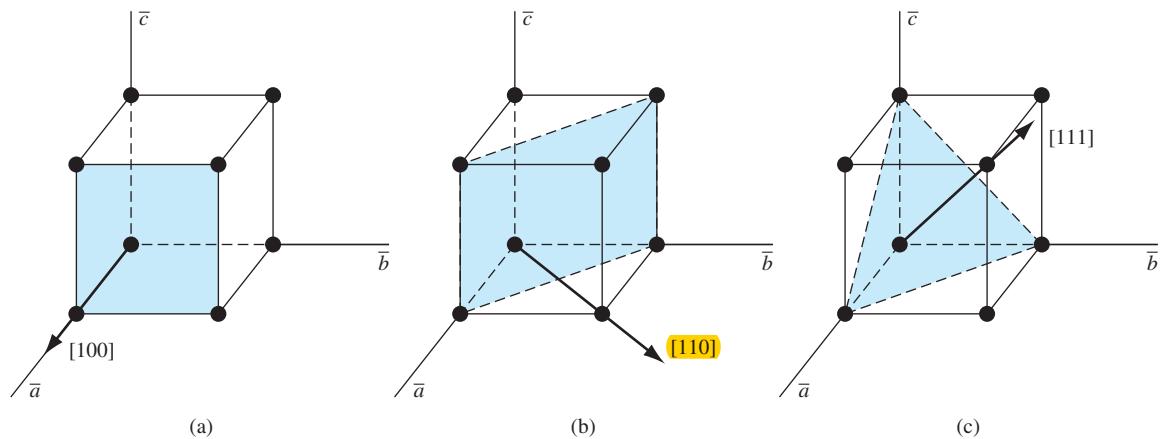


Figure 1.10 | Three lattice directions and planes: (a) (100) plane and [100] direction, (b) (110) plane and [110] direction, (c) (111) plane and [111] direction.

1.4 | THE DIAMOND STRUCTURE

As already stated, silicon is the most common semiconductor material. Silicon is referred to as a group IV element and has a diamond crystal structure. Germanium is also a group IV element and has the same diamond structure. A unit cell of the diamond structure, shown in Figure 1.11, is more complicated than the simple cubic structures that we have considered up to this point.

We may begin to understand the diamond lattice by considering the tetrahedral structure shown in Figure 1.12. This structure is basically a body-centered cubic with four of the corner atoms missing. Every atom in the tetrahedral structure has four nearest neighbors and it is this structure that is the basic building block of the diamond lattice.

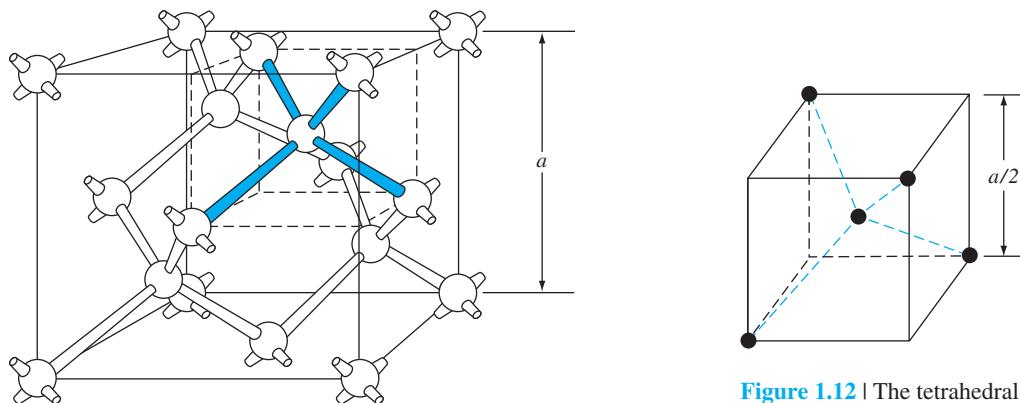


Figure 1.11 | The diamond structure.

Figure 1.12 | The tetrahedral structure of closest neighbors in the diamond lattice.

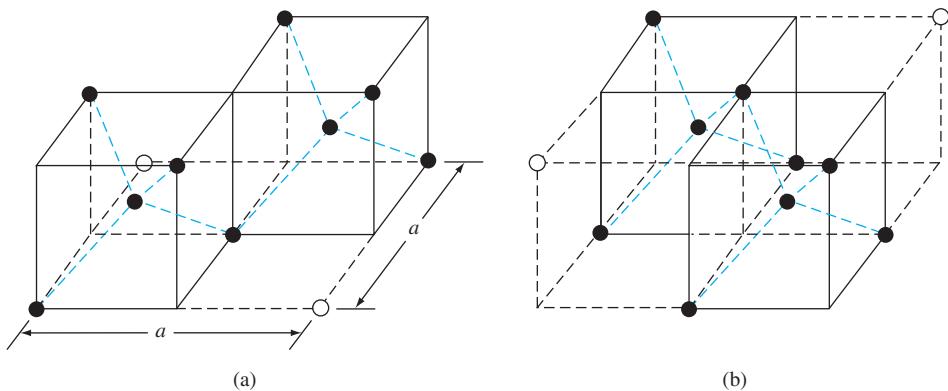


Figure 1.13 Portions of the diamond lattice: (a) bottom half and (b) top half.

There are several ways to visualize the diamond structure. One way to gain a further understanding of the diamond lattice is by considering Figure 1.13. Figure 1.13a shows two body-centered cubic, or tetrahedral, structures diagonally adjacent to each other. The open circles represent atoms in the lattice that are generated when the structure is translated to the right or left, one lattice constant, a . Figure 1.13b represents the top half of the diamond structure. The top half again consists of two tetrahedral structures joined diagonally, but which are at 90° with respect to the bottom-half diagonal. An important characteristic of the diamond lattice is that any atom within the diamond structure will have four nearest neighboring atoms. We will note this characteristic again in our discussion of atomic bonding in the next section.

The diamond structure refers to the particular lattice in which all atoms are of the same species, such as silicon or germanium. The zincblende (sphalerite) structure differs from the diamond structure only in that there are two different types of atoms in the lattice. Compound semiconductors, such as gallium arsenide, have the zincblende structure shown in Figure 1.14. The important feature of both the diamond and the zincblende structures is that the atoms are joined together to form a tetrahedron. Figure 1.15 shows the basic tetrahedral structure of GaAs in which each Ga atom has four nearest As neighbors and each As atom has four nearest Ga neighbors. This figure also begins to show the interpenetration of two sublattices that can be used to generate the diamond or zincblende lattice.

TEST YOUR UNDERSTANDING

- TYU 1.4** Consider the diamond unit cell shown in Figure 1.11. Determine the (a) number of corner atoms, (b) number of face-centered atoms, and (c) number of atoms totally enclosed in the unit cell. [Ans. (a) 8; (b) 6; (c) 4]

- TYU 1.5** The lattice constant of silicon is 5.43 Å. Calculate the volume density of silicon atoms. ($\text{Ans. } 5 \times 10^{22} \text{ cm}^{-3}$)

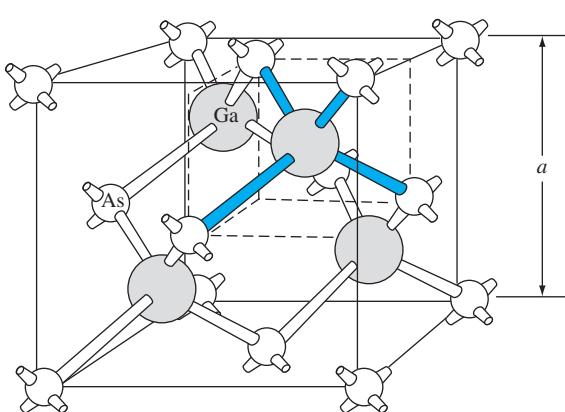


Figure 1.14 | The zincblende (sphalerite) lattice of GaAs.

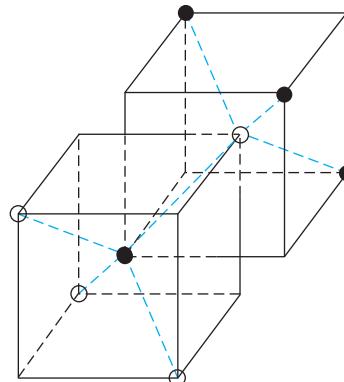


Figure 1.15 | The tetrahedral structure of closest neighbors in the zincblende lattice.

1.5 | ATOMIC BONDING

We have been considering various single-crystal structures. The question arises as to why one particular crystal structure is favored over another for a particular assembly of atoms. A fundamental law of nature is that the total energy of a system in thermal equilibrium tends to reach a minimum value. The interaction that occurs between atoms to form a solid and to reach the minimum total energy depends on the type of atom or atoms involved. The type of bond, or interaction, between atoms, then, depends on the particular atom or atoms in the crystal. If there is not a strong bond between atoms, they will not “stick together” to create a solid.

The interaction between atoms can be described by quantum mechanics. Although an introduction to quantum mechanics is presented in the next chapter, the quantum-mechanical description of the atomic bonding interaction is still beyond the scope of this text. We can nevertheless obtain a qualitative understanding of how various atoms interact by considering the valence, or outermost, electrons of an atom.

The atoms at the two extremes of the periodic table (excepting the inert elements) tend to lose or gain valence electrons, thus forming ions. These ions then essentially have complete outer energy shells. The elements in group I of the periodic table tend to lose their one electron and become positively charged, while the elements in group VII tend to gain an electron and become negatively charged. These oppositely charged ions then experience a coulomb attraction and form a bond referred to as an *ionic bond*. If the ions were to get too close, a repulsive force would become dominant, so an equilibrium distance results between these two ions. In a crystal, negatively charged ions tend to be surrounded by positively charged ions and positively charged ions tend to be surrounded by negatively charged ions, so a periodic array of the atoms is formed to create the lattice. A classic example of ionic bonding is sodium chloride.

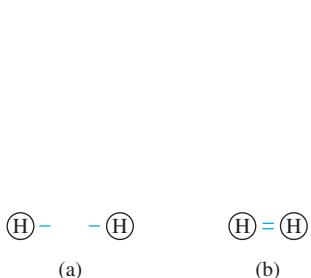


Figure 1.16 | Representation of (a) hydrogen valence electrons and (b) covalent bonding in a hydrogen molecule.

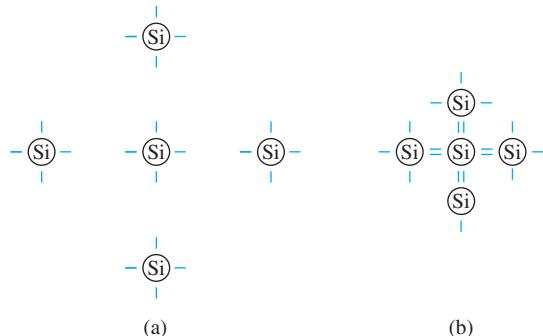


Figure 1.17 | Representation of (a) silicon valence electrons and (b) covalent bonding in the silicon crystal.

The interaction of atoms tends to form closed valence shells such as we see in ionic bonding. Another atomic bond that tends to achieve closed-valence energy shells is *covalent bonding*, an example of which is found in the hydrogen molecule. A hydrogen atom has one electron and needs one more electron to complete the lowest energy shell. A schematic of two noninteracting hydrogen atoms, and the hydrogen molecule with the covalent bonding, is shown in Figure 1.16. Covalent bonding results in electrons being shared between atoms, so that in effect the valence energy shell of each atom is full.

Atoms in group IV of the periodic table, such as silicon and germanium, also tend to form covalent bonds. Each of these elements has four valence electrons and needs four more electrons to complete the valence energy shell. If a silicon atom, for example, has four nearest neighbors, with each neighbor atom contributing one valence electron to be shared, then the center atom will in effect have eight electrons in its outer shell. Figure 1.17a schematically shows five noninteracting silicon atoms with the four valence electrons around each atom. A two-dimensional representation of the covalent bonding in silicon is shown in Figure 1.17b. The center atom has eight shared valence electrons.

A significant difference between the covalent bonding of hydrogen and of silicon is that, when the hydrogen molecule is formed, it has no additional electrons to form additional covalent bonds, while the outer silicon atoms always have valence electrons available for additional covalent bonding. The silicon array may then be formed into an infinite crystal, with each silicon atom having four nearest neighbors and eight shared electrons. The four nearest neighbors in silicon forming the covalent bond correspond to the tetrahedral structure and the diamond lattice, which were shown in Figures 1.12 and 1.11 respectively. Atomic bonding and crystal structure are obviously directly related.

The third major atomic bonding scheme is referred to as *metallic bonding*. Group I elements have one valence electron. If two sodium atoms ($Z = 11$), for example, are brought into close proximity, the valence electrons interact in a way similar to that in covalent bonding. When a third sodium atom is brought into close proximity with the

first two, the valence electrons can also interact and continue to form a bond. Solid sodium has a body-centered cubic structure, so each atom has eight nearest neighbors with each atom sharing many valence electrons. We may think of the positive metallic ions as being surrounded by a sea of negative electrons, the solid being held together by the electrostatic forces. This description gives a qualitative picture of the metallic bond.

A fourth type of atomic bond, called the *Van der Waals* bond, is the weakest of the chemical bonds. A hydrogen fluoride (HF) molecule, for example, is formed by an ionic bond. The effective center of the positive charge of the molecule is not the same as the effective center of the negative charge. This nonsymmetry in the charge distribution results in a small electric dipole that can interact with the dipoles of other HF molecules. With these weak interactions, solids formed by the Van der Waals bonds have a relatively low melting temperature—in fact, most of these materials are in gaseous form at room temperature.

*1.6 | IMPERFECTIONS AND IMPURITIES IN SOLIDS

Up to this point, we have been considering an ideal single-crystal structure. In a real crystal, the lattice is not perfect, but contains imperfections or defects; that is, the perfect geometric periodicity is disrupted in some manner. Imperfections tend to alter the electrical properties of a material and, in some cases, electrical parameters can be dominated by these defects or impurities.

1.6.1 Imperfections in Solids

One type of imperfection that all crystals have in common is atomic thermal vibration. A perfect single crystal contains atoms at particular lattice sites, the atoms separated from each other by a distance we have assumed to be constant. The atoms in a crystal, however, have a certain thermal energy, which is a function of temperature. The thermal energy causes the atoms to vibrate in a random manner about an equilibrium lattice point. This random thermal motion causes the distance between atoms to randomly fluctuate, slightly disrupting the perfect geometric arrangement of atoms. This imperfection, called *lattice vibrations*, affects some electrical parameters, as we will see later in our discussion of semiconductor material characteristics.

Another type of defect is called a *point defect*. There are several of this type that we need to consider. Again, in an ideal single-crystal lattice, the atoms are arranged in a perfect periodic arrangement. However, in a real crystal, an atom may be missing from a particular lattice site. This defect is referred to as a *vacancy*; it is schematically shown in Figure 1.18a. In another situation, an atom may be located between lattice sites. This defect is referred to as an *interstitial* and is schematically shown in Figure 1.18b. In the case of vacancy and interstitial defects, not only is the perfect

*Indicates sections that will aid in the total summation of understanding of semiconductor devices, but may be skipped the first time through the text without loss of continuity.

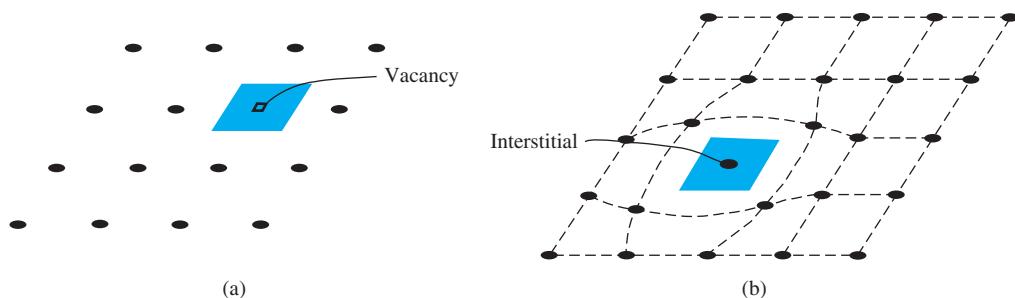


Figure 1.18 | Two-dimensional representation of a single-crystal lattice showing (a) a vacancy defect and (b) an interstitial defect.

geometric arrangement of atoms broken but also the ideal chemical bonding between atoms is disrupted, which tends to change the electrical properties of the material. A vacancy and interstitial may be in close enough proximity to exhibit an interaction between the two point defects. This vacancy–interstitial defect, also known as a *Frenkel defect*, produces different effects than the simple vacancy or interstitial.

The point defects involve single atoms or single-atom locations. In forming single-crystal materials, more complex defects may occur. A line defect, for example, occurs when an entire row of atoms is missing from its normal lattice site. This defect is referred to as a *line dislocation* and is shown in Figure 1.19. As with a point defect, a line dislocation disrupts both the normal geometric periodicity of the lattice and the ideal atomic bonds in the crystal. This dislocation can also alter the electrical properties of the material, usually in a more unpredictable manner than the simple point defects.

Other complex dislocations can also occur in a crystal lattice. However, this introductory discussion is intended only to present a few of the basic types of defect, and to show that a real crystal is not necessarily a perfect lattice structure. The effect of these imperfections on the electrical properties of a semiconductor will be considered in later chapters.

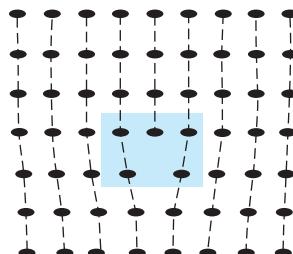


Figure 1.19 | A two-dimensional representation of a line dislocation.

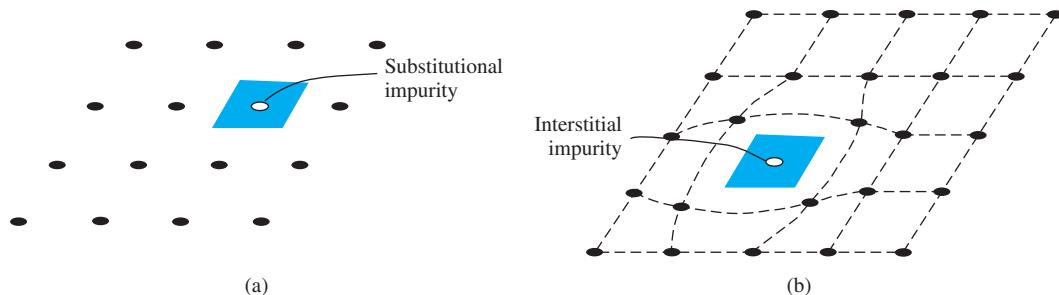


Figure 1.20 | Two-dimensional representation of a single-crystal lattice showing (a) a substitutional impurity and (b) an interstitial impurity.

1.6.2 Impurities in Solids

Foreign atoms, or impurity atoms, may be present in a crystal lattice. Impurity atoms may be located at normal lattice sites, in which case they are called *substitutional* impurities. Impurity atoms may also be located between normal sites, in which case they are called *interstitial* impurities. Both these impurities are lattice defects and are schematically shown in Figure 1.20. Some impurities, such as oxygen in silicon, tend to be essentially inert; however, other impurities, such as gold or phosphorus in silicon, can drastically alter the electrical properties of the material.

In Chapter 4 we will see that, by adding controlled amounts of particular impurity atoms, the electrical characteristics of a semiconductor material can be favorably altered. The technique of adding impurity atoms to a semiconductor material in order to change its conductivity is called *doping*. There are two general methods of doping: impurity diffusion and ion implantation.

The actual diffusion process depends to some extent on the material but, in general, impurity diffusion occurs when a semiconductor crystal is placed in a high-temperature ($\approx 1000^\circ\text{C}$) gaseous atmosphere containing the desired impurity atom. At this high temperature, many of the crystal atoms can randomly move in and out of their single-crystal lattice sites. Vacancies may be created by this random motion so that impurity atoms can move through the lattice by hopping from one vacancy to another. Impurity diffusion is the process by which impurity particles move from a region of high concentration near the surface to a region of lower concentration within the crystal. When the temperature decreases, the impurity atoms become permanently frozen into the substitutional lattice sites. Diffusion of various impurities into selected regions of a semiconductor allows us to fabricate complex electronic circuits in a single semiconductor crystal.

Ion implantation generally takes place at a lower temperature than diffusion. A beam of impurity ions is accelerated to kinetic energies in the range of 50 keV or greater and then directed to the surface of the semiconductor. The high-energy impurity ions enter the crystal and come to rest at some average depth from the surface. One advantage of ion implantation is that controlled numbers of impurity atoms can be introduced into specific regions of the crystal. A disadvantage of this technique is that the incident impurity atoms collide with the crystal atoms, causing

lattice-displacement damage. However, most of the lattice damage can be removed by thermal annealing, in which the temperature of the crystal is raised for a short time. Thermal annealing is a required step after implantation.

*1.7 | GROWTH OF SEMICONDUCTOR MATERIALS

The success in fabricating very large scale integrated (VLSI) circuits is a result, to a large extent, of the development of and improvement in the formation or growth of pure single-crystal semiconductor materials. Semiconductors are some of the purest materials. Silicon, for example, has concentrations of most impurities of less than 1 part in 10^{10} atoms. The high purity requirement means that extreme care is necessary in the growth and the treatment of the material at each step of the fabrication process. The mechanics and kinetics of crystal growth are extremely complex and will be described in only very general terms in this text. However, a general knowledge of the growth techniques and terminology is valuable.

1.7.1 Growth from a Melt

A common technique for growing single-crystal materials is called the *Czochralski method*. In this technique, a small piece of single-crystal material, known as a *seed*, is brought into contact with the surface of the same material in liquid phase, and then slowly pulled from the melt. As the seed is slowly pulled, solidification occurs along the plane between the solid–liquid interface. Usually the crystal is also rotated slowly as it is being pulled, to provide a slight stirring action to the melt, resulting in a more uniform temperature. Controlled amounts of specific impurity atoms, such as boron or phosphorus, may be added to the melt so that the grown semiconductor crystal is intentionally doped with the impurity atom. Figure 1.21a shows a schematic of the Czochralski growth process and a silicon ingot or boule grown by this process.

Some impurities may be present in the ingot that are undesirable. Zone refining is a common technique for purifying material. A high-temperature coil, or r-f induction coil, is slowly passed along the length of the boule. The temperature induced by the coil is high enough so that a thin layer of liquid is formed. At the solid–liquid interface, there is a distribution of impurities between the two phases. The parameter that describes this distribution is called the *segregation coefficient*: the ratio of the concentration of impurities in the solid to the concentration in the liquid. If the segregation coefficient is 0.1, for example, the concentration of impurities in the liquid is a factor of 10 greater than that in the solid. As the liquid zone moves through the material, the impurities are driven along with the liquid. After several passes of the r-f coil, most impurities are at the end of the bar, which can then be cut off. The moving molten zone, or the zone-refining technique, can result in considerable purification.

After the semiconductor is grown, the boule is mechanically trimmed to the proper diameter and a flat is ground over the entire length of the boule to denote

*Indicates sections that will aid in the total summation of understanding of semiconductor devices, but may be skipped the first time through the text without loss of continuity.

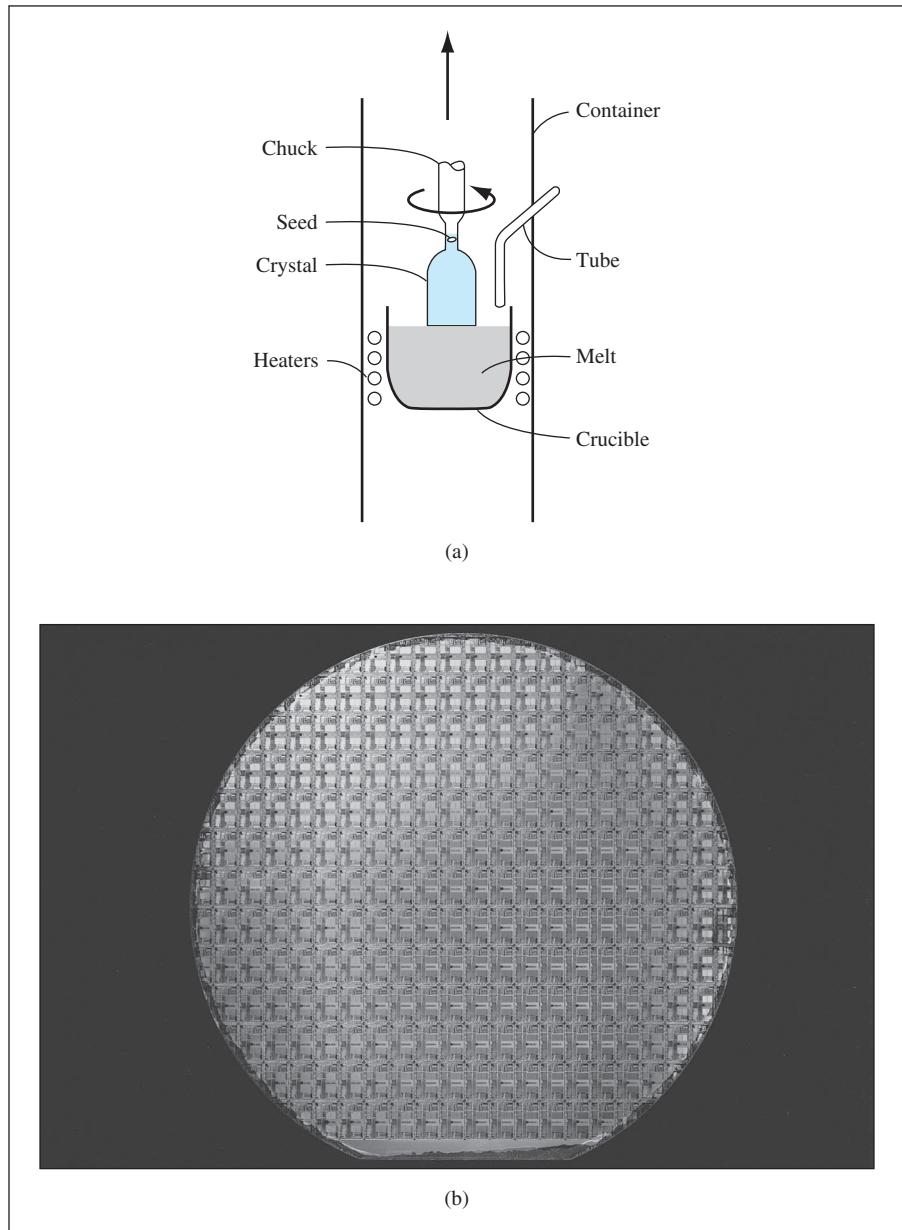


Figure 1.21 | (a) Model of a crystal puller and (b) photograph of a silicon wafer with an array of integrated circuits. The circuits are tested on the wafer then sawed apart into chips that are mounted into packages. (Photo courtesy of Intel Corporation.)

the crystal orientation. The flat is perpendicular to the [110] direction or indicates the (110) plane. (See Figure 1.21b.) This then allows the individual chips to be fabricated along given crystal planes so that the chips can be sawed apart more easily. The boule is then sliced into wafers. The wafer must be thick enough to mechanically support itself. A mechanical two-sided lapping operation produces a flat wafer of uniform thickness. Since the lapping procedure can leave a surface damaged and contaminated by the mechanical operation, the surface must be removed by chemical etching. The final step is polishing. This provides a smooth surface on which devices may be fabricated or further growth processes may be carried out. This final semiconductor wafer is called the substrate material.

1.7.2 Epitaxial Growth

A common and versatile growth technique that is used extensively in device and integrated circuit fabrication is epitaxial growth. *Epitaxial growth* is a process whereby a thin, single-crystal layer of material is grown on the surface of a single-crystal substrate. In the epitaxial process, the single-crystal substrate acts as the seed, although the process takes place far below the melting temperature. When an epitaxial layer is grown on a substrate of the same material, the process is termed *homoepitaxy*. Growing silicon on a silicon substrate is one example of a homoepitaxy process. At present, a great deal of work is being done with *heteroepitaxy*. In a heteroepitaxy process, although the substrate and epitaxial materials are not the same, the two crystal structures should be very similar if single-crystal growth is to be obtained and if a large number of defects are to be avoided at the epitaxial–substrate interface. Growing epitaxial layers of the ternary alloy AlGaAs on a GaAs substrate is one example of a heteroepitaxy process.

One epitaxial growth technique that has been used extensively is called *chemical vapor-phase deposition* (CVD). Silicon epitaxial layers, for example, are grown on silicon substrates by the controlled deposition of silicon atoms onto the surface from a chemical vapor containing silicon. In one method, silicon tetrachloride reacts with hydrogen at the surface of a heated substrate. The silicon atoms are released in the reaction and can be deposited onto the substrate, while the other chemical reactant, HCl, is in gaseous form and is swept out of the reactor. A sharp demarcation between the impurity doping in the substrate and in the epitaxial layer can be achieved using the CVD process. This technique allows great flexibility in the fabrication of semiconductor devices.

Liquid-phase epitaxy is another epitaxial growth technique. A compound of the semiconductor with another element may have a melting temperature lower than that of the semiconductor itself. The semiconductor substrate is held in the liquid compound and, since the temperature of the melt is lower than the melting temperature of the substrate, the substrate does not melt. As the solution is slowly cooled, a single-crystal semiconductor layer grows on the seed crystal. This technique, which occurs at a lower temperature than the Czochralski method, is useful in growing group III–V compound semiconductors.

A versatile technique for growing epitaxial layers is the *molecular beam epitaxy* (MBE) process. A substrate is held in vacuum at a temperature normally in the range

of 400 to 800°C, a relatively low temperature compared with many semiconductor-processing steps. Semiconductor and dopant atoms are then evaporated onto the surface of the substrate. In this technique, the doping can be precisely controlled resulting in very complex doping profiles. Complex ternary compounds, such as AlGaAs, can be grown on substrates, such as GaAs, where abrupt changes in the crystal composition are desired. Many layers of various types of epitaxial compositions can be grown on a substrate in this manner. These structures are extremely beneficial in optical devices such as laser diodes.

1.8 | SUMMARY

- A few of the most common semiconductor materials were listed. Silicon is the most common semiconductor material and appears in column IV of the periodic table.
- The properties of semiconductors and other materials are determined to a large extent by the single-crystal lattice structure. The unit cell is a small volume of the crystal that is used to reproduce the entire crystal. Three basic unit cells are the simple cubic, body-centered cubic, and face-centered cubic.
- Silicon has the diamond crystal structure. Atoms are formed in a tetrahedral configuration with four nearest neighbor atoms. The binary semiconductors have a zincblende lattice that is basically the same as the diamond lattice.
- Miller indices are used to describe planes in a crystal lattice. These planes may be used to describe the surface of a semiconductor material. The Miller indices are also used to describe directions in a crystal.
- Imperfections do exist in semiconductor materials. A few of these imperfections are vacancies, substitutional impurities, and interstitial impurities. Small amounts of controlled substitutional impurities can favorably alter semiconductor properties as we will see in later chapters.
- A brief description of semiconductor growth methods was given. Bulk growth, such as the Czochralski method, produces the starting semiconductor material or substrate. Epitaxial growth can be used to control the surface properties of a semiconductor. Most semiconductor devices are fabricated in the epitaxial layer.

GLOSSARY OF IMPORTANT TERMS

binary semiconductor A two-element compound semiconductor, such as gallium arsenide (GaAs).

covalent bonding The bonding between atoms in which valence electrons are shared.

diamond lattice The atomic crystal structure of silicon, for example, in which each atom has four nearest neighbors in a tetrahedral configuration.

doping The process of adding specific types of atoms to a semiconductor to favorably alter the electrical characteristics.

elemental semiconductor A semiconductor composed of a single species of atom, such as silicon or germanium.

epitaxial layer A thin, single-crystal layer of material formed on the surface of a substrate.

ion implantation One particular process of doping a semiconductor.

lattice The periodic arrangement of atoms in a crystal.

Miller indices The set of integers used to describe a crystal plane.

primitive cell The smallest unit cell that can be repeated to form a lattice.

substrate A semiconductor wafer or other material used as the starting material for further semiconductor processing, such as epitaxial growth or diffusion.

ternary semiconductor A three-element compound semiconductor, such as aluminum gallium arsenide (AlGaAs).

unit cell A small volume of a crystal that can be used to reproduce the entire crystal.

zincblende lattice A lattice structure identical to the diamond lattice except that there are two types of atoms instead of one.

CHECKPOINT

After studying this chapter, the reader should have the ability to:

- List the most common elemental semiconductor material.
- Describe the concept of a unit cell.
- Determine the volume density of atoms for various lattice structures.
- Determine the Miller indices of a crystal-lattice plane.
- Sketch a lattice plane given the Miller indices.
- Determine the surface density of atoms on a given crystal-lattice plane.
- Describe the tetrahedral configuration of silicon atoms.
- Understand and describe various defects in a single-crystal lattice.

REVIEW QUESTIONS

1. List two elemental semiconductor materials and two compound semiconductor materials.
2. Sketch three lattice structures: (a) simple cubic, (b) body-centered cubic, and (c) face-centered cubic.
3. Describe the procedure for finding the volume density of atoms in a crystal.
4. Describe the procedure for obtaining the Miller indices that describe a plane in a crystal.
5. Describe the procedure for finding the surface density of atoms on a particular lattice plane.
6. Describe why a unit cell, that is not a primitive unit cell, might be preferable to a primitive unit cell.
7. Describe covalent bonding in silicon.
8. What is meant by a substitutional impurity in a crystal? What is meant by an interstitial impurity?

PROBLEMS

Section 1.3 Space Lattices

- 1.1 Determine the number of atoms per unit cell in a (a) face-centered cubic, (b) body-centered cubic, and (c) diamond lattice.
- 1.2 Assume that each atom is a hard sphere with the surface of each atom in contact with the surface of its nearest neighbor. Determine the percentage of total unit cell volume

that is occupied in (a) a simple cubic lattice, (b) a face-centered cubic lattice, (c) a body-centered cubic lattice, and (d) a diamond lattice.

- 1.3 If the lattice constant of silicon is 5.43 Å, calculate (a) the distance from the center of one silicon atom to the center of its nearest neighbor, (b) the number density of silicon atoms (#/cm³), and (c) the mass density (g/cm³) of silicon.
- 1.4 (a) The lattice constant of GaAs is 5.65 Å. Determine the number of Ga atoms and As atoms per cm³. (b) Determine the volume density of germanium atoms in a germanium semiconductor. The lattice constant of germanium is 5.65 Å.
- 1.5 The lattice constant of GaAs is $a = 5.65 \text{ \AA}$. Calculate (a) the distance between the centers of the nearest Ga and As atoms, and (b) the distance between the centers of the nearest As atoms.
- 1.6 Calculate the angle between any pair of bonds in the tetrahedral structure.
- 1.7 Assume the radius of an atom, which can be represented as a hard sphere, is $r = 1.95 \text{ \AA}$. The atom is placed in a (a) simple cubic, (b) fcc, (c) bcc, and (d) diamond lattice. Assuming that nearest atoms are touching each other, what is the lattice constant of each lattice?
- 1.8 A crystal is composed of two elements, A and B. The basic crystal structure is a face-centered cubic with element A at each of the corners and element B in the center of each face. The effective radius of element A is $r_A = 1.035 \text{ \AA}$. Assume that the elements are hard spheres with the surface of each A-type atom in contact with the surface of its nearest A-type neighbor. Calculate (a) the maximum radius of the B-type element that will fit into this structure, (b) the lattice constant, and (c) the volume density (#/cm³) of both the A-type atoms and the B-type atoms.
- 1.9 (a) A crystal with a simple cubic lattice structure is composed of atoms with an effective radius of $r = 2.25 \text{ \AA}$ and has an atomic weight of 12.5. Determine the mass density assuming the atoms are hard spheres and nearest neighbors are touching each other. (b) Repeat part (a) for a body-centered cubic structure.
- 1.10 A material, with a volume of 1 cm³, is composed of an fcc lattice with a lattice constant of 2.5 mm. The “atoms” in this material are actually coffee beans. Assume the coffee beans are hard spheres with each bean touching its nearest neighbor. Determine the volume of coffee after the coffee beans have been ground. (Assume 100% packing density of the ground coffee.)
- 1.11 The crystal structure of sodium chloride (NaCl) is a simple cubic with the Na and Cl atoms alternating positions. Each Na atom is then surrounded by six Cl atoms and likewise each Cl atom is surrounded by six Na atoms. (a) Sketch the atoms in a (100) plane. (b) Assume the atoms are hard spheres with nearest neighbors touching. The effective radius of Na is 1.0 Å and the effective radius of Cl is 1.8 Å. Determine the lattice constant. (c) Calculate the volume density of Na and Cl atoms. (d) Calculate the mass density of NaCl.
- 1.12 (a) A material is composed of two types of atoms. Atom A has an effective radius of 2.2 Å and atom B has an effective radius of 1.8 Å. The lattice is a bcc with atoms A at the corners and atom B in the center. Determine the lattice constant and the volume densities of A atoms and B atoms. (b) Repeat part (a) with atoms B at the corners and atom A in the center. (c) What comparison can be made of the materials in parts (a) and (b)?
- 1.13 (a) Consider the materials described in Problem 1.12(a) and 1.12(b). For each case, calculate the surface density of A atoms and B atoms in the (100) plane. What comparison can be made of the two materials? (b) Repeat part (a) for the (110) plane.

- 1.14** (a) The crystal structure of a particular material consists of a single atom in the center of a cube. The lattice constant is a_0 and the diameter of the atom is a_0 . Determine the volume density of atoms and the surface density of atoms in the (110) plane. (b) Compare the results of part (a) to the results for the case of the simple cubic structure shown in Figure 1.5a with the same lattice constant.
- 1.15** The lattice constant of a simple cubic lattice is a_0 . (a) Sketch the following planes: (i) (110), (ii) (111), (iii) (220), and (iv) (321). (b) Sketch the following directions: (i) [110], (ii) [111], (iii) [220], and (iv) [321].
- 1.16** For a simple cubic lattice, determine the Miller indices for the planes shown in Figure P1.16.
- 1.17** A body-centered cubic lattice has a lattice constant of 4.83 Å. A plane cutting the lattice has intercepts of 9.66 Å, 19.32 Å, and 14.49 Å along the three cartesian coordinates. What are the Miller indices of the plane?
- 1.18** The lattice constant of a simple cubic primitive cell is 5.28 Å. Determine the distance between the nearest parallel (a) (100), (b) (110), and (c) (111) planes.
- 1.19** The lattice constant of a single crystal is 4.73 Å. Calculate the surface density (#/cm²) of atoms on the (i) (100), (ii) (110), and (iii) (111) plane for a (a) simple cubic, (b) body-centered cubic, and (c) face-centered cubic lattice.
- 1.20** Determine the surface density of atoms for silicon on the (a) (100) plane, (b) (110) plane, and (c) (111) plane.
- 1.21** Consider a face-centered cubic lattice. Assume the atoms are hard spheres with the surfaces of the nearest neighbors touching. Assume the effective radius of the atom is 2.37 Å. (a) Determine the volume density of atoms in the crystal. (b) Calculate the surface density of atoms in the (110) plane. (c) Determine the distance between nearest (110) planes. (d) Repeat parts (b) and (c) for the (111) plane.

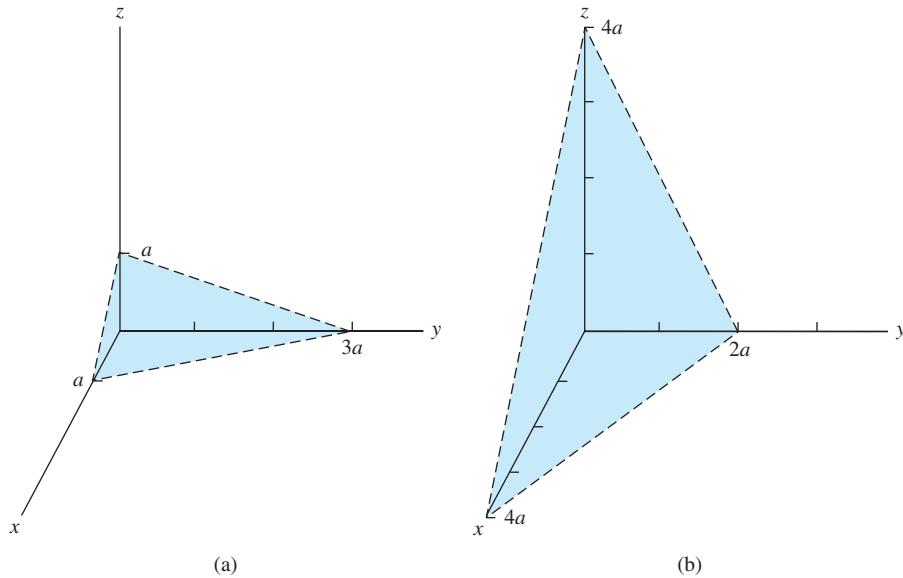


Figure P1.16 | Figure for Problem 1.16.

Section 1.5 Atomic Bonding

- 1.22** Calculate the density of valence electrons in silicon.
- 1.23** The structure of GaAs is the zincblende lattice. The lattice constant is 5.65 Å. Calculate the density of valence electrons in GaAs.

Section 1.6 Imperfections and Impurities in Solids

- 1.24** (a) If 5×10^{17} phosphorus atoms per cm³ are added to silicon as a substitutional impurity, determine the percentage of silicon atoms per unit volume that are displaced in the single crystal lattice. (b) Repeat part (a) for 2×10^{15} boron atoms per cm³ added to silicon.
- 1.25** (a) Assume that 2×10^{16} cm⁻³ of boron atoms are distributed homogeneously throughout single crystal silicon. What is the fraction by weight of boron in the crystal? (b) If phosphorus atoms, at a concentration of 10^{18} cm⁻³, are added to the material in part (a), determine the fraction by weight of phosphorus.
- 1.26** If 2×10^{16} cm⁻³ boron atoms are added to silicon as a substitutional impurity and are distributed uniformly throughout the semiconductor, determine the distance between boron atoms in terms of the silicon lattice constant. (Assume the boron atoms are distributed in a rectangular or cubic array.)
- 1.27** Repeat Problem 1.26 for 4×10^{15} cm⁻³ phosphorus atoms being added to silicon.

READING LIST

1. Azaroff, L. V., and J. J. Brophy. *Electronic Processes in Materials*. New York: McGraw-Hill, 1963.
2. Campbell, S. A. *The Science and Engineering of Microelectronic Fabrication*. New York: Oxford University Press, 1996.
3. Dimitrijev, S. *Principles of Semiconductor Devices*. New York: Oxford University Press, 2006.
4. Kittel, C. *Introduction to Solid State Physics*, 7th ed. Berlin: Springer-Verlag, 1993.
- *5. Li, S. S. *Semiconductor Physical Electronics*. New York: Plenum Press, 1993.
6. McKelvey, J. P. *Solid State Physics for Engineering and Materials Science*. Malabar, FL: Krieger, 1993.
7. Pierret, R. F. *Semiconductor Device Fundamentals*. Reading, MA: Addison-Wesley, 1996.
8. Runyan, W. R., and K. E. Bean. *Semiconductor Integrated Circuit Processing and Technology*. Reading, MA: Addison-Wesley, 1990.
9. Singh, J. *Semiconductor Devices: Basic Principles*. New York: John Wiley and Sons, 2001.
10. Streetman, B. G., and S. K. Banerjee. *Solid State Electronic Devices*, 6th ed. Upper Saddle River, NJ: Pearson Prentice Hall, 2006.
11. Sze, S. M. *VLSI Technology*. New York: McGraw-Hill, 1983.
- *12. Wolfe, C. M., N. Holonyak, Jr., and G. E. Stillman. *Physical Properties of Semiconductors*. Englewood Cliffs, NJ: Prentice Hall, 1989.

*Indicates references that are at an advanced level compared to this text.

C H A P T E R



Introduction to the Quantum Theory of Solids

In the last chapter, we applied quantum mechanics and Schrodinger's wave equation to determine the behavior of electrons in the presence of various potential functions. We found one important characteristic of an electron bound to an atom or bound within a finite space to be that the electron can take on only discrete values of energy; that is, the energies are quantized. We also discussed the Pauli exclusion principle, which stated that only one electron is allowed to occupy any given quantum state. In this chapter, we will generalize these concepts to the electron in a crystal lattice.

One of our goals is to determine the electrical properties of a semiconductor material, which we will then use to develop the current–voltage characteristics of semiconductor devices. Toward this end, we have two tasks in this chapter: to determine the properties of electrons in a crystal lattice and to determine the statistical characteristics of the very large number of electrons in a crystal. ■

3.0 | PREVIEW

In this chapter, we will:

- Develop the concept of allowed and forbidden electron energy bands in a single-crystal material, and describe conduction and valence energy bands in a semiconductor material.
- Discuss the concept of negatively charged electrons and positively charged holes as two distinct charge carriers in a semiconductor material.
- Develop electron energy versus momentum curves in a single-crystal material, which yields the concept of direct and indirect bandgap semiconductor materials.
- Discuss the concept of effective mass of an electron and a hole.
- Derive the density of quantum states in the allowed energy bands.

- Develop the Fermi-Dirac probability function, which describes the statistical distribution of electrons among the allowed energy levels, and define the Fermi energy level.

3.1 | ALLOWED AND FORBIDDEN ENERGY BANDS

In the last chapter, we considered the one-electron, or hydrogen, atom. That analysis showed that the energy of the bound electron is quantized: Only discrete values of electron energy are allowed. The radial probability density for the electron was also determined. This function gives the probability of finding the electron at a particular distance from the nucleus and shows that the electron is not localized at a given radius. We can extrapolate these single-atom results to a crystal and qualitatively derive the concepts of allowed and forbidden energy bands. We will then apply quantum mechanics and Schrodinger's wave equation to the problem of an electron in a single crystal. We find that the electronic energy states occur in bands of allowed states that are separated by forbidden energy bands.

3.1.1 Formation of Energy Bands

Figure 3.1a shows the radial probability density function for the lowest electron energy state of the single, noninteracting hydrogen atom, and Figure 3.1b shows the same probability curves for two atoms that are in close proximity to each other. The wave functions of the electrons of the two atoms overlap, which means that the two electrons will interact. This interaction or perturbation results in the discrete quantized energy level splitting into two discrete energy levels, schematically shown in Figure 3.1c. The splitting of the discrete state into two states is consistent with the Pauli exclusion principle.

A simple analogy of the splitting of energy levels by interacting particles is the following. Two identical race cars and drivers are far apart on a race track. There is no interaction between the cars, so they both must provide the same power to achieve a given speed. However, if one car pulls up close behind the other car, there is an interaction called *draft*. The second car will be pulled to an extent by the lead car. The lead car will therefore require more power to achieve the same speed since it is pulling the second car, and the second car will require less power since it is being

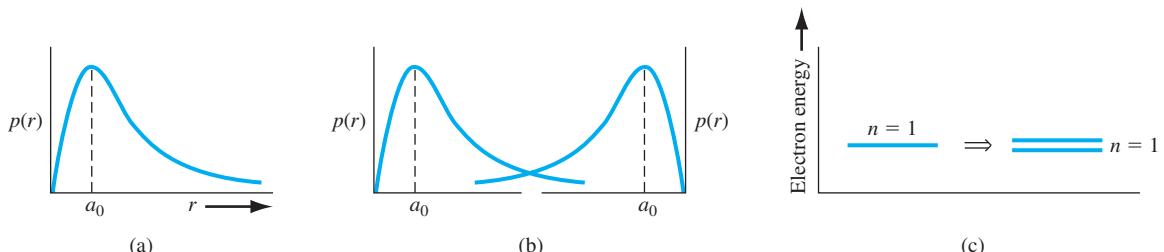


Figure 3.1 | (a) Probability density function of an isolated hydrogen atom. (b) Overlapping probability density functions of two adjacent hydrogen atoms. (c) The splitting of the $n = 1$ state.

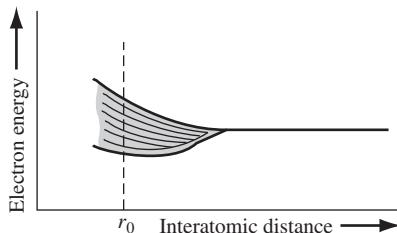


Figure 3.2 | The splitting of an energy state into a band of allowed energies.

pulled by the lead car. So there is a “splitting” of power (energy) of the two interacting race cars. (Keep in mind not to take analogies too literally.)

Now, if we somehow start with a regular periodic arrangement of hydrogen-type atoms that are initially very far apart, and begin pushing the atoms together, the initial quantized energy level will split into a band of discrete energy levels. This effect is shown schematically in Figure 3.2, where the parameter r_0 represents the equilibrium interatomic distance in the crystal. At the equilibrium interatomic distance, there is a band of allowed energies, but within the allowed band, the energies are at discrete levels. The Pauli exclusion principle states that the joining of atoms to form a system (crystal) does not alter the total number of quantum states regardless of size. However, since no two electrons can have the same quantum number, the discrete energy must split into a band of energies in order that each electron can occupy a distinct quantum state.

We have seen previously that, at any energy level, the number of allowed quantum states is relatively small. In order to accommodate all of the electrons in a crystal, we must have many energy levels within the allowed band. As an example, suppose that we have a system with 10^{19} one-electron atoms and also suppose that, at the equilibrium interatomic distance, the width of the allowed energy band is 1 eV. For simplicity, we assume that each electron in the system occupies a different energy level and, if the discrete energy states are equidistant, then the energy levels are separated by 10^{-19} eV. This energy difference is extremely small, so that for all practical purposes, we have a quasi-continuous energy distribution through the allowed energy band. The fact that 10^{-19} eV is a very small difference between two energy states can be seen from the following example.

EXAMPLE 3.1

Objective: Calculate the change in kinetic energy of an electron when the velocity changes by a small amount.

Consider an electron traveling at a velocity of 10^7 cm/s. Assume that the velocity increases by a value of 1 cm/s. The increase in kinetic energy is given by

$$\Delta E = \frac{1}{2}mv_2^2 - \frac{1}{2}mv_1^2 = \frac{1}{2}m(v_2^2 - v_1^2)$$

Let $v_2 = v_1 + \Delta v$. Then

$$v_2^2 = (v_1 + \Delta v)^2 = v_1^2 + 2v_1\Delta v + (\Delta v)^2$$

But $\Delta v \ll v_1$, so we have that

$$\Delta E \approx \frac{1}{2} m(2v_1\Delta v) = mv_1\Delta v$$

■ Solution

Substituting the number into this equation, we obtain

$$\Delta E = (9.11 \times 10^{-31})(10^5)(0.01) = 9.11 \times 10^{-28} \text{ J}$$

which may be converted to units of electron volts as

$$\Delta E = \frac{9.11 \times 10^{-28}}{1.6 \times 10^{-19}} = 5.7 \times 10^{-9} \text{ eV}$$

■ Comment

A change in velocity of 1 cm/s compared with 10^7 cm/s results in a change in energy of 5.7×10^{-9} eV, which is orders of magnitude larger than the change in energy of 10^{-19} eV between energy states in the allowed energy band. This example serves to demonstrate that a difference in adjacent energy states of 10^{-19} eV is indeed very small, so that the discrete energies within an allowed band may be treated as a quasi-continuous distribution.

■ EXERCISE PROBLEM

Ex 3.1 The initial velocity of an electron is 10^7 cm/s. If the kinetic energy of the electron increases by $\Delta E = 10^{-12}$ eV, determine the increase in velocity.
(s/mc₂) × 9.7 × 10⁻¹² = a/V

Consider again a regular periodic arrangement of atoms, in which each atom now contains more than one electron. Suppose the atom in this imaginary crystal contains electrons up through the $n = 3$ energy level. If the atoms are initially very far apart, the electrons in adjacent atoms will not interact and will occupy the discrete energy levels. If these atoms are brought closer together, the outermost electrons in the $n = 3$ energy shell will begin to interact initially, so that this discrete energy level will split into a band of allowed energies. If the atoms continue to move closer together, the electrons in the $n = 2$ shell may begin to interact and will also split into a band of allowed energies. Finally, if the atoms become sufficiently close together, the innermost electrons in the $n = 1$ level may interact, so that this energy level may also split into a band of allowed energies. The splitting of these discrete energy levels is qualitatively shown in Figure 3.3. If the equilibrium interatomic distance is r_0 , then we have bands of allowed energies that the electrons may occupy separated by bands of forbidden energies. This energy-band splitting and the formation of allowed and forbidden bands is the energy-band theory of single-crystal materials.

The actual band splitting in a crystal is much more complicated than indicated in Figure 3.3. A schematic representation of an isolated silicon atom is shown in Figure 3.4a. Ten of the 14 silicon atom electrons occupy deep-lying energy levels close to the nucleus. The four remaining valence electrons are relatively weakly bound and are the electrons involved in chemical reactions. Figure 3.4b shows the band splitting of silicon. We need only consider the $n = 3$ level for the valence

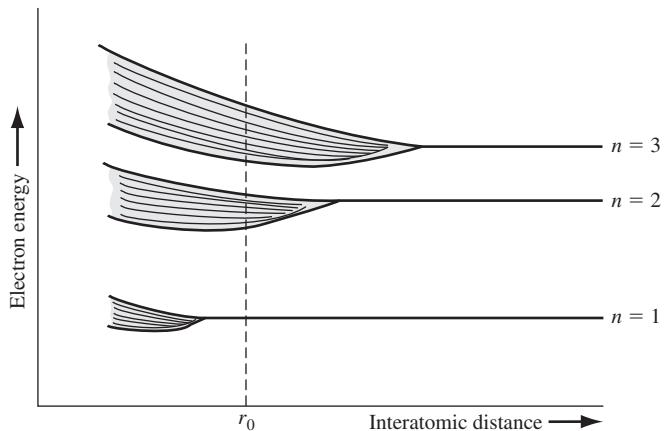


Figure 3.3 | Schematic showing the splitting of three energy states into allowed bands of energies.

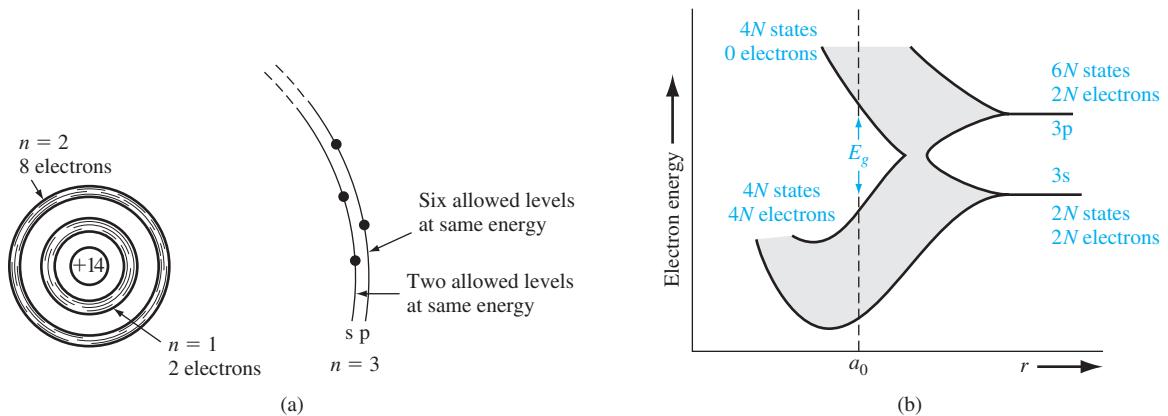


Figure 3.4 | (a) Schematic of an isolated silicon atom. (b) The splitting of the $3s$ and $3p$ states of silicon into the allowed and forbidden energy bands.

(From Shockley [6].)

electrons, since the first two energy shells are completely full and are tightly bound to the nucleus. The $3s$ state corresponds to $n = 3$ and $l = 0$ and contains two quantum states per atom. This state will contain two electrons at $T = 0$ K. The $3p$ state corresponds to $n = 3$ and $l = 1$ and contains six quantum states per atom. This state will contain the remaining two electrons in the individual silicon atom.

As the interatomic distance decreases, the $3s$ and $3p$ states interact and overlap. At the equilibrium interatomic distance, the bands have again split, but now four quantum states per atom are in the lower band and four quantum states per atom are in the upper band. At absolute zero degrees, electrons are in the lowest energy state, so that all states in the lower band (the valence band) will be full and all states in the

upper band (the conduction band) will be empty. The bandgap energy E_g between the top of the valence band and the bottom of the conduction band is the width of the forbidden energy band.

We have discussed qualitatively how and why bands of allowed and forbidden energies are formed in a crystal. The formation of these energy bands is directly related to the electrical characteristics of the crystal, as we will see later in our discussion.

*3.1.2 The Kronig–Penney Model¹

In the previous section, we discussed qualitatively the splitting of allowed electron energies as atoms are brought together to form a crystal. The concept of allowed and forbidden energy bands can be developed more rigorously by considering quantum mechanics and Schrodinger's wave equation. It may be easy for the reader to "get lost" in the following derivation, but the result forms the basis for the energy-band theory of semiconductors.

The potential function of a single, noninteracting, one-electron atom is shown in Figure 3.5a. Also indicated on the figure are the discrete energy levels allowed for the electron. Figure 3.5b shows the same type of potential function for the case when several atoms in close proximity are arranged in a one-dimensional array. The potential functions of adjacent atoms overlap, and the net potential function for this case is shown in Figure 3.5c. It is this potential function we would need to use in Schrodinger's wave equation to model a one-dimensional single-crystal material.

The solution to Schrodinger's wave equation, for this one-dimensional single-crystal lattice, is made more tractable by considering a simpler potential function. Figure 3.6 is the one-dimensional Kronig–Penney model of the periodic potential function, which is used to represent a one-dimensional single-crystal lattice. We need to solve Schrodinger's wave equation in each region. As with previous quantum mechanical problems, the more interesting solution occurs for the case when $E < V_0$, which corresponds to a particle being bound within the crystal. The electrons are contained in the potential wells, but we have the possibility of tunneling between wells. The Kronig–Penney model is an idealized periodic potential representing a one-dimensional single crystal, but the results will illustrate many of the important features of the quantum behavior of electrons in a periodic lattice.

To obtain the solution to Schrodinger's wave equation, we make use of a mathematical theorem by Bloch. The theorem states that all one-electron wave functions, for problems involving periodically varying potential energy functions, must be of the form

$$\psi(x) = u(x)e^{ikx} \quad (3.1)$$

^{*}Indicates sections that will aid in the total summation of understanding of semiconductor devices, but may be skipped the first time through the text without loss of continuity.

¹Other techniques, such as the nearly free electron model, can be used to predict the energy-band theory of semiconductor materials. See, for example, Kittel [3] or Wolfe et al. [14].

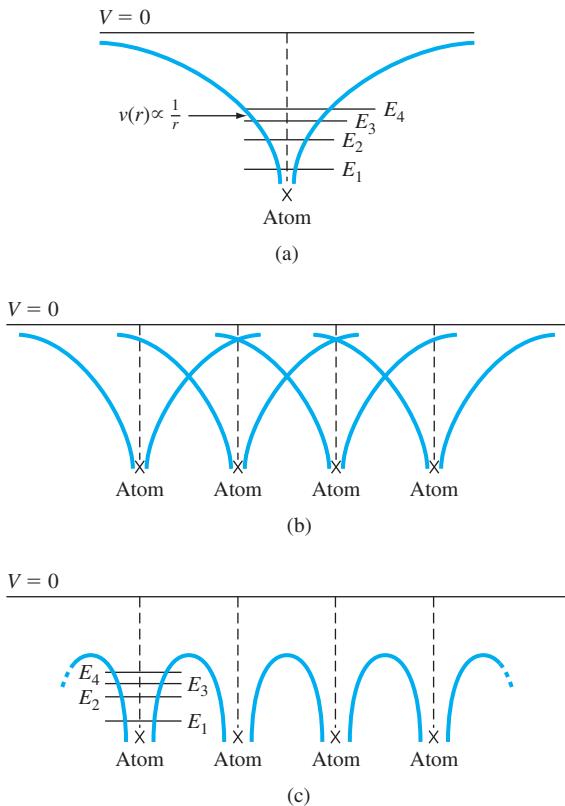


Figure 3.5 | (a) Potential function of a single isolated atom. (b) Overlapping potential functions of adjacent atoms. (c) Net potential function of a one-dimensional single crystal.

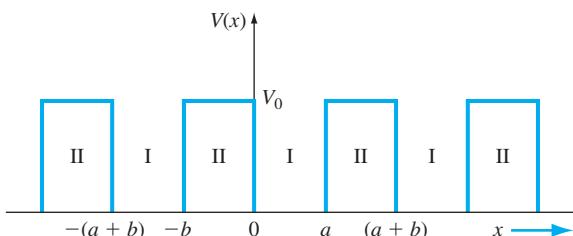


Figure 3.6 | The one-dimensional periodic potential function of the Kronig–Penney model.

The parameter k is called a constant of motion and will be considered in more detail as we develop the theory. The function $u(x)$ is a periodic function with period $(a + b)$.

We stated in Chapter 2 that the total solution to the wave equation is the product of the time-independent solution and the time-dependent solution, or

$$\Psi(x, t) = \psi(x)\phi(t) = u(x)e^{ikx} \cdot e^{-j(E/\hbar)t} \quad (3.2)$$

which may be written as

$$\Psi(x, t) = u(x)e^{j(kx - (E/\hbar)t)} \quad (3.3)$$

This traveling-wave solution represents the motion of an electron in a single-crystal material. The amplitude of the traveling wave is a periodic function and the parameter k is also referred to as a wave number.

We can now begin to determine a relation between the parameter k , the total energy E , and the potential V_0 . If we consider region I in Figure 3.6 ($0 < x < a$) in which $V(x) = 0$, take the second derivative of Equation (3.1), and substitute this result into the time-independent Schrodinger's wave equation given by Equation (2.13), we obtain the relation

$$\frac{d^2u_1(x)}{dx^2} + 2jk\frac{du_1(x)}{dx} - (k^2 - \alpha^2)u_1(x) = 0 \quad (3.4)$$

The function $u_1(x)$ is the amplitude of the wave function in region I and the parameter α is defined as

$$\alpha^2 = \frac{2mE}{\hbar^2} \quad (3.5)$$

Consider now a specific region II, $-b < x < 0$, in which $V(x) = V_0$, and apply Schrodinger's wave equation. We obtain the relation

$$\frac{d^2u_2(x)}{dx^2} + 2jk\frac{du_2(x)}{dx} - \left(k^2 - \alpha^2 + \frac{2mV_0}{\hbar^2}\right)u_2(x) = 0 \quad (3.6)$$

where $u_2(x)$ is the amplitude of the wave function in region II. We may define

$$\frac{2m}{\hbar^2}(E - V_0) = \alpha^2 - \frac{2mV_0}{\hbar^2} = \beta^2 \quad (3.7)$$

so that Equation (3.6) may be written as

$$\frac{d^2u_2(x)}{dx^2} + 2jk\frac{du_2(x)}{dx} - (k^2 - \beta^2)u_2(x) = 0 \quad (3.8)$$

Note that from Equation (3.7), if $E > V_0$, the parameter β is real, whereas if $E < V_0$, then β is imaginary.

The solution to Equation (3.4), for region I, is of the form

$$u_1(x) = Ae^{j(\alpha-k)x} + Be^{-j(\alpha+k)x} \quad \text{for } (0 < x < a) \quad (3.9)$$

and the solution to Equation (3.8), for region II, is of the form

$$u_2(x) = Ce^{j(\beta-k)x} + De^{-j(\beta+k)x} \quad \text{for } (-b < x < 0) \quad (3.10)$$

Since the potential function $V(x)$ is everywhere finite, both the wave function $\psi(x)$ and its first derivative $\partial\psi(x)/\partial x$ must be continuous. This continuity condition implies that the wave amplitude function $u(x)$ and its first derivative $\partial u(x)/\partial x$ must also be continuous.

If we consider the boundary at $x = 0$ and apply the continuity condition to the wave amplitude, we have

$$u_1(0) = u_2(0) \quad (3.11)$$

Substituting Equations (3.9) and (3.10) into Equation (3.11), we obtain

$$A + B - C - D = 0 \quad (3.12)$$

Now applying the condition that

$$\frac{du_1}{dx} \Big|_{x=0} = \frac{du_2}{dx} \Big|_{x=0} \quad (3.13)$$

we obtain

$$(\alpha - k)A - (\alpha + k)B - (\beta - k)C + (\beta + k)D = 0 \quad (3.14)$$

We have considered region I as $0 < x < a$ and region II as $-b < x < 0$. The periodicity and the continuity condition mean that the function u_1 , as $x \rightarrow a$, is equal to the function u_2 , as $x \rightarrow -b$. This condition may be written as

$$u_1(a) = u_2(-b) \quad (3.15)$$

Applying the solutions for $u_1(x)$ and $u_2(x)$ to the boundary condition in Equation (3.15) yields

$$Ae^{j(\alpha-k)a} + Be^{-j(\alpha+k)a} - Ce^{-j(\beta-k)b} - De^{j(\beta+k)b} = 0 \quad (3.16)$$

The last boundary condition is

$$\frac{du_1}{dx} \Big|_{x=a} = \frac{du_2}{dx} \Big|_{x=-b} \quad (3.17)$$

which gives

$$\begin{aligned} &(\alpha - k)Ae^{j(\alpha-k)a} - (\alpha + k)Be^{-j(\alpha+k)a} - (\beta - k)Ce^{-j(\beta-k)b} \\ &+ (\beta + k)De^{j(\beta+k)b} = 0 \end{aligned} \quad (3.18)$$

We now have four homogeneous equations, Equations (3.12), (3.14), (3.16), and (3.18), with four unknowns as a result of applying the four boundary conditions. In a set of simultaneous, linear, homogeneous equations, there is a nontrivial solution if, and only if, the determinant of the coefficients is zero. In our case, the coefficients in question are the coefficients of the parameters A , B , C , and D .

The evaluation of this determinant is extremely laborious and will not be considered in detail. The result is

$$\frac{-(\alpha^2 + \beta^2)}{2\alpha\beta} (\sin \alpha a)(\sin \beta b) + (\cos \alpha a)(\cos \beta b) = \cos k(a + b) \quad (3.19)$$

Equation (3.19) relates the parameter k to the total energy E (through the parameter α) and the potential function V_0 (through the parameter β).

As we mentioned, the more interesting solutions occur for $E < V_0$, which applies to the electron bound within the crystal. From Equation (3.7), the parameter β is then an imaginary quantity. We may define

$$\beta = j\gamma \quad (3.20)$$

where γ is a real quantity. Equation (3.19) can be written in terms of γ as

$$\frac{\gamma^2 - \alpha^2}{2\alpha\gamma} (\sin \alpha a)(\sinh \gamma b) + (\cos \alpha a)(\cosh \gamma b) = \cos k(a + b) \quad (3.21)$$

Equation (3.21) does not lend itself to an analytical solution, but must be solved using numerical or graphical techniques to obtain the relation between k , E , and V_0 . The solution of Schrodinger's wave equation for a single bound particle resulted in discrete allowed energies. The solution of Equation (3.21) will result in a band of allowed energies.

To obtain an equation that is more susceptible to a graphical solution and thus will illustrate the nature of the results, let the potential barrier width $b \rightarrow 0$ and the barrier height $V_0 \rightarrow \infty$, but such that the product bV_0 remains finite. Equation (3.21) then reduces to

$$\left(\frac{mV_0ba}{\hbar^2} \right) \frac{\sin \alpha a}{\alpha a} + \cos \alpha a = \cos ka \quad (3.22)$$

We may define a parameter P' as

$$P' = \frac{mV_0ba}{\hbar^2} \quad (3.23)$$

Then, finally, we have the relation

$$P' \frac{\sin \alpha a}{\alpha a} + \cos \alpha a = \cos ka \quad (3.24)$$

Equation (3.24) again gives the relation between the parameter k , total energy E (through the parameter α), and the potential barrier bV_0 . We may note that Equation (3.24) is not a solution of Schrodinger's wave equation but gives the conditions for which Schrodinger's wave equation will have a solution. If we assume that the crystal is infinitely large, then k in Equation (3.24) can assume a continuum of values and must be real.

3.1.3 The k -Space Diagram

To begin to understand the nature of the solution, initially consider the special case for which $V_0 = 0$. In this case $P' = 0$, which corresponds to a free particle since there are no potential barriers. From Equation (3.24), we have that

$$\cos \alpha a = \cos ka \quad (3.25)$$

or

$$\alpha = k \quad (3.26)$$

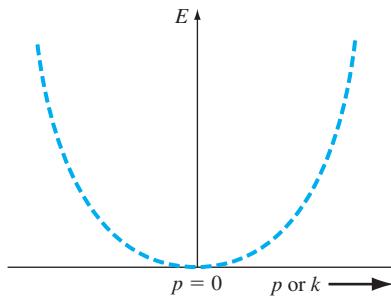


Figure 3.7 | The parabolic E versus k curve for the free electron.

Since the potential is equal to zero, the total energy E is equal to the kinetic energy, so that, from Equation (3.5), Equation (3.26) may be written as

$$\alpha = \sqrt{\frac{2mE}{\hbar^2}} = \sqrt{\frac{2m(\frac{1}{2}mv^2)}{\hbar^2}} = \frac{p}{\hbar} = k \quad (3.27)$$

where p is the particle momentum. The constant of the motion parameter k is related to the particle momentum for the free electron. The parameter k is also referred to as a wave number.

We can also relate the energy and momentum as

$$E = \frac{p^2}{2m} = \frac{k^2\hbar^2}{2m} \quad (3.28)$$

Figure 3.7 shows the parabolic relation of Equation (3.28) between the energy E and momentum p for the free particle. Since the momentum and wave number are linearly related, Figure 3.7 is also the E versus k curve for the free particle.

We now want to consider the relation between E and k from Equation (3.24) for the particle in the single-crystal lattice. As the parameter P' increases, the particle becomes more tightly bound to the potential well or atom. We may define the left side of Equation (3.24) to be a function $f(\alpha a)$, so that

$$f(\alpha a) = P' \frac{\sin \alpha a}{\alpha a} + \cos \alpha a \quad (3.29)$$

Figure 3.8a is a plot of the first term of Equation (3.29) versus αa . Figure 3.8b shows a plot of the $\cos \alpha a$ term and Figure 3.8c is the sum of the two terms, or $f(\alpha a)$.

Now from Equation (3.24), we also have that

$$f(\alpha a) = \cos ka \quad (3.30)$$

For Equation (3.30) to be valid, the allowed values of the $f(\alpha a)$ function must be bounded between $+1$ and -1 . Figure 3.8c shows the allowed values of $f(\alpha a)$ and the allowed values of αa in the shaded areas. Also shown on the figure are the values of ka from the right side of Equation (3.30), which correspond to the allowed values of $f(\alpha a)$.

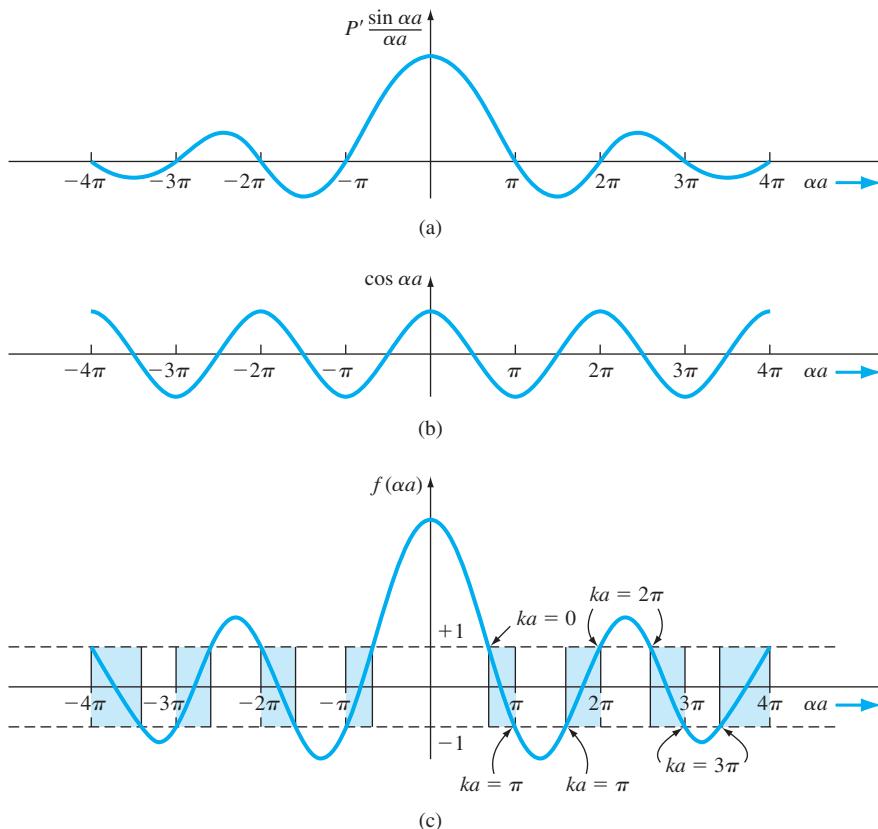


Figure 3.8 | A plot of (a) the first term in Equation (3.29), (b) the second term in Equation (3.29), and (c) the entire $f(\alpha a)$ function. The shaded areas show the allowed values of (αa) corresponding to real values of k .

The parameter α is related to the total energy E of the particle through Equation (3.5), which is $\alpha^2 = 2mE/\hbar^2$. A plot of the energy E of the particle as a function of the wave number k can be generated from Figure 3.8c. Figure 3.9 shows this plot and thus shows the concept of allowed energy bands for the particle propagating in the crystal lattice. Since the energy E has discontinuities, we also have the concept of forbidden energies for the particles in the crystal.

Objective: Determine the width (in eV) of a forbidden energy band.

Determine the width of the forbidden bandgap that exists at $ka = \pi$ (see Figure 3.9). Assume that the coefficient $P' = 8$ and the potential width is $a = 4.5 \text{ \AA}$.

EXAMPLE 3.2

■ Solution

Combining Equations (3.29) and (3.30), we have

$$\cos ka = P' \frac{\sin \alpha a}{\alpha a} + \cos \alpha a$$

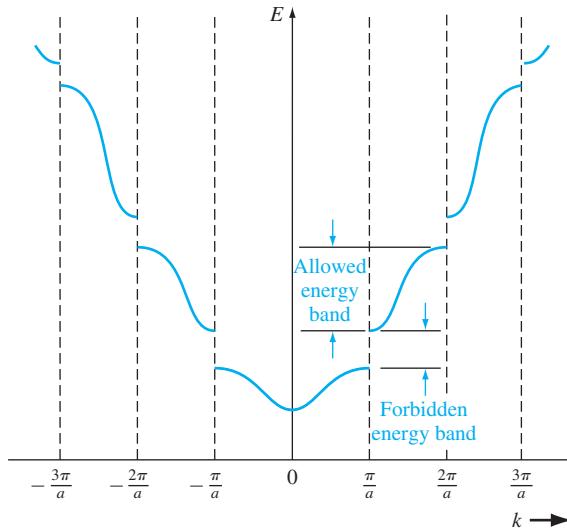


Figure 3.9 | The E versus k diagram generated from Figure 3.8. The allowed energy bands and forbidden energy bandgaps are indicated.

At $ka = \pi$ and using $P' = 8$, we have

$$-1 = 8 \frac{\sin \alpha a}{\alpha a} + \cos \alpha a$$

We need to find the smallest values of αa that satisfy this equation and then relate α to the energy E to find the bandgap energy. From Figure 3.8, we see that, at one value of $ka = \pi$, we have $\alpha a = \pi \equiv \alpha_1 a$. Then

$$\alpha_1 a = \sqrt{\frac{2mE_1}{\hbar^2}} \cdot a = \pi$$

or

$$E_1 = \frac{\pi^2 \hbar^2}{2ma^2} = \frac{\pi^2 (1.054 \times 10^{-34})^2}{2(9.11 \times 10^{-31})(4.5 \times 10^{-10})^2} = 2.972 \times 10^{-19} \text{ J}$$

From Figure 3.8, we see that, at the other value of $ka = \pi$, αa is in the range $\pi < \alpha a < 2\pi$. By trial and error, we find $\alpha a = 5.141 \equiv \alpha_2 a$. Then

$$\alpha_2 a = \sqrt{\frac{2mE_2}{\hbar^2}} \cdot a = 5.141$$

or

$$E_2 = \frac{(5.141)^2 \hbar^2}{2ma^2} = \frac{(5.141)^2 (1.054 \times 10^{-34})^2}{2(9.11 \times 10^{-31})(4.5 \times 10^{-10})^2} = 7.958 \times 10^{-19} \text{ J}$$

The bandgap energy is then

$$E_g = E_2 - E_1 = 7.958 \times 10^{-19} - 2.972 \times 10^{-19} = 4.986 \times 10^{-19} \text{ J}$$

or

$$E_g = \frac{4.986 \times 10^{-19}}{1.6 \times 10^{-19}} = 3.12 \text{ eV}$$

Comment

The results of this example give an order of magnitude of forbidden energy band widths.

EXERCISE PROBLEM

Ex 3.2 Using the parameters given in Example 3.2, determine the width of the allowed energy band in the range $\pi < ka < 2\pi$. ($\Delta E = 24 \text{ eV}$)

Consider again the right side of Equation (3.24), which is the function $\cos ka$. The cosine function is periodic so that

$$\cos ka = \cos(ka + 2n\pi) = \cos(ka - 2n\pi) \quad (3.31)$$

where n is a positive integer. We may consider Figure 3.9 and displace portions of the curve by 2π . Mathematically, Equation (3.24) is still satisfied. Figure 3.10 shows how various segments of the curve can be displaced by the 2π factor. Figure 3.11 shows the case in which the entire E versus k plot is contained within $-\pi/a < k < \pi/a$. This plot is referred to as a reduced k -space diagram, or a reduced-zone representation.

We noted in Equation (3.27) that for a free electron, the particle momentum and the wave number k are related by $p = \hbar k$. Given the similarity between the free

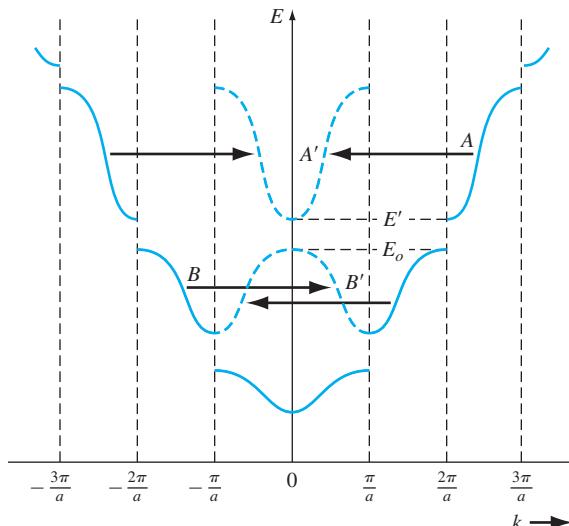


Figure 3.10 | The E versus k diagram showing 2π displacements of several sections of allowed energy bands.

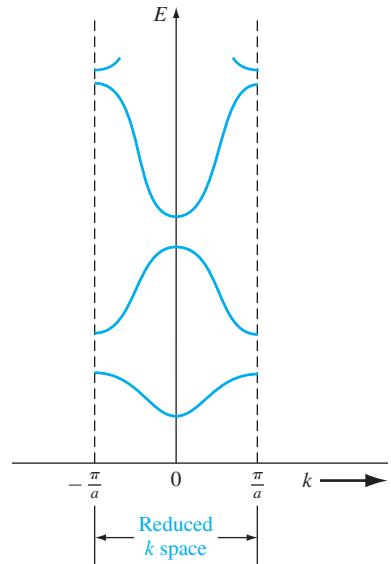


Figure 3.11 | The E versus k diagram in the reduced-zone representation.

electron solution and the results of the single crystal shown in Figure 3.9, the parameter $\hbar k$ in a single crystal is referred to as the *crystal momentum*. This parameter is not the actual momentum of the electron in the crystal but is a constant of the motion that includes the crystal interaction.

We have been considering the Kronig–Penney model, which is a one-dimensional periodic potential function used to model a single-crystal lattice. The principal result of this analysis, so far, is that electrons in the crystal occupy certain allowed energy bands and are excluded from the forbidden energy bands. For real three-dimensional single-crystal materials, a similar energy-band theory exists. We will obtain additional electron properties from the Kronig–Penney model in the next sections.

TEST YOUR UNDERSTANDING

- TYU 3.1** Using the parameters given in Example 3.2, determine the width (in eV) of the second forbidden energy band existing at $ka = 2\pi$ (see Figure 3.8(c)).
 (Ans. $E = 4.23 \text{ eV}$)
- TYU 3.2** Using the parameters given in Example 3.2, determine the width (in eV) of the allowed energy band in the range $0 < ka < \pi$ (see Figure 3.8(c)).
 (Ans. $E = 0.654 \text{ eV}$)

3.2 | ELECTRICAL CONDUCTION IN SOLIDS

Again, we are eventually interested in determining the current–voltage characteristics of semiconductor devices. We will need to consider electrical conduction in solids as it relates to the band theory we have just developed. Let us begin by considering the motion of electrons in the various allowed energy bands.

3.2.1 The Energy Band and the Bond Model

In Chapter 1, we discussed the covalent bonding of silicon. Figure 3.12 shows a two-dimensional representation of the covalent bonding in a single-crystal silicon lattice. This figure represents silicon at $T = 0 \text{ K}$ in which each silicon atom is surrounded by eight valence electrons that are in their lowest energy state and are directly involved in the covalent bonding. Figure 3.4b represented the splitting of the discrete silicon energy states into bands of allowed energies as the silicon crystal is formed. At $T = 0 \text{ K}$, the $4N$ states in the lower band, the valence band, are filled with the valence electrons. All of the valence electrons schematically shown in Figure 3.12 are in the valence band. The upper energy band, the conduction band, is completely empty at $T = 0 \text{ K}$.

As the temperature increases above 0 K , a few valence band electrons may gain enough thermal energy to break the covalent bond and jump into the conduction band. Figure 3.13a shows a two-dimensional representation of this bond-breaking

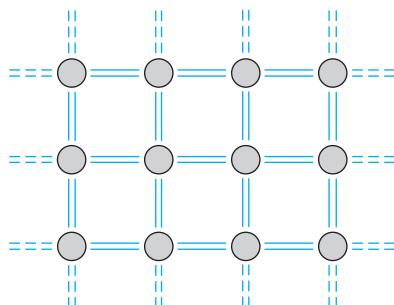


Figure 3.12 | Two-dimensional representation of the covalent bonding in a semiconductor at $T = 0\text{ K}$.

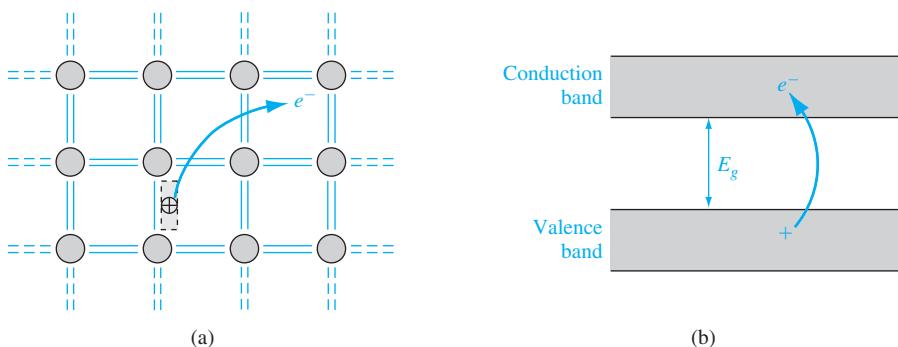


Figure 3.13 | (a) Two-dimensional representation of the breaking of a covalent bond. (b) Corresponding line representation of the energy band and the generation of a negative and positive charge with the breaking of a covalent bond.

effect and Figure 3.13b, a simple line representation of the energy-band model, shows the same effect.

The semiconductor is neutrally charged. This means that, as the negatively charged electron breaks away from its covalent bonding position, a positively charged “empty state” is created in the original covalent bonding position in the valence band. As the temperature further increases, more covalent bonds are broken, more electrons jump to the conduction band, and more positive “empty states” are created in the valence band.

We can also relate this bond breaking to the E versus k energy bands. Figure 3.14a shows the E versus k diagram of the conduction and valence bands at $T = 0\text{ K}$. The energy states in the valence band are completely full and the states in the conduction band are empty. Figure 3.14b shows these same bands for $T > 0\text{ K}$, in which some electrons have gained enough energy to jump to the conduction band and have left empty states in the valence band. We are assuming at this point that no external forces are applied so the electron and “empty state” distributions are symmetrical with k .

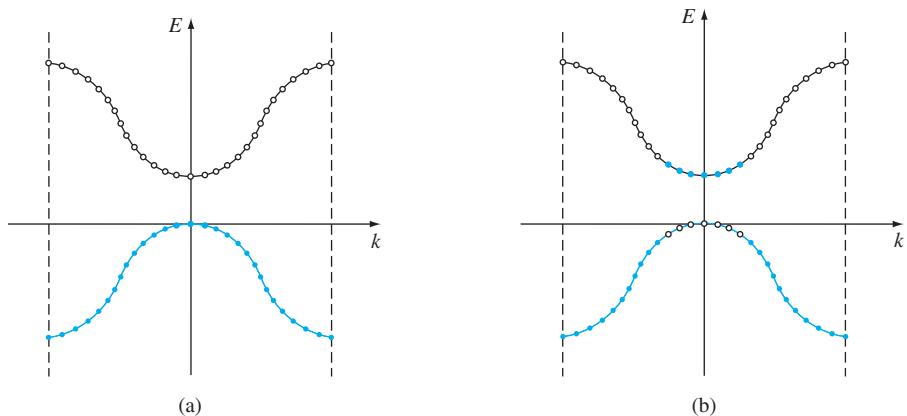


Figure 3.14 | The E versus k diagram of the conduction and valence bands of a semiconductor at (a) $T = 0$ K and (b) $T > 0$ K.

3.2.2 Drift Current

Current is due to the net flow of charge. If we had a collection of positively charged ions with a volume density N (cm^{-3}) and an average drift velocity v_d (cm/s), then the drift current density would be

$$J = qNv_d \quad \text{A/cm}^2 \quad (3.32)$$

If, instead of considering the average drift velocity, we considered the individual ion velocities, then we could write the drift current density as

$$J = q \sum_{i=1}^N v_i \quad (3.33)$$

where v_i is the velocity of the i th ion. The summation in Equation (3.33) is taken over a unit volume so that the current density J is still in units of A/cm^2 .

Since electrons are charged particles, a net drift of electrons in the conduction band will give rise to a current. The electron distribution in the conduction band, as shown in Figure 3.14b, is an even function of k when no external force is applied. Recall that k for a free electron is related to momentum so that, since there are as many electrons with a $+|k|$ value as there are with a $-|k|$ value, the net drift current density due to these electrons is zero. This result is certainly expected since there is no externally applied force.

If a force is applied to a particle and the particle moves, it must gain energy. This effect is expressed as

$$dE = F dx = Fv dt \quad (3.34)$$

where F is the applied force, dx is the differential distance the particle moves, v is the velocity, and dE is the increase in energy. If an external force is applied to the electrons in the conduction band, there are empty energy states into which the electrons can move; therefore, because of the external force, electrons can gain energy and a

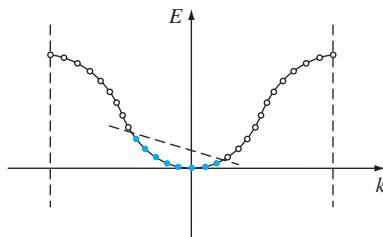


Figure 3.15 | The asymmetric distribution of electrons in the E versus k diagram when an external force is applied.

net momentum. The electron distribution in the conduction band may look like that shown in Figure 3.15, which implies that the electrons have gained a net momentum.

We may write the drift current density due to the motion of electrons as

$$J = -e \sum_{i=1}^n v_i \quad (3.35)$$

where e is the magnitude of the electronic charge and n is the number of electrons per unit volume in the conduction band. Again, the summation is taken over a unit volume so that the current density is still in units of A/cm^2 . We may note from Equation (3.35) that the current is directly related to the electron velocity; that is, the current is related to how well the electron can move in the crystal.

3.2.3 Electron Effective Mass

The movement of an electron in a lattice will, in general, be different from that of an electron in free space. In addition to an externally applied force, there are internal forces in the crystal due to positively charged ions or protons and negatively charged electrons, which will influence the motion of electrons in the lattice. We can write

$$F_{\text{total}} = F_{\text{ext}} + F_{\text{int}} = ma \quad (3.36)$$

where F_{total} , F_{ext} , and F_{int} are the total force, the externally applied force, and the internal forces, respectively, acting on a particle in a crystal. The parameter a is the acceleration and m is the rest mass of the particle.

Since it is difficult to take into account all of the internal forces, we will write the equation

$$F_{\text{ext}} = m^*a \quad (3.37)$$

where the acceleration a is now directly related to the external force. The parameter m^* , called the *effective mass*, takes into account the particle mass and also takes into account the effect of the internal forces.

To use an analogy for the effective mass concept, consider the difference in motion between a glass marble in a container filled with water and in a container filled with oil. In general, the marble will drop through the water at a faster rate than through the oil. The external force in this example is the gravitational force and the internal

forces are related to the viscosity of the liquids. Because of the difference in motion of the marble in these two cases, the mass of the marble would appear to be different in water than in oil. (As with any analogy, we must be careful not to be too literal.)

We can also relate the effective mass of an electron in a crystal to the E versus k curves, such as is shown in Figure 3.11. In a semiconductor material, we will be dealing with allowed energy bands that are almost empty of electrons and other energy bands that are almost full of electrons.

To begin, consider the case of a free electron whose E versus k curve is shown in Figure 3.7. Recalling Equation (3.28), the energy and momentum are related by $E = p^2/2m = \hbar^2k^2/2m$, where m is the mass of the electron. The momentum and wave number k are related by $p = \hbar k$. If we take the derivative of Equation (3.28) with respect to k , we obtain

$$\frac{dE}{dk} = \frac{\hbar^2 k}{m} = \frac{\hbar p}{m} \quad (3.38)$$

Relating momentum to velocity, Equation (3.38) can be written as

$$\frac{1}{\hbar} \frac{dE}{dk} = \frac{p}{m} = v \quad (3.39)$$

where v is the velocity of the particle. The first derivative of E with respect to k is related to the velocity of the particle.

If we now take the second derivative of E with respect to k , we have

$$\frac{d^2E}{dk^2} = \frac{\hbar^2}{m} \quad (3.40)$$

We may rewrite Equation (3.40) as

$$\boxed{\frac{1}{\hbar^2} \frac{d^2E}{dk^2} = \frac{1}{m}} \quad (3.41)$$

The second derivative of E with respect to k is inversely proportional to the mass of the particle. For the case of a free electron, the mass is a constant (nonrelativistic effect), so the second derivative function is a constant. We may also note from Figure 3.7 that d^2E/dk^2 is a positive quantity, which implies that the mass of the electron is also a positive quantity.

If we apply an electric field to the free electron and use Newton's classical equation of motion, we can write

$$F = ma = -eE \quad (3.42)$$

where a is the acceleration, E is the applied electric field, and e is the magnitude of the electronic charge. Solving for the acceleration, we have

$$a = \frac{-eE}{m} \quad (3.43)$$

The motion of the free electron is in the opposite direction to the applied electric field because of the negative charge.

We may now apply the results to the electron in the bottom of an allowed energy band. Consider the allowed energy band in Figure 3.16a. The energy near the bottom

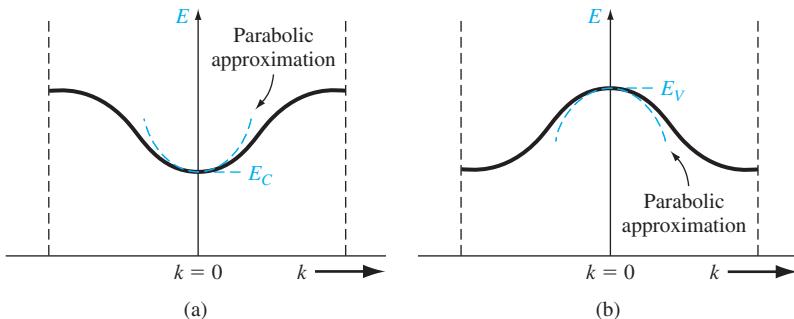


Figure 3.16 | (a) The conduction band in reduced k space, and the parabolic approximation. (b) The valence band in reduced k space, and the parabolic approximation.

of this energy band may be approximated by a parabola, just as that of a free particle. We may write

$$E - E_c = C_1(k)^2 \quad (3.44)$$

The energy E_c is the energy at the bottom of the band. Since $E > E_c$, the parameter C_1 is a positive quantity.

Taking the second derivative of E with respect to k from Equation (3.44), we obtain

$$\frac{d^2E}{dk^2} = 2C_1 \quad (3.45)$$

We may put Equation (3.45) in the form

$$\frac{1}{\hbar^2} \frac{d^2E}{dk^2} = \frac{2C_1}{\hbar^2} \quad (3.46)$$

Comparing Equation (3.46) with Equation (3.41), we may equate $\hbar^2/2C_1$ to the mass of the particle. However, the curvature of the curve in Figure 3.16a will not, in general, be the same as the curvature of the free-particle curve. We may write

$$\frac{1}{\hbar^2} \frac{d^2E}{dk^2} = \frac{2C_1}{\hbar^2} = \frac{1}{m^*} \quad (3.47)$$

where m^* is called the effective mass. Since $C_1 > 0$, we have that $m^* > 0$ also.

The effective mass is a parameter that relates the quantum mechanical results to the classical force equations. In most instances, the electron in the bottom of the conduction band can be thought of as a classical particle whose motion can be modeled by Newtonian mechanics, provided that the internal forces and quantum mechanical properties are taken into account through the effective mass. If we apply an electric field to the electron in the bottom of the allowed energy band, we may write the acceleration as

$$a = \frac{-eE}{m_n^*} \quad (3.48)$$

where m_n^* is the effective mass of the electron. The effective mass m_n^* of the electron near the bottom of the conduction band is a constant.

3.2.4 Concept of the Hole

In considering the two-dimensional representation of the covalent bonding shown in Figure 3.13a, a positively charged “empty state” was created when a valence electron was elevated into the conduction band. For $T > 0$ K, all valence electrons may gain thermal energy; if a valence electron gains a small amount of thermal energy, it may hop into the empty state. The movement of a valence electron into the empty state is equivalent to the movement of the positively charged empty state itself. Figure 3.17 shows the movement of valence electrons in the crystal, alternately filling one empty state and creating a new empty state—a motion equivalent to a positive charge moving in the valence band. The crystal now has a second equally important charge carrier that can give rise to a current. This charge carrier is called a *hole* and, as we will see, can also be thought of as a classical particle whose motion can be modeled using Newtonian mechanics.

The drift current density due to electrons in the valence band, such as shown in Figure 3.14b, can be written as

$$J = -e \sum_{i \text{ (filled)}} v_i \quad (3.49)$$

where the summation extends over all filled states. This summation is inconvenient since it extends over a nearly full valence band and takes into account a very large number of states. We may rewrite Equation (3.49) in the form

$$J = -e \sum_{i \text{ (total)}} v_i + e \sum_{i \text{ (empty)}} v_i \quad (3.50)$$

If we consider a band that is totally full, all available states are occupied by electrons. The individual electrons can be thought of as moving with a velocity as given by Equation (3.39):

$$v(E) = \left(\frac{1}{\hbar}\right) \left(\frac{dE}{dk}\right) \quad (3.39)$$

The band is symmetric in k and each state is occupied so that, for every electron with a velocity $|v|$, there is a corresponding electron with a velocity $-|v|$. Since the band is full, the distribution of electrons with respect to k cannot be changed with an

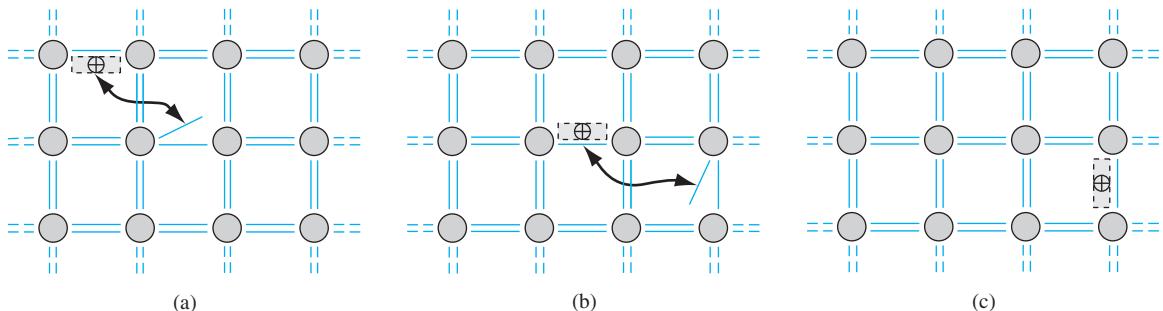


Figure 3.17 | Visualization of the movement of a hole in a semiconductor.

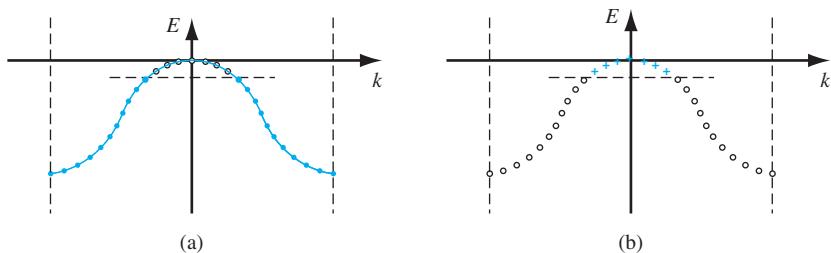


Figure 3.18 | (a) Valence band with conventional electron-filled states and empty states.
(b) Concept of positive charges occupying the original empty states.

externally applied force. The net drift current density generated from a completely full band, then, is zero, or

$$-e \sum_{i(\text{total})} v_i \equiv 0 \quad (3.51)$$

We can now write the drift current density from Equation (3.50) for an almost full band as

$$J = +e \sum_{i(\text{empty})} v_i \quad (3.52)$$

where the v_i in the summation is the

$$v(E) = \left(\frac{1}{\hbar} \right) \left(\frac{dE}{dk} \right)$$

associated with the empty state. Equation (3.52) is entirely equivalent to placing a positively charged particle in the empty states and assuming all other states in the band are empty, or neutrally charged. This concept is shown in Figure 3.18. Figure 3.18a shows the valence band with the conventional electron-filled states and empty states, whereas Figure 3.18b shows the new concept of positive charges occupying the original empty states. This concept is consistent with the discussion of the positively charged “empty state” in the valence band, as shown in Figure 3.17.

The v_i in the summation of Equation (3.52) is related to how well this positively charged particle moves in the semiconductor. Now consider an electron near the top of the allowed energy band shown in Figure 3.16b. The energy near the top of the allowed energy band may again be approximated by a parabola so that we may write

$$(E - E_v) = -C_2(k)^2 \quad (3.53)$$

The energy E_v is the energy at the top of the energy band. Since $E < E_v$ for electrons in this band, the parameter C_2 must be a positive quantity.

Taking the second derivative of E with respect to k from Equation (3.53), we obtain

$$\frac{d^2E}{dk^2} = -2C_2 \quad (3.54)$$

We may rearrange this equation so that

$$\frac{1}{\hbar^2} \frac{d^2E}{dk^2} = \frac{-2C_2}{\hbar^2} \quad (3.55)$$

Comparing Equation (3.55) with Equation (3.41), we may write

$$\frac{1}{\hbar^2} \frac{d^2E}{dk^2} = \frac{-2C_2}{\hbar^2} = \frac{1}{m^*} \quad (3.56)$$

where m^* is again an effective mass. We have argued that C_2 is a positive quantity, which now implies that m^* is a negative quantity. An electron moving near the top of an allowed energy band behaves as if it has a negative mass.

We must keep in mind that the effective mass parameter is used to relate quantum mechanics and classical mechanics. The attempt to relate these two theories leads to this strange result of a negative effective mass. However, we must recall that solutions to Schrodinger's wave equation also led to results that contradicted classical mechanics. The negative effective mass is another such example.

In discussing the concept of effective mass in the previous section, we used an analogy of marbles moving through two liquids. Now consider placing an ice cube in the center of a container filled with water: the ice cube will move upward toward the surface in a direction opposite to the gravitational force. The ice cube appears to have a negative effective mass since its acceleration is opposite to the external force. The effective mass parameter takes into account all internal forces acting on the particle.

If we again consider an electron near the top of an allowed energy band and use Newton's force equation for an applied electric field, we will have

$$F = m^*a = -eE \quad (3.57)$$

However, m^* is now a negative quantity, so we may write

$$a = \frac{-eE}{-|m^*|} = \frac{+eE}{|m^*|} \quad (3.58)$$

An electron moving near the top of an allowed energy band moves in the same direction as the applied electric field.

The net motion of electrons in a nearly full band can be described by considering just the empty states, provided that a positive electronic charge is associated with each state and that the negative of m^* from Equation (3.56) is associated with each state. We now can model this band as having particles with a positive electronic charge and a positive effective mass. The density of these particles in the valence band is the same as the density of empty electronic energy states. This new particle is the *hole*. The hole, then, has a positive effective mass denoted by m_p^* and a positive electronic charge, so it will move in the same direction as an applied field.

3.2.5 Metals, Insulators, and Semiconductors

Each crystal has its own energy-band structure. We noted that the splitting of the energy states in silicon, for example, to form the valence and conduction bands, was complex. Complex band splitting occurs in other crystals, leading to large variations in band structures between various solids and to a wide range of electrical characteristics observed in these various materials. We can qualitatively begin to understand some basic differences in electrical characteristics caused by variations in band structure by considering some simplified energy bands.

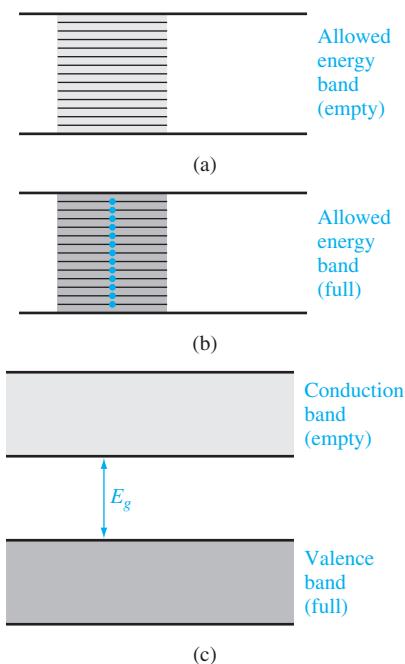


Figure 3.19 | Allowed energy bands showing (a) an empty band, (b) a completely full band, and (c) the bandgap energy between the two allowed bands.

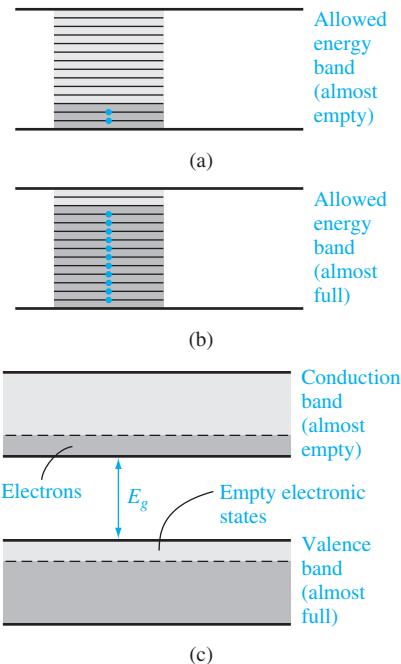


Figure 3.20 | Allowed energy bands showing (a) an almost empty band, (b) an almost full band, and (c) the bandgap energy between the two allowed bands.

There are several possible energy-band conditions to consider. Figure 3.19a shows an allowed energy band that is completely empty of electrons. If an electric field is applied, there are no particles to move, so there will be no current. Figure 3.19b shows another allowed energy band whose energy states are completely full of electrons. We argued in the previous section that a completely full energy band will also not give rise to a current. A material that has energy bands either completely empty or completely full is an insulator. The resistivity of an insulator is very large or, conversely, the conductivity of an insulator is very small. There are essentially no charged particles that can contribute to a drift current. Figure 3.19c shows a simplified energy-band diagram of an insulator. The bandgap energy E_g of an insulator is usually on the order of 3.5 to 6 eV or larger, so that at room temperature, there are essentially no electrons in the conduction band and the valence band remains completely full. There are very few thermally generated electrons and holes in an insulator.

Figure 3.20a shows an energy band with relatively few electrons near the bottom of the band. Now, if an electric field is applied, the electrons can gain energy, move to higher energy states, and move through the crystal. The net flow of charge is a current. Figure 3.20b shows an allowed energy band that is almost full of electrons, which means that we can consider the holes in this band. If an electric field is applied, the holes can move and give rise to a current. Figure 3.20c shows the simplified energy-band diagram for this case. The bandgap energy may be on the order of 1 eV.

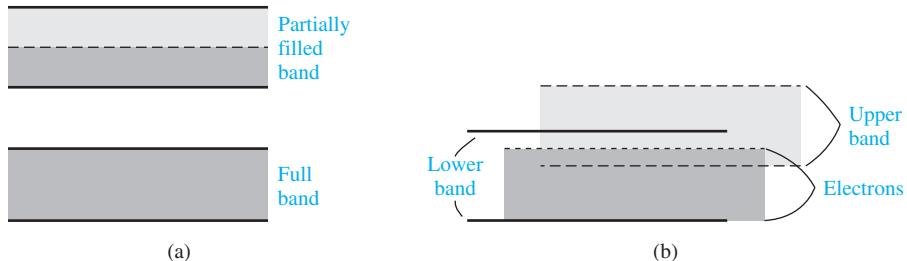


Figure 3.21 | Two possible energy bands of a metal showing (a) a partially filled band and (b) overlapping allowed energy bands.

This energy-band diagram represents a semiconductor for $T > 0$ K. The resistivity of a semiconductor, as we will see in the next chapter, can be controlled and varied over many orders of magnitude.

The characteristics of a metal include a very low resistivity. The energy-band diagram for a metal may be in one of two forms. Figure 3.21a shows the case of a partially full band in which there are many electrons available for conduction, so that the material can exhibit a large electrical conductivity. Figure 3.21b shows another possible energy-band diagram of a metal. The band splitting into allowed and forbidden energy bands is a complex phenomenon, and Figure 3.21b shows a case in which the conduction and valence bands overlap at the equilibrium interatomic distance. As in the case shown in Figure 3.21a, there are large numbers of electrons as well as large numbers of empty energy states into which the electrons can move, so this material can also exhibit a very high electrical conductivity.

TEST YOUR UNDERSTANDING

TYU 3.3 A simplified E versus k curve for an electron in the conduction band is given. The value of a is 10 Å. Determine the relative effective mass m^*/m_0 .
 $(\text{Ans: } 1.17)$

TYU 3.4 A simplified E versus k curve for a hole in the valence band is given. Assume a value of $a = 12$ Å. Determine the relative effective mass $|m^*/m_0|$.
 $(\text{Ans: } 0.29)$

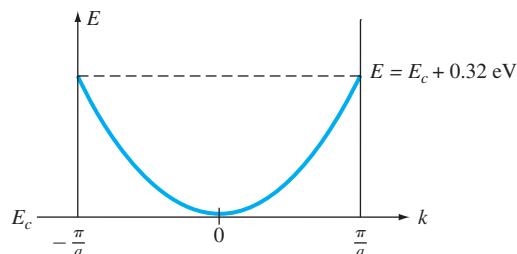


Figure 3.22 | Figure for Exercise TYU 3.3.

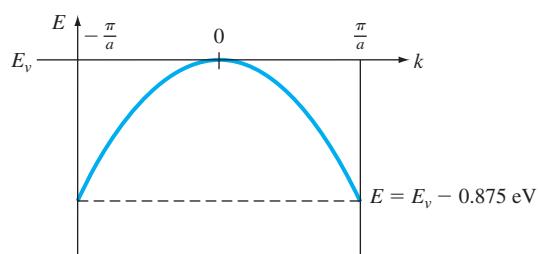


Figure 3.23 | Figure for Exercise TYU 3.4.

3.3 | EXTENSION TO THREE DIMENSIONS

The basic concepts of allowed and forbidden energy bands and effective mass have been developed in the previous sections. In this section, we will extend these concepts to three dimensions and to real crystals. We will qualitatively consider particular characteristics of the three-dimensional crystal in terms of the E versus k plots, bandgap energy, and effective mass. We must emphasize that we will only briefly touch on the basic three-dimensional concepts; therefore, many details will not be considered.

One problem encountered in extending the potential function to a three-dimensional crystal is that the distance between atoms varies as the direction through the crystal changes. Figure 3.24 shows a face-centered cubic structure with the [100] and [110] directions indicated. Electrons traveling in different directions encounter different potential patterns and therefore different k -space boundaries. The E versus k diagrams are, in general, a function of the k -space direction in a crystal.

3.3.1 The k -Space Diagrams of Si and GaAs

Figure 3.25 shows an E versus k diagram of gallium arsenide (GaAs) and of silicon (Si). These simplified diagrams show the basic properties considered in this text but do not show many of the details more appropriate for advanced-level courses.

Note that in place of the usual positive and negative k axes, we now show two different crystal directions. The E versus k diagram for the one-dimensional model was symmetric in k so that no new information is obtained by displaying the negative axis. It is normal practice to plot the [100] direction along the normal $+k$ axis and to plot the [111] portion of the diagram so the $+k$ points to the left. In the case of diamond or zincblende lattices, the maxima in the valence band energy and minima in the conduction band energy occur at $k = 0$ or along one of these two directions.

Figure 3.25a shows the E versus k diagram for GaAs. The valence band maximum and the conduction band minimum both occur at $k = 0$. The electrons in the

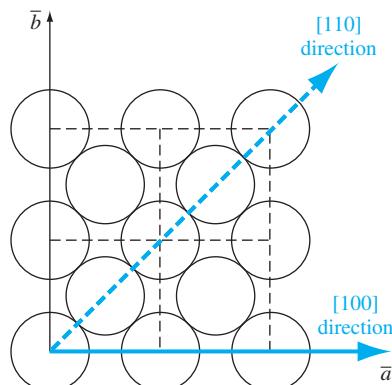


Figure 3.24 | The (100) plane of a face-centered cubic crystal showing the [100] and [110] directions.

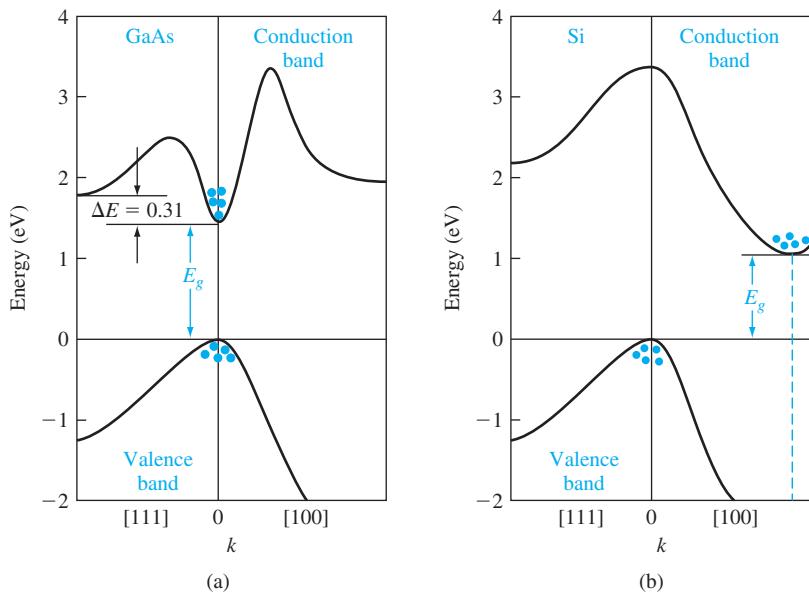


Figure 3.25 | Energy-band structures of (a) GaAs and (b) Si.
(From Sze [12].)

conduction band tend to settle at the minimum conduction band energy that is at $k = 0$. Similarly, holes in the valence band tend to congregate at the uppermost valence band energy. In GaAs, the minimum conduction band energy and maximum valence band energy occur at the same k value. A semiconductor with this property is said to be a *direct* bandgap semiconductor; transitions between the two allowed bands can take place with no change in crystal momentum. This direct nature has significant effect on the optical properties of the material. GaAs and other direct bandgap materials are ideally suited for use in semiconductor lasers and other optical devices.

The E versus k diagram for silicon is shown in Figure 3.25b. The maximum in the valence band energy occurs at $k = 0$ as before. The minimum in the conduction band energy occurs not at $k = 0$, but along the [100] direction. The difference between the minimum conduction band energy and the maximum valence band energy is still defined as the bandgap energy E_g . A semiconductor whose maximum valence band energy and minimum conduction band energy do not occur at the same k value is called an *indirect* bandgap semiconductor. When electrons make a transition between the conduction and valence bands, we must invoke the law of conservation of momentum. A transition in an indirect bandgap material must necessarily include an interaction with the crystal so that crystal momentum is conserved.

Germanium is also an indirect bandgap material, whose valence band maximum occurs at $k = 0$ and whose conduction band minimum occurs along the [111] direction. GaAs is a direct bandgap semiconductor, but other compound semiconductors, such as GaP and AlAs, have indirect bandgaps.

3.3.2 Additional Effective Mass Concepts

The curvature of the E versus k diagrams near the minimum of the conduction band energy is related to the effective mass of the electron. We may note from Figure 3.25 that the curvature of the conduction band at its minimum value for GaAs is larger than that of silicon, so the effective mass of an electron in the conduction band of GaAs will be smaller than that in silicon.

For the one-dimensional E versus k diagram, the effective mass was defined by Equation (3.41) as $1/m^* = 1/\hbar^2 \cdot dE/dk^2$. A complication occurs in the effective mass concept in a real crystal. A three-dimensional crystal can be described by three k vectors. The curvature of the E versus k diagram at the conduction band minimum may not be the same in the three k directions. In later sections and chapters, the effective mass parameters used in calculations will be a kind of statistical average that is adequate for most device calculations.²

3.4 | DENSITY OF STATES FUNCTION

As we have stated, we eventually wish to describe the current–voltage characteristics of semiconductor devices. Since current is due to the flow of charge, an important step in the process is to determine the number of electrons and holes in the semiconductor that will be available for conduction. The number of carriers that can contribute to the conduction process is a function of the number of available energy or quantum states since, by the Pauli exclusion principle, only one electron can occupy a given quantum state. When we discussed the splitting of energy levels into bands of allowed and forbidden energies, we indicated that the band of allowed energies was actually made up of discrete energy levels. We must determine the density of these allowed energy states as a function of energy in order to calculate the electron and hole concentrations.

3.4.1 Mathematical Derivation

To determine the density of allowed quantum states as a function of energy, we need to consider an appropriate mathematical model. Electrons are allowed to move relatively freely in the conduction band of a semiconductor but are confined to the crystal. As a first step, we will consider a free electron confined to a three-dimensional infinite potential well, where the potential well represents the crystal. The potential of the infinite potential well is defined as

$$\begin{aligned} V(x, y, z) &= 0 && \text{for } 0 < x < a \\ && 0 < y < a \\ && 0 < z < a \\ V(x, y, z) &= \infty && \text{elsewhere} \end{aligned} \tag{3.59}$$

where the crystal is assumed to be a cube with length a . Schrodinger's wave equation in three dimensions can be solved by using the separation of variables technique.

²See Appendix F for further discussion of effective mass concepts.

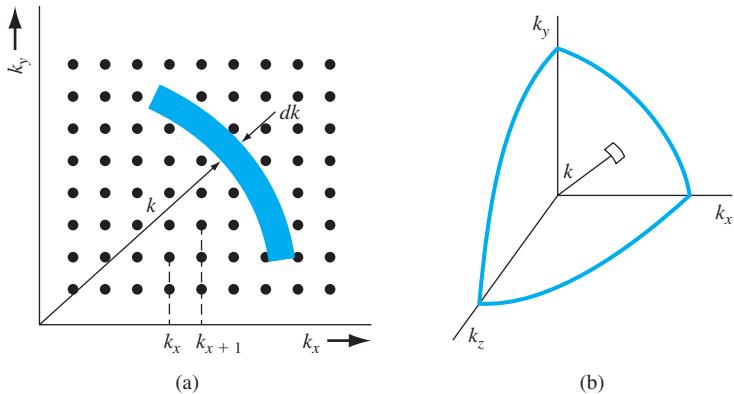


Figure 3.26 | (a) A two-dimensional array of allowed quantum states in k space. (b) The positive one-eighth of the spherical k space.

Extrapolating the results from the one-dimensional infinite potential well, we can show (see Problem 3.23) that

$$\frac{2mE}{\hbar^2} = k^2 = k_x^2 + k_y^2 + k_z^2 = (n_x^2 + n_y^2 + n_z^2) \left(\frac{\pi^2}{a^2} \right) \quad (3.60)$$

where n_x , n_y , and n_z are positive integers. (Negative values of n_x , n_y , and n_z yield the same wave function, except for the sign, as the positive integer values, resulting in the same probability function and energy; therefore the negative integers do not represent a different quantum state.)

We can schematically plot the allowed quantum states in k space. Figure 3.26a shows a two-dimensional plot as a function of k_x and k_y . Each point represents an allowed quantum state corresponding to various integral values of n_x and n_y . Positive and negative values of k_x , k_y , or k_z have the same energy and represent the same energy state. Since negative values of k_x , k_y , or k_z do not represent additional quantum states, the density of quantum states will be determined by considering only the positive one-eighth of the spherical k space as shown in Figure 3.26b.

The distance between two quantum states in the k_x direction, for example, is given by

$$k_{x+1} - k_x = (n_x + 1) \left(\frac{\pi}{a} \right) - n_x \left(\frac{\pi}{a} \right) = \frac{\pi}{a} \quad (3.61)$$

Generalizing this result to three dimensions, the volume V_k of a single quantum state is

$$V_k = \left(\frac{\pi}{a} \right)^3 \quad (3.62)$$

We can now determine the density of quantum states in k space. A differential volume in k space is shown in Figure 3.26b and is given by $4\pi k^2 dk$, so the differential density of quantum states in k space can be written as

$$g_T(k) dk = 2 \left(\frac{1}{8} \right) \frac{4\pi k^2 dk}{\left(\frac{\pi}{a} \right)^3} \quad (3.63)$$

The first factor, 2, takes into account the two spin states allowed for each quantum state; the next factor, $\frac{1}{8}$, takes into account that we are considering only the quantum states for positive values of k_x , k_y , and k_z . The factor $4\pi k^2 dk$ is again the differential volume and the factor $(\pi/a)^3$ is the volume of one quantum state. Equation (3.63) may be simplified to

$$g_T(k) dk = \frac{\pi k^2 dk}{\pi^3} \cdot a^3 \quad (3.64)$$

Equation (3.64) gives the density of quantum states as a function of momentum, through the parameter k . We can now determine the density of quantum states as a function of energy E . For a free electron, the parameters E and k are related by

$$k^2 = \frac{2mE}{\hbar^2} \quad (3.65a)$$

or

$$k = \frac{1}{\hbar} \sqrt{2mE} \quad (3.65b)$$

The differential dk is

$$dk = \frac{1}{\hbar} \sqrt{\frac{m}{2E}} dE \quad (3.66)$$

Then, substituting the expressions for k^2 and dk into Equation (3.64), the number of energy states between E and $E + dE$ is given by

$$g_T(E) dE = \frac{\pi a^3}{\pi^3} \left(\frac{2mE}{\hbar^2} \right) \cdot \frac{1}{\hbar} \sqrt{\frac{m}{2E}} dE \quad (3.67)$$

Since $\hbar = h/2\pi$, Equation (3.67) becomes

$$g_T(E) dE = \frac{4\pi a^3}{h^3} \cdot (2m)^{3/2} \cdot \sqrt{E} dE \quad (3.68)$$

Equation (3.68) gives the total number of quantum states between the energy E and $E + dE$ in the crystal space volume of a^3 . If we divide by the volume a^3 , then we will obtain the density of quantum states per unit volume of the crystal. Equation (3.68) then becomes

$$g(E) = \frac{4\pi(2m)^{3/2}}{h^3} \sqrt{E} \quad (3.69)$$

The density of quantum states is a function of energy E . As the energy of this free electron becomes small, the number of available quantum states decreases. This density function is really a double density, in that the units are given in terms of states per unit energy per unit volume.

Objective: Calculate the density of states per unit volume over a particular energy range.

Consider the density of states for a free electron given by Equation (3.69). Calculate the density of states per unit volume with energies between 0 and 1 eV.

EXAMPLE 3.3

■ Solution

The volume density of quantum states, from Equation (3.69), is

$$N = \int_0^{1 \text{ eV}} g(E) dE = \frac{4\pi(2m)^{3/2}}{\hbar^3} \cdot \int_0^{1 \text{ eV}} \sqrt{E} dE$$

or

$$N = \frac{4\pi(2m)^{3/2}}{\hbar^3} \cdot \frac{2}{3} \cdot E^{3/2}$$

The density of states is now

$$N = \frac{4\pi[2(9.11 \times 10^{-31})]^{3/2}}{(6.625 \times 10^{-34})^3} \cdot \frac{2}{3} \cdot (1.6 \times 10^{-19})^{3/2} = 4.5 \times 10^{27} \text{ m}^{-3}$$

or

$$N = 4.5 \times 10^{21} \text{ states/cm}^3$$

■ Comment

The density of quantum states is typically a large number. An effective density of states in a semiconductor, as we will see in the following sections and in the next chapter, is also a large number but is usually less than the density of atoms in the semiconductor crystal.

■ EXERCISE PROBLEM

Ex 3.3 For a free electron, calculate the density of quantum states (#/cm³)

over the energy range of (a) $0 \leq E \leq 2.0 \text{ eV}$ and (b) $1 \leq E \leq 2 \text{ eV}$.

[Ans. (a) $N = 1.28 \times 10^{22} \text{ cm}^{-3}$; (b) $N = 8.29 \times 10^{21} \text{ cm}^{-3}$]

3.4.2 Extension to Semiconductors

In the previous section, we derived a general expression for the density of allowed electron quantum states using the model of a free electron with mass m bounded in a three-dimensional infinite potential well. We can extend this same general model to a semiconductor to determine the density of quantum states in the conduction band and the density of quantum states in the valence band. Electrons and holes are confined within the semiconductor crystal, so we will again use the basic model of the infinite potential well.

The parabolic relationship between energy and momentum of a free electron is given in Equation (3.28) as $E = p^2/2m = \hbar^2 k^2/2m$. Figure 3.16a shows the conduction energy band in the reduced k space. The E versus k curve near $k = 0$ at the bottom of the conduction band can be approximated as a parabola, so we may write

$$E = E_c + \frac{\hbar^2 k^2}{2m_n^*} \quad (3.70)$$

where E_c is the bottom edge of the conduction band and m_n^* is the electron density of states effective mass.³ Equation (3.70) may be rewritten to give

$$E - E_c = \frac{\hbar^2 k^2}{2m_n^*} \quad (3.71)$$

³Again, see Appendix F for further discussion of effective mass concepts.

The general form of the E versus k relation for an electron in the bottom of a conduction band is the same as the free electron, except the mass is replaced by the effective mass. We can then think of the electron in the bottom of the conduction band as being a “free” electron with its own particular mass. The right side of Equation (3.71) is of the same form as the right side of Equation (3.28), which was used in the derivation of the density of states function. Because of this similarity, which yields the “free” conduction electron model, we may generalize the free electron results of Equation (3.69) and write the density of allowed electronic energy states in the conduction band as

$$g_c(E) = \frac{4\pi(2m_n^*)^{3/2}}{h^3} \sqrt{E - E_c} \quad (3.72)$$

Equation (3.72) is valid for $E \geq E_c$. As the energy of the electron in the conduction band decreases, the number of available quantum states also decreases.

The density of quantum states in the valence band can be obtained by using the same infinite potential well model, since the hole is also confined in the semiconductor crystal and can be treated as a “free” particle. The density of states effective mass of the hole is m_p^* . Figure 3.16b shows the valence energy band in the reduced k space. We may also approximate the E versus k curve near $k = 0$ by a parabola for a “free” hole, so that

$$E = E_v - \frac{\hbar^2 k^2}{2m_p^*} \quad (3.73)$$

Equation (3.73) may be rewritten to give

$$E_v - E = \frac{\hbar^2 k^2}{2m_p^*} \quad (3.74)$$

Again, the right side of Equation (3.74) is of the same form used in the general derivation of the density of states function. We may then generalize the density of states function from Equation (3.69) to apply to the valence band, so that

$$g_v(E) = \frac{4\pi(2m_p^*)^{3/2}}{h^3} \sqrt{E_v - E} \quad (3.75)$$

Equation (3.75) is valid for $E \leq E_v$.

We have argued that quantum states do not exist within the forbidden energy band, so $g(E) = 0$ for $E_v < E < E_c$. Figure 3.27 shows the plot of the density of quantum states as a function of energy. If the electron and hole effective masses were equal, then the functions $g_c(E)$ and $g_v(E)$ would be symmetrical about the energy midway between E_c and E_v , or the midgap energy, E_{midgap} .

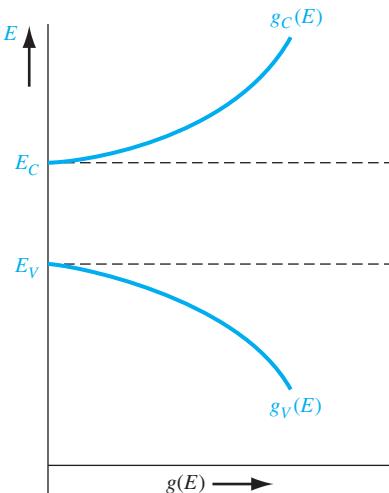


Figure 3.27 | The density of energy states in the conduction band and the density of energy states in the valence band as a function of energy.

EXAMPLE 3.4

Objective: Determine the number (#/cm³) of quantum states in silicon between E_c and $E_c + kT$ at $T = 300$ K.

■ Solution

Using Equation (3.72), we can write

$$\begin{aligned} N &= \int_{E_c}^{E_c + kT} \frac{4\pi(2m_n^*)^{3/2}}{\hbar^3} \sqrt{E - E_c} \cdot dE \\ &= \frac{4\pi(2m_n^*)^{3/2}}{\hbar^3} \cdot \frac{2}{3} \cdot (E - E_c)^{3/2} \Big|_{E_c}^{E_c + kT} \\ &= \frac{4\pi[2(1.08)(9.11 \times 10^{-31})]^{3/2}}{(6.625 \times 10^{-34})^3} \cdot \frac{2}{3} \cdot [(0.0259)(1.6 \times 10^{-19})]^{3/2} \\ &= 2.12 \times 10^{25} \text{ m}^{-3} \end{aligned}$$

or

$$N = 2.12 \times 10^{19} \text{ cm}^{-3}$$

■ Comment

The result of this example shows the order of magnitude of the density of quantum states in a semiconductor.

■ EXERCISE PROBLEM

Ex3.4 Determine the number (#/cm³) of quantum states in silicon between $(E_v - kT)$ and E_v at $T = 300$ K. **Ans. $N = 7.92 \times 10^{18} \text{ cm}^{-3}$**

3.5 | STATISTICAL MECHANICS

In dealing with large numbers of particles, we are interested only in the statistical behavior of the group as a whole rather than in the behavior of each individual particle. For example, gas within a container will exert an average pressure on the walls of the vessel. The pressure is actually due to the collisions of the individual gas molecules with the walls, but we do not follow each individual molecule as it collides with the wall. Likewise in a crystal, the electrical characteristics will be determined by the statistical behavior of a large number of electrons.

3.5.1 Statistical Laws

In determining the statistical behavior of particles, we must consider the laws that the particles obey. There are three distribution laws determining the distribution of particles among available energy states.

One distribution law is the Maxwell–Boltzmann probability function. In this case, the particles are considered to be distinguishable by being numbered, for example, from 1 to N , with no limit to the number of particles allowed in each energy state. The behavior of gas molecules in a container at fairly low pressure is an example of this distribution.

A second distribution law is the Bose–Einstein function. The particles in this case are indistinguishable and, again, there is no limit to the number of particles permitted in each quantum state. The behavior of photons, or black body radiation, is an example of this law.

The third distribution law is the Fermi–Dirac probability function. In this case, the particles are again indistinguishable, but now only one particle is permitted in each quantum state. Electrons in a crystal obey this law. In each case, the particles are assumed to be noninteracting.

3.5.2 The Fermi–Dirac Probability Function

Figure 3.28 shows the i th energy level with g_i quantum states. A maximum of one particle is allowed in each quantum state by the Pauli exclusion principle. There are g_i ways of choosing where to place the first particle, $(g_i - 1)$ ways of choosing where to place the second particle, $(g_i - 2)$ ways of choosing where to place the third particle, and so on. Then the total number of ways of arranging N_i particles in the i th energy level (where $N_i \leq g_i$) is

$$(g_i)(g_i - 1) \cdots (g_i - (N_i - 1)) = \frac{g_i!}{(g_i - N_i)!} \quad (3.76)$$

This expression includes all permutations of the N_i particles among themselves.

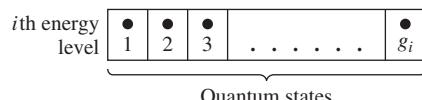


Figure 3.28 | The i th energy level with g_i quantum states.

However, since the particles are indistinguishable, the $N_i!$ number of permutations that the particles have among themselves in any given arrangement do not count as separate arrangements. The interchange of any two electrons, for example, does not produce a new arrangement. Therefore, the actual number of independent ways of realizing a distribution of N_i particles in the i th level is

$$W_i = \frac{g_i!}{N_i!(g_i - N_i)!} \quad (3.77)$$

EXAMPLE 3.5

Objective: Determine the possible number of ways of realizing a particular distribution for (a) $g_i = N_i = 10$ and (b) $g_i = 10, N_i = 9$.

■ Solution

(a) $g_i = N_i = 10$: We may note that $(g_i - N_i)! = 0! = 1$. Then, from Equation (3.77), we find

$$\frac{g_i!}{N_i!(g_i - N_i)!} = \frac{10!}{10!} = 1$$

(b) $g_i = 10, N_i = 9$: We may note that $(g_i - N_i)! = 1! = 1$. Then, we find

$$\frac{g_i!}{N_i!(g_i - N_i)!} = \frac{10!}{(9!)(1)} = \frac{(10)(9!)}{(9!)} = 10$$

■ Comment

In part (a), we have 10 particles to be arranged in 10 quantum states. There is only one possible arrangement. Each quantum state contains one particle. In part (b), we have 9 particles to be arranged in 10 quantum states. There is one empty quantum state, and there are 10 possible positions in which that empty state may occur. Thus, there are 10 possible arrangements for this case.

■ EXERCISE PROBLEM

Ex3.5 Determine the possible number of ways of realizing a particular distribution if $g_i = 10$ and $N_i = 8$. (Ans. 45)

Equation (3.77) gives the number of independent ways of realizing a distribution of N_i particles in the i th level. The total number of ways of arranging $(N_1, N_2, N_3, \dots, N_n)$ indistinguishable particles among n energy levels is the product of all distributions, or

$$W = \prod_{i=1}^n \frac{g_i!}{N_i!(g_i - N_i)!} \quad (3.78)$$

The parameter W is the total number of ways in which N electrons can be arranged in this system, where $N = \sum_{i=1}^n N_i$ is the total number of electrons in the system. We want to find the most probable distribution, which means that we want to find the maximum W . The maximum W is found by varying N_i among the E_i levels, which varies the distribution, but at the same time, we will keep the total number of particles and total energy constant.

We may write the most probable distribution function as

$$\frac{N(E)}{g(E)} = f_F(E) = \frac{1}{1 + \exp\left(\frac{E - E_F}{kT}\right)} \quad (3.79)$$

The number density $N(E)$ is the number of particles per unit volume per unit energy and the function $g(E)$ is the number of quantum states per unit volume per unit energy. The function $f_F(E)$ is called the *Fermi–Dirac distribution* or probability function and gives the probability that a quantum state at the energy E will be occupied by an electron. The energy E_F is called the *Fermi energy*. Another interpretation of the distribution function is that $f_F(E)$ is the ratio of filled to total quantum states at any energy E .

3.5.3 The Distribution Function and the Fermi Energy

To begin to understand the meaning of the distribution function and the Fermi energy, we can plot the distribution function versus energy. Initially, let $T = 0$ K and consider the case when $E < E_F$. The exponential term in Equation (3.79) becomes $\exp[(E - E_F)/kT] \rightarrow \exp(-\infty) = 0$. The resulting distribution function is $f_F(E < E_F) = 1$. Again let $T = 0$ K and consider the case when $E > E_F$. The exponential term in the distribution function becomes $\exp[(E - E_F)/kT] \rightarrow \exp(+\infty) \rightarrow +\infty$. The resulting Fermi–Dirac distribution function now becomes $f_F(E > E_F) = 0$.

The Fermi–Dirac distribution function for $T = 0$ K is plotted in Figure 3.29. This result shows that, for $T = 0$ K, the electrons are in their lowest possible energy states. The probability of a quantum state being occupied is unity for $E < E_F$ and the probability of a state being occupied is zero for $E > E_F$. All electrons have energies below the Fermi energy at $T = 0$ K.

Figure 3.30 shows discrete energy levels of a particular system as well as the number of available quantum states at each energy. If we assume, for this case, that the system contains 13 electrons, then Figure 3.30 shows how these electrons are distributed among the various quantum states at $T = 0$ K. The electrons will be in the lowest possible energy state, so the probability of a quantum state being occupied in energy levels E_1 through E_4 is unity, and the probability of a quantum state being occupied in energy level E_5 is zero. The Fermi energy, for this case, must be above E_4 but less than E_5 . The Fermi energy determines the statistical distribution of electrons and does not have to correspond to an allowed energy level.

Now consider a case in which the density of quantum states $g(E)$ is a continuous function of energy, as shown in Figure 3.31. If we have N_0 electrons in this system,

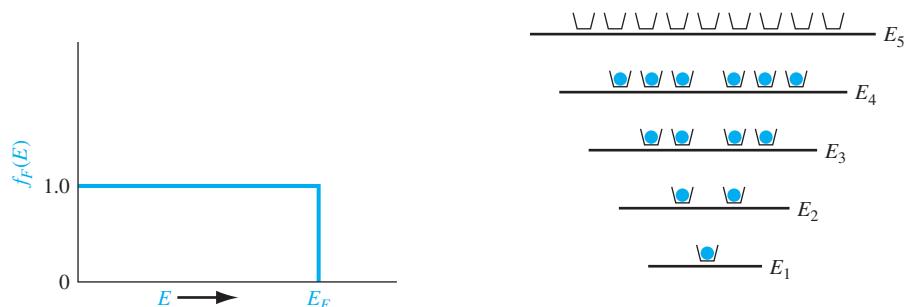


Figure 3.29 | The Fermi probability function versus energy for $T = 0$ K.

Figure 3.30 | Discrete energy states and quantum states for a particular system at $T = 0$ K.

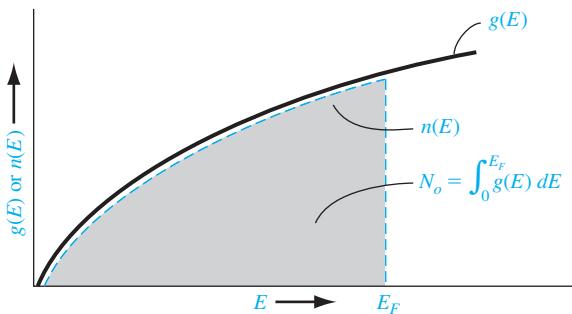


Figure 3.31 | Density of quantum states and electrons in a continuous energy system at $T = 0$ K.

then the distribution of these electrons among the quantum states at $T = 0$ K is shown by the dashed line. The electrons are in the lowest possible energy state so that all states below E_F are filled and all states above E_F are empty. If $g(E)$ and N_0 are known for this particular system, then the Fermi energy E_F can be determined.

Consider the situation when the temperature increases above $T = 0$ K. Electrons gain a certain amount of thermal energy so that some electrons can jump to higher energy levels, which means that the distribution of electrons among the available energy states will change. Figure 3.32 shows the same discrete energy levels and quantum states as in Figure 3.30. The distribution of electrons among the quantum states has changed from the $T = 0$ K case. Two electrons from the E_4 level have gained enough energy to jump to E_5 , and one electron from E_3 has jumped to E_4 . As the temperature changes, the distribution of electrons versus energy changes.

The change in the electron distribution among energy levels for $T > 0$ K can be seen by plotting the Fermi–Dirac distribution function. If we let $E = E_F$ and $T > 0$ K, then Equation (3.79) becomes

$$f_F(E = E_F) = \frac{1}{1 + \exp(0)} = \frac{1}{1 + 1} = \frac{1}{2}$$

The probability of a state being occupied at $E = E_F$ is $\frac{1}{2}$. Figure 3.33 shows the Fermi–Dirac distribution function plotted for several temperatures, assuming that the Fermi energy is independent of temperature.

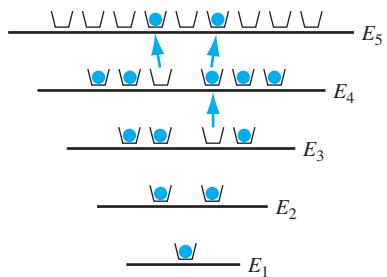


Figure 3.32 | Discrete energy states and quantum states for the same system shown in Figure 3.30 for $T > 0$ K.

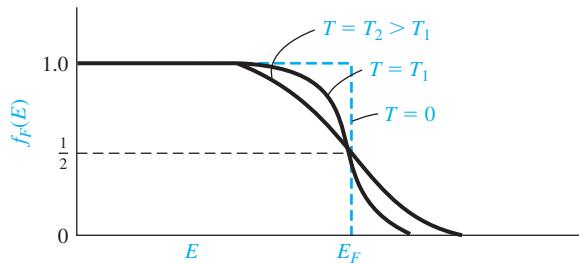


Figure 3.33 | The Fermi probability function versus energy for different temperatures.

We can see that for temperatures above absolute zero, there is a nonzero probability that some energy states above E_F will be occupied by electrons and some energy states below E_F will be empty. This result again means that some electrons have jumped to higher energy levels with increasing thermal energy.

Objective: Calculate the probability that an energy state above E_F is occupied by an electron.

Let $T = 300$ K. Determine the probability that an energy level $3kT$ above the Fermi energy is occupied by an electron.

EXAMPLE 3.6

■ Solution

From Equation (3.79), we can write

$$f_F(E) = \frac{1}{1 + \exp\left(\frac{E - E_F}{kT}\right)} = \frac{1}{1 + \exp\left(\frac{3kT}{kT}\right)}$$

which becomes

$$f_F(E) = \frac{1}{1 + 20.09} = 0.0474 = 4.74\%$$

■ Comment

At energies above E_F , the probability of a state being occupied by an electron can become significantly less than unity, or the ratio of electrons to available quantum states can be quite small.

■ EXERCISE PROBLEM

Ex3.6 Assume the Fermi energy level is 0.30 eV below the conduction band energy E_c .

Assume $T = 300$ K. (a) Determine the probability of a state being occupied by an electron at $E = E_c + kT/4$. (b) Repeat part (a) for an energy state at $E = E_c + kT$.

[Ans. (a) 7.26×10^{-6} ; (b) 3.43×10^{-6}]

We can see from Figure 3.33 that the probability of an energy above E_F being occupied increases as the temperature increases and the probability of a state below E_F being empty increases as the temperature increases.

Objective: Determine the temperature at which there is 1 percent probability that an energy state is empty.

EXAMPLE 3.7

Assume that the Fermi energy level for a particular material is 6.25 eV and that the electrons in this material follow the Fermi–Dirac distribution function. Calculate the temperature at which there is a 1 percent probability that a state 0.30 eV below the Fermi energy level will not contain an electron.

■ Solution

The probability that a state is empty is

$$1 - f_F(E) = 1 - \frac{1}{1 + \exp\left(\frac{E - E_F}{kT}\right)}$$

Then

$$0.01 = 1 - \frac{1}{1 + \exp\left(\frac{5.95 - 6.25}{kT}\right)}$$

Solving for kT , we find $kT = 0.06529$ eV, so that the temperature is $T = 756$ K.

Comment

The Fermi probability function is a strong function of temperature.

EXERCISE PROBLEM

- Ex 3.7** Assume that E_F is 0.3 eV below E_c . Determine the temperature at which the probability of an electron occupying an energy state at $E = (E_c + 0.025)$ eV is 8×10^{-6} .
 Ans. $T = 321$ K

We may note that the probability of a state a distance dE above E_F being occupied is the same as the probability of a state a distance dE below E_F being empty. The function $f_F(E)$ is symmetrical with the function $1 - f_F(E)$ about the Fermi energy, E_F . This symmetry effect is shown in Figure 3.34 and will be used in the next chapter.

Consider the case when $E - E_F \gg kT$, where the exponential term in the denominator of Equation (3.79) is much greater than unity. We may neglect the 1 in the denominator, so the Fermi–Dirac distribution function becomes

$$f_F(E) \approx \exp\left[\frac{-(E - E_F)}{kT}\right] \quad (3.80)$$

Equation (3.80) is known as the Maxwell–Boltzmann approximation, or simply the Boltzmann approximation, to the Fermi–Dirac distribution function. Figure 3.35 shows the Fermi–Dirac probability function and the Boltzmann approximation. This figure gives an indication of the range of energies over which the approximation is valid.

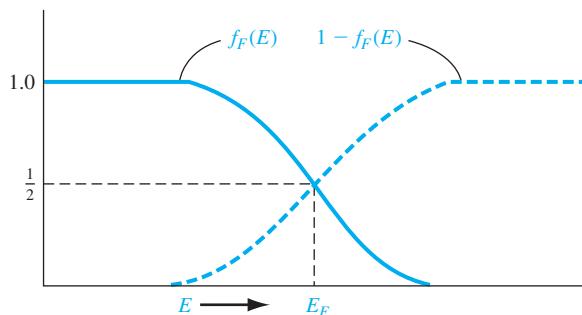


Figure 3.34 | The probability of a state being occupied, $f_F(E)$, and the probability of a state being empty, $1 - f_F(E)$.

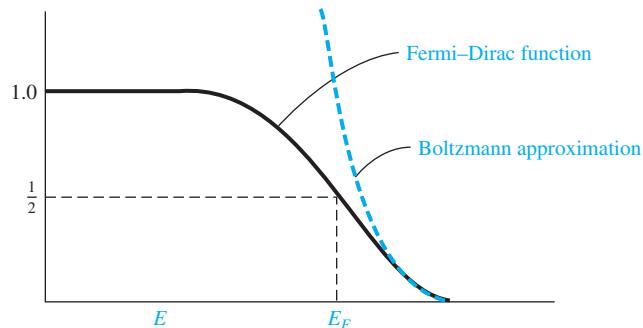


Figure 3.35 | The Fermi-Dirac probability function and the Maxwell-Boltzmann approximation.

Objective: Determine the energy at which the Boltzmann approximation may be considered valid.

EXAMPLE 3.8

Calculate the energy, in terms of kT and E_F , at which the difference between the Boltzmann approximation and the Fermi-Dirac function is 5 percent of the Fermi function.

■ Solution

We can write

$$\frac{\exp\left[\frac{-(E - E_F)}{kT}\right] - \frac{1}{1 + \exp\left(\frac{E - E_F}{kT}\right)}}{\frac{1}{1 + \exp\left(\frac{E - E_F}{kT}\right)}} = 0.05$$

If we multiply both numerator and denominator by the $1 + \exp()$ function, we have

$$\exp\left[\frac{-(E - E_F)}{kT}\right] \cdot \left\{1 + \exp\left[\frac{E - E_F}{kT}\right]\right\} - 1 = 0.05$$

which becomes

$$\exp\left[\frac{-(E - E_F)}{kT}\right] = 0.05$$

or

$$(E - E_F) = kT \ln\left(\frac{1}{0.05}\right) \approx 3kT$$

■ Comment

As seen in this example and in Figure 3.35, the $E - E_F \gg kT$ notation is somewhat misleading. The Maxwell-Boltzmann and Fermi-Dirac functions are within 5 percent of each other when $E - E_F \approx 3kT$.

■ EXERCISE PROBLEM

Ex 3.8 Repeat Example 3.8 for the case when the difference between the Boltzmann approximation and the Fermi-Dirac function is 2 percent of the Fermi function.

(Ans. $E - E_F = 3.9kT$)

The actual Boltzmann approximation is valid when $\exp[(E - E_c)/kT] \gg 1$. However, it is still common practice to use the $E - E_c \gg kT$ notation when applying the Boltzmann approximation. We will use this Boltzmann approximation in our discussion of semiconductors in the next chapter.

TEST YOUR UNDERSTANDING

- TYU 3.5** Assume that the Fermi energy level is 0.35 eV above the valence band energy. Let $T = 300$ K. (a) Determine the probability of a state being empty of an electron at $E = E_v - kT/2$. (b) Repeat part (a) for an energy state at $E = E_v - 3kT/2$.
[Ans. (a) 8.20×10^{-7} ; (b) 3.02×10^{-7}]
- TYU 3.6** Repeat Exercise Problem Ex 3.6 for $T = 400$ K.
[Ans. (a) 1.31×10^{-4} ; (b) 6.21×10^{-5}]
- TYU 3.7** Repeat Exercise Problem TYU 3.5 for $T = 400$ K.
[Ans. (a) 2.41×10^{-5} ; (b) 8.85×10^{-6}]

3.6 | SUMMARY

- Discrete allowed electron energies split into a band of allowed energies as atoms are brought together to form a crystal.
- The concept of allowed and forbidden energy bands was developed more rigorously by considering quantum mechanics and Schrodinger's wave equation using the Kronig–Penney model representing the potential function of a single-crystal material. This result forms the basis of the energy-band theory of semiconductors.
- The concept of effective mass was developed. Effective mass relates the motion of a particle in a crystal to an externally applied force and takes into account the effect of the crystal lattice on the motion of the particle.
- Two charged particles exist in a semiconductor. An electron is a negatively charged particle with a positive effective mass existing at the bottom of an allowed energy band. A hole is a positively charged particle with a positive effective mass existing at the top of an allowed energy band.
- The E versus k diagrams of silicon and gallium arsenide were given and the concept of direct and indirect bandgap semiconductors was discussed.
- Energies within an allowed energy band are actually at discrete levels and each contains a finite number of quantum states. The density per unit energy of quantum states was determined by using the three-dimensional infinite potential well as a model.
- In dealing with large numbers of electrons and holes, we must consider the statistical behavior of these particles. The Fermi–Dirac probability function was developed, which gives the probability of a quantum state at an energy E of being occupied by an electron. The Fermi energy was defined.

GLOSSARY OF IMPORTANT TERMS

allowed energy band A band or range of energy levels that an electron in a crystal is allowed to occupy based on quantum mechanics.

density of states function The density of available quantum states as a function of energy, given in units of number per unit energy per unit volume.

electron effective mass The parameter that relates the acceleration of an electron in the conduction band of a crystal to an external force; a parameter that takes into account the effect of internal forces in the crystal.

Fermi–Dirac probability function The function describing the statistical distribution of electrons among available energy states and the probability that an allowed energy state is occupied by an electron.

fermi energy In the simplest definition, the energy below which all states are filled with electrons and above which all states are empty at $T = 0$ K.

forbidden energy band A band or range of energy levels that an electron in a crystal is not allowed to occupy based on quantum mechanics.

hole The positively charged “particle” associated with an empty state in the top of the valence band.

hole effective mass The parameter that relates the acceleration of a hole in the valence band of a crystal to an applied external force (a positive quantity); a parameter that takes into account the effect of internal forces in a crystal.

k -space diagram The plot of electron energy in a crystal versus k , where k is the momentum-related constant of the motion that incorporates the crystal interaction.

Kronig–Penney model The mathematical model of a periodic potential function representing a one-dimensional single-crystal lattice by a series of periodic step functions.

Maxwell–Boltzmann approximation The condition in which the energy is several kT above the Fermi energy or several kT below the Fermi energy so that the Fermi–Dirac probability function can be approximated by a simple exponential function.

CHECKPOINT

After studying this chapter, the reader should have the ability to

- Discuss the concept of allowed and forbidden energy bands in a single crystal both qualitatively and more rigorously from the results of using the Kronig–Penney model.
- Discuss the splitting of energy bands in silicon.
- State the definition of effective mass from the E versus k diagram and discuss its meaning in terms of the movement of a particle in a crystal.
- Discuss the concept of a hole.
- Discuss the characteristics of a direct and an indirect bandgap semiconductor.
- Qualitatively, in terms of energy bands, discuss the difference between a metal, an insulator, and semiconductor.
- What is meant by the density of states function?
- Understand the meaning of the Fermi–Dirac distribution function and the Fermi energy.

REVIEW QUESTIONS

1. What is the Kronig–Penney model? What does it represent?
2. State two results of using the Kronig–Penney model with Schrodinger’s wave equation.
3. What is effective mass? How is effective mass defined in terms of the E versus k diagram?
4. What is a direct bandgap semiconductor? What is an indirect bandgap semiconductor?
5. What is the meaning of the density of states function?

6. What was the mathematical model used in deriving the density of states function?
7. In general, what is the relation between density of states and energy?
8. What is the meaning of the Fermi–Dirac probability function?
9. What is the Fermi energy?

PROBLEMS

Section 3.1 Allowed and Forbidden Energy Bands

- 3.1 Consider Figure 3.4b, which shows the energy-band splitting of silicon. If the equilibrium lattice spacing were to change by a small amount, discuss how you would expect the electrical properties of silicon to change. Determine at what point the material would behave like an insulator or like a metal.
- 3.2 Show that Equations (3.4) and (3.6) are derived from Schrodinger's wave equation, using the form of solution given by Equation (3.3).
- 3.3 Show that Equations (3.9) and (3.10) are solutions of the differential equations given by Equations (3.4) and (3.8), respectively.
- 3.4 Show that Equations (3.12), (3.14), (3.16), and (3.18) result from the boundary conditions in the Kronig–Penney model.
- 3.5 (a) Plot the function $f(\alpha a) = 12(\sin \alpha a)/\alpha a + \cos \alpha a$ for $0 \leq \alpha a \leq 4\pi$. Also, given the function $f(\alpha a) = \cos ka$, indicate the allowed values of αa that will satisfy this equation. (b) Determine the values of αa at (i) $ka = \pi$ and (ii) $ka = 2\pi$.
- 3.6 Repeat Problem 3.5 for the function $f(\alpha a) = 5(\sin \alpha a)/\alpha a + \cos \alpha a = \cos ka$.
- 3.7 Using Equation (3.24), show that $dE/dk = 0$ at $k = n\pi/a$, where $n = 0, 1, 2, \dots$
- 3.8 Using the parameters of Problem 3.5 for a free electron and letting $a = 4.2 \text{ \AA}$, determine the width (in eV) of the forbidden energy bands that exist at (a) $ka = \pi$ and (b) $ka = 2\pi$. (Refer to Figure 3.8c).
- 3.9 Using the parameters in Problem 3.5 for a free electron and letting $a = 4.2 \text{ \AA}$, determine the width (in eV) of the allowed energy bands that exist for (a) $0 < ka < \pi$ and (b) $\pi < ka < 2\pi$.
- 3.10 Repeat Problem 3.8 using the parameters in Problem 3.6.
- 3.11 Repeat Problem 3.9 using the parameters in Problem 3.6.
- 3.12 The bandgap energy in a semiconductor is usually a slight function of temperature. In some cases, the bandgap energy versus temperature can be modeled by

$$E_g = E_g(0) - \frac{\alpha T^2}{(\beta + T)}$$

where $E_g(0)$ is the value of the bandgap energy at $T = 0 \text{ K}$. For silicon, the parameter values are $E_g(0) = 1.170 \text{ eV}$, $\alpha = 4.73 \times 10^{-4} \text{ eV/K}$, and $\beta = 636 \text{ K}$. Plot E_g versus T over the range $0 \leq T \leq 600 \text{ K}$. In particular, note the value at $T = 300 \text{ K}$.

Section 3.2 Electrical Conduction in Solids

- 3.13 Two possible conduction bands are shown in the E versus k diagram given in Figure P3.13. State which band will result in the heavier electron effective mass; state why.

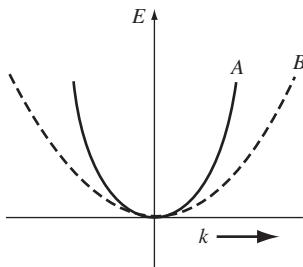


Figure P3.13 | Conduction bands for Problem 3.13.

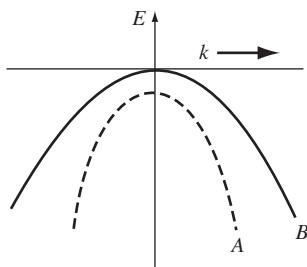


Figure P3.14 | Valence bands for Problem 3.14.

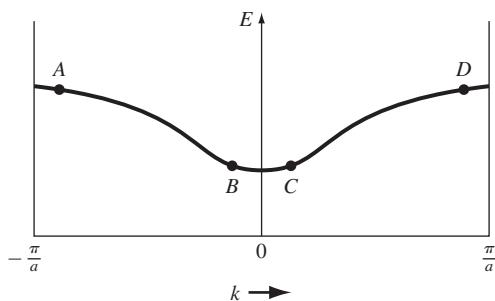


Figure P3.15 | Figure for Problem 3.15.

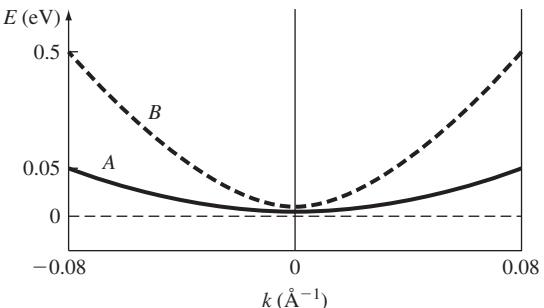


Figure P3.16 | Figure for Problem 3.16.

- 3.14** Two possible valence bands are shown in the E versus k diagram given in Figure P3.14. State which band will result in the heavier hole effective mass; state why.
- 3.15** The E versus k diagram for a particular allowed energy band is shown in Figure P3.15. Determine (a) the sign of the effective mass and (b) the direction of velocity for a particle at each of the four positions shown.
- 3.16** Figure P3.16 shows the parabolic E versus k relationship in the conduction band for an electron in two particular semiconductor materials. Determine the effective mass (in units of the free electron mass) of the two electrons.
- 3.17** Figure P3.17 shows the parabolic E versus k relationship in the valence band for a hole in two particular semiconductor materials. Determine the effective mass (in units of the free electron mass) of the two holes.
- 3.18** (a) The forbidden bandgap energy in GaAs is 1.42 eV. (i) Determine the minimum frequency of an incident photon that can interact with a valence electron and elevate the electron to the conduction band. (ii) What is the corresponding wavelength? (b) Repeat part (a) for silicon with a bandgap energy of 1.12 eV.
- 3.19** The E versus k diagrams for a free electron (curve A) and for an electron in a semiconductor (curve B) are shown in Figure P3.19. Sketch (a) dE/dk versus k and (b) d^2E/dk^2 versus k for each curve. (c) What conclusion can you make concerning a comparison in effective masses for the two cases?

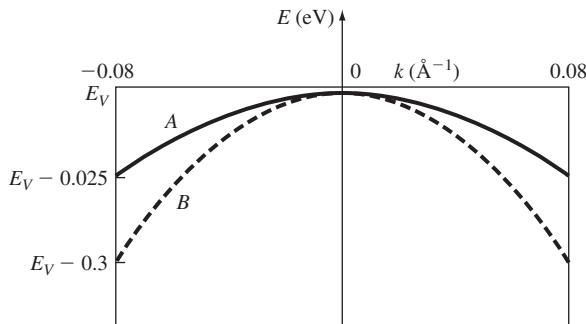


Figure P3.17 | Figure for Problem 3.17.

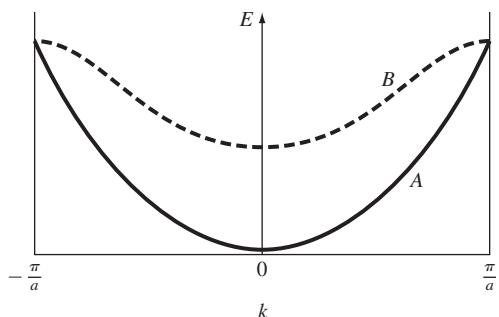


Figure P3.19 | Figure for Problem 3.19.

Section 3.3 Extension to Three Dimensions

- 3.20** The energy-band diagram for silicon is shown in Figure 3.25b. The minimum energy in the conduction band is in the [100] direction. The energy in this one-dimensional direction near the minimum value can be approximated by

$$E = E_0 - E_1 \cos \alpha(k - k_0)$$

where k_0 is the value of k at the minimum energy. Determine the effective mass of the particle at $k = k_0$ in terms of the equation parameters.

- 3.21** The constant kinetic energy curves for electrons in the conduction band of germanium consists of four ellipsoids similar to those in silicon (see Appendix F). The longitudinal and transverse effective masses are $m_l = 1.64m_0$ and $m_t = 0.082m_0$, respectively. Determine the (a) density of states effective mass and (b) conductivity effective mass.
- 3.22** Heavy and light holes exist in GaAs with effective masses $m_{hh} = 0.45m_0$ and $m_{lh} = 0.082m_0$, respectively. Determine the (a) density of states effective mass and (b) conductivity effective mass.

Section 3.4 Density of States Function

- 3.23** Starting with the three-dimensional infinite potential well function given by Equation (3.59) and using the separation of variables technique, derive Equation (3.60).
- 3.24** Show that Equation (3.69) can be derived from Equation (3.64).
- 3.25** Derive the density of states function for a one-dimensional electron gas in GaAs ($m_n^* = 0.067m_0$). Note that the kinetic energy may be written as $E = (\pm p)^2/2m_n^*$, which means that there are two momentum states for each energy level.
- 3.26** (a) Determine the total number (#/cm³) of energy states in silicon between E_c and $E_c + 2kT$ at (i) $T = 300$ K and (ii) $T = 400$ K. (b) Repeat part (a) for GaAs.
- 3.27** (a) Determine the total number (#/cm³) of energy states in silicon between E_v and $E_v - 3kT$ at (i) $T = 300$ K and (ii) $T = 400$ K. (b) Repeat part (a) for GaAs.
- 3.28** (a) Plot the density of states in the conduction band of silicon over the range $E_c < E < E_c + 0.4$ eV. (b) Repeat part (a) for the density of states in the valence band over the range $E_v - 0.4$ eV $< E < E_v$.

- 3.29** (a) For silicon, find the ratio of the density of states in the conduction band at $E = E_c + kT$ to the density of states in the valence band at $E = E_v - kT$. (b) Repeat part (a) for GaAs.

Section 3.5 Statistical Mechanics

- 3.30** Plot the Fermi–Dirac probability function, given by Equation (3.79), over the range $-0.2 \leq (E - E_F) \leq 0.2$ eV for (a) $T = 200$ K, (b) $T = 300$ K, and (c) $T = 400$ K.
- 3.31** (a) Repeat Example 3.5 for the case when $g_i = 10$ and $N_i = 7$. (b) Repeat part (a) for (i) $g_i = 12, N_i = 10$ and (ii) $g_i = 12, N_i = 8$.
- 3.32** Determine the probability that an energy level is occupied by an electron if the state is above the Fermi level by (a) kT , (b) $5kT$, and (c) $10kT$.
- 3.33** Determine the probability that an energy level is empty of an electron if the state is below the Fermi level by (a) kT , (b) $5kT$, and (c) $10kT$.
- 3.34** (a) The Fermi energy in silicon is 0.30 eV below the conduction band energy E_c at $T = 300$ K. Plot the probability of a state being occupied by an electron in the conduction band over the range $E_c \leq E \leq E_c + 2kT$. (b) The Fermi energy in silicon is 0.25 eV above the valence band energy E_v . Plot the probability of a state being empty by an electron in the valence band over the range $E_v - 2kT \text{ eV} \leq E \leq E_v$.
- 3.35** The probability that a state at $E_c + kT$ is occupied by an electron is equal to the probability that a state at $E_v - kT$ is empty. Determine the position of the Fermi energy level as a function of E_c and E_v .
- 3.36** Six free electrons exist in a one-dimensional infinite potential well of width $a = 12$ Å. Determine the Fermi energy level at $T = 0$ K.
- 3.37** (a) Five free electrons exist in a three-dimensional infinite potential well with all three widths equal to $a = 12$ Å. Determine the Fermi energy level at $T = 0$ K. (b) Repeat part (a) for 13 electrons.
- 3.38** Show that the probability of an energy state being occupied ΔE above the Fermi energy is the same as the probability of a state being empty ΔE below the Fermi level.
- 3.39** (a) Determine for what energy above E_F (in terms of kT) the Fermi–Dirac probability function is within 1 percent of the Boltzmann approximation. (b) Give the value of the probability function at this energy.
- 3.40** The Fermi energy level for a particular material at $T = 300$ K is 5.50 eV. The electrons in this material follow the Fermi–Dirac distribution function. (a) Find the probability of an electron occupying an energy at 5.80 eV. (b) Repeat part (a) if the temperature is increased to $T = 700$ K. (Assume that E_F is a constant.) (c) Determine the temperature at which there is a 2 percent probability that a state 0.25 eV below the Fermi level will be empty of an electron.
- 3.41** The Fermi energy for copper at $T = 300$ K is 7.0 eV. The electrons in copper follow the Fermi–Dirac distribution function. (a) Find the probability of an energy level at 7.15 eV being occupied by an electron. (b) Repeat part (a) for $T = 1000$ K. (Assume that E_F is a constant.) (c) Repeat part (a) for $E = 6.85$ eV and $T = 300$ K. (d) Determine the probability of the energy state at $E = E_F$ being occupied at $T = 300$ K and at $T = 1000$ K.
- 3.42** Consider the energy levels shown in Figure P3.42. Let $T = 300$ K. (a) If $E_1 - E_F = 0.30$ eV, determine the probability that an energy state at $E = E_1$ is occupied by

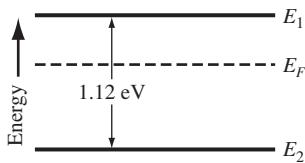


Figure P3.42 | Energy levels for Problem 3.42.

an electron and the probability that an energy state at $E = E_2$ is empty. (b) Repeat part (a) if $E_F - E_2 = 0.40$ eV.

- 3.43 Repeat problem 3.42 for the case when $E_1 - E_2 = 1.42$ eV.
- 3.44 Determine the derivative with respect to energy of the Fermi–Dirac distribution function. Plot the derivative with respect to energy for (a) $T = 0$ K, (b) $T = 300$ K, and (c) $T = 500$ K.
- 3.45 Assume that the Fermi energy level is exactly in the center of the bandgap energy of a semiconductor at $T = 300$ K. (a) Calculate the probability that an energy state in the bottom of the conduction band is occupied by an electron for Si, Ge, and GaAs. (b) Calculate the probability that an energy state in the top of the valence band is empty for Si, Ge, and GaAs.
- 3.46 (a) Calculate the temperature at which there is a 10^{-8} probability that an energy state 0.60 eV above the Fermi energy level is occupied by an electron. (b) Repeat part (a) for a probability of 10^{-6} .
- 3.47 Calculate the energy range (in eV) between $f_F = 0.95$ and $f_F = 0.05$ for $E_F = 5.0$ eV at (a) $T = 200$ K and (b) $T = 400$ K.

READING LIST

1. Dimitrijev, S. *Principles of Semiconductor Devices*. New York: Oxford University, 2006.
2. Kano, K. *Semiconductor Devices*. Upper Saddle River, NJ: Prentice Hall, 1998.
3. Kittel, C. *Introduction to Solid State Physics*, 7th ed. Berlin: Springer-Verlag, 1993.
4. McKelvey, J. P. *Solid State Physics for Engineering and Materials Science*. Malabar, FL: Krieger, 1993.
5. Pierret, R. F. *Semiconductor Device Fundamentals*. Reading, MA: Addison-Wesley, 1996.
- *6. Shockley, W. *Electrons and Holes in Semiconductors*. New York: D. Van Nostrand, 1950.
7. Shur, M. *Introduction to Electronic Devices*. New York: John Wiley and Sons, 1996.
- *8. Shur, M. *Physics of Semiconductor Devices*. Englewood Cliffs, NJ: Prentice Hall, 1990.
9. Singh, J. *Semiconductor Devices: An Introduction*. New York: McGraw-Hill, 1994.
10. Singh, J. *Semiconductor Devices: Basic Principles*. New York: John Wiley and Sons, 2001.

11. Streetman, B. G., and S. K. Banerjee. *Solid State Electronic Devices*, 6th ed. Upper Saddle River, NJ: Pearson Prentice Hall, 2006.
12. Sze, S. M. *Semiconductor Devices: Physics and Technology*, 2nd ed. New York: John Wiley and Sons, 2001.
- *13. Wang, S. *Fundamentals of Semiconductor Theory and Device Physics*. Englewood Cliffs, NJ: Prentice Hall, 1988.
14. Wolfe, C. M., N. Holonyak, Jr., and G. E. Stillman. *Physical Properties of Semiconductors*. Englewood Cliffs, NJ: Prentice Hall, 1989.

*Indicates references that are at an advanced level compared to this text.

4

The Semiconductor in Equilibrium

So far, we have been considering a general crystal and applying to it the concepts of quantum mechanics in order to determine a few of the characteristics of electrons in a single-crystal lattice. In this chapter, we apply these concepts specifically to a semiconductor material. In particular, we use the density of quantum states in the conduction band and the density of quantum states in the valence band along with the Fermi–Dirac probability function to determine the concentration of electrons and holes in the conduction and valence bands, respectively. We also apply the concept of the Fermi energy to the semiconductor material.

This chapter deals with the semiconductor in equilibrium. Equilibrium, or thermal equilibrium, implies that no external forces such as voltages, electric fields, magnetic fields, or temperature gradients are acting on the semiconductor. All properties of the semiconductor will be independent of time in this case. ■

4.0 | PREVIEW

In this chapter, we will:

- Derive the thermal-equilibrium concentrations of electrons and holes in a semiconductor as a function of the Fermi energy level.
- Discuss the process by which the properties of a semiconductor material can be favorably altered by adding specific impurity atoms to the semiconductor.
- Determine the thermal-equilibrium concentrations of electrons and holes in a semiconductor as a function of the concentration of dopant atoms added to the semiconductor.
- Determine the position of the Fermi energy level as a function of the concentrations of dopant atoms added to the semiconductor.

4.1 | CHARGE CARRIERS IN SEMICONDUCTORS

Current is the rate at which charge flows. In a semiconductor, two types of charge carrier, the electron and the hole, can contribute to a current. Since the current in a semiconductor is determined largely by the number of electrons in the conduction band and the number of holes in the valence band, an important characteristic of the semiconductor is the density of these charge carriers. The density of electrons and holes is related to the density of states function and the Fermi distribution function, both of which we have considered. A qualitative discussion of these relationships will be followed by a more rigorous mathematical derivation of the thermal-equilibrium concentration of electrons and holes.

4.1.1 Equilibrium Distribution of Electrons and Holes

The distribution (with respect to energy) of electrons in the conduction band is given by the density of allowed quantum states times the probability that a state is occupied by an electron. This statement is written in equation form as

$$n(E) = g_c(E)f_F(E) \quad (4.1)$$

where $f_F(E)$ is the Fermi–Dirac probability function and $g_c(E)$ is the density of quantum states in the conduction band. The total electron concentration per unit volume in the conduction band is then found by integrating Equation (4.1) over the entire conduction-band energy.

Similarly, the distribution (with respect to energy) of holes in the valence band is the density of allowed quantum states in the valence band multiplied by the probability that a state is *not* occupied by an electron. We may express this as

$$p(E) = g_v(E)[1 - f_F(E)] \quad (4.2)$$

The total hole concentration per unit volume is found by integrating this function over the entire valence-band energy.

To find the thermal-equilibrium electron and hole concentrations, we need to determine the position of the Fermi energy E_F with respect to the bottom of the conduction-band energy E_c and the top of the valence-band energy E_v . To address this question, we will initially consider an intrinsic semiconductor. An ideal intrinsic semiconductor is a pure semiconductor with no impurity atoms and no lattice defects in the crystal (e.g., pure silicon). We have argued in the previous chapter that, for an intrinsic semiconductor at $T = 0$ K, all energy states in the valence band are filled with electrons and all energy states in the conduction band are empty of electrons. The Fermi energy must, therefore, be somewhere between E_c and E_v . (The Fermi energy does not need to correspond to an allowed energy.)

As the temperature begins to increase above 0 K, the valence electrons will gain thermal energy. A few electrons in the valence band may gain sufficient energy to jump to the conduction band. As an electron jumps from the valence band to the conduction band, an empty state, or hole, is created in the valence band. In an intrinsic

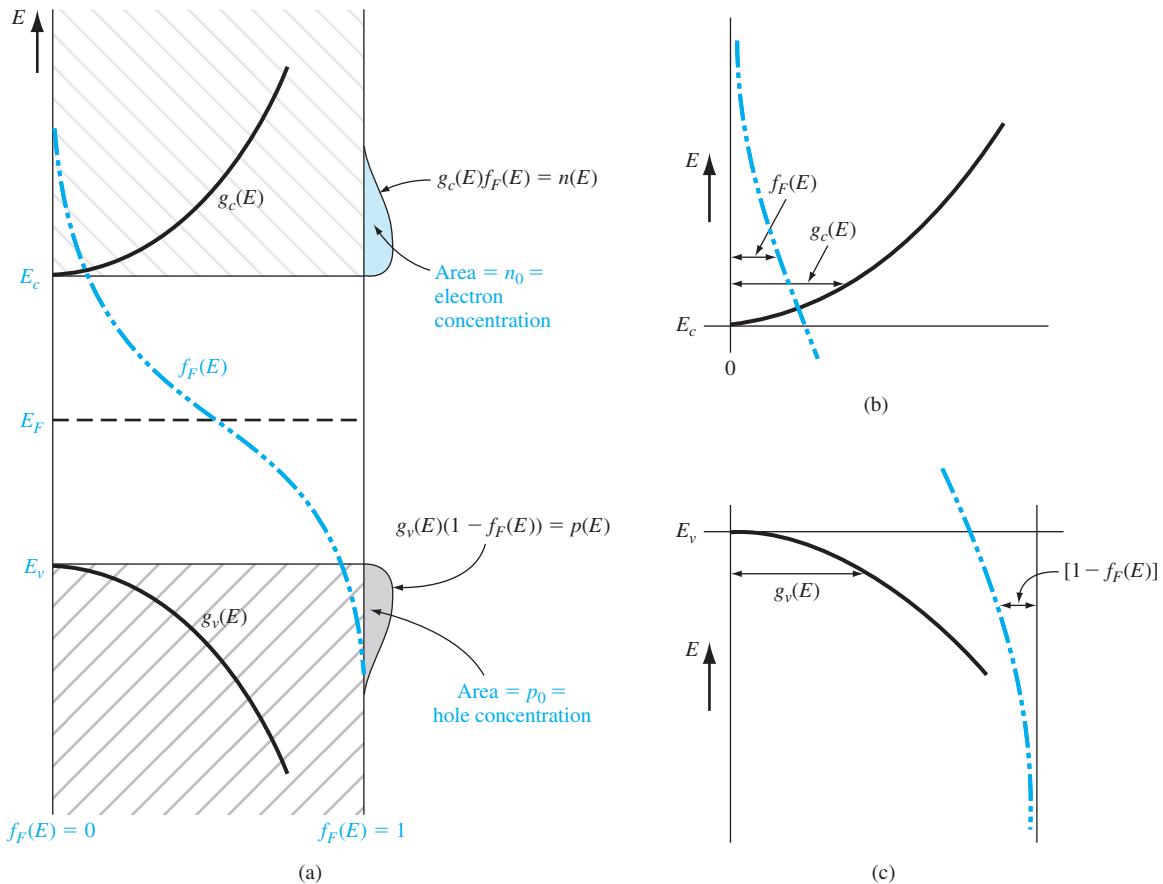


Figure 4.1 | (a) Density of states functions, Fermi–Dirac probability function, and areas representing electron and hole concentrations for the case when E_F is near the midgap energy; (b) expanded view near the conduction-band energy; and (c) expanded view near the valence-band energy.

semiconductor, then, electrons and holes are created in pairs by the thermal energy so that the number of electrons in the conduction band is equal to the number of holes in the valence band.

Figure 4.1a shows a plot of the density of states function in the conduction-band $g_c(E)$, the density of states function in the valence-band $g_v(E)$, and the Fermi–Dirac probability function for $T > 0$ K when E_F is approximately halfway between E_c and E_v . If we assume, for the moment, that the electron and hole effective masses are equal, then $g_c(E)$ and $g_v(E)$ are symmetrical functions about the midgap energy (the energy midway between E_c and E_v). We noted previously that the function $f_F(E)$ for $E > E_F$ is symmetrical to the function $1 - f_F(E)$ for $E < E_F$ about the energy $E = E_F$. This also means that the function $f_F(E)$ for $E = E_F + dE$ is equal to the function $1 - f_F(E)$ for $E = E_F - dE$.

Figure 4.1b is an expanded view of the plot in Figure 4.1a showing $f_F(E)$ and $g_c(E)$ above the conduction-band energy E_c . The product of $g_c(E)$ and $f_F(E)$ is the distribution of electrons $n(E)$ in the conduction band given by Equation (4.1). This product is plotted in Figure 4.1a. Figure 4.1c, which is an expanded view of the plot in Figure 4.1a shows $[1 - f_F(E)]$ and $g_v(E)$ below the valence-band energy E_v . The product of $g_v(E)$ and $[1 - f_F(E)]$ is the distribution of holes $p(E)$ in the valence band given by Equation (4.2). This product is also plotted in Figure 4.1a. The areas under these curves are then the total density of electrons in the conduction band and the total density of holes in the valence band. From this we see that if $g_c(E)$ and $g_v(E)$ are symmetrical, the Fermi energy must be at the midgap energy in order to obtain equal electron and hole concentrations. If the effective masses of the electron and hole are not exactly equal, then the effective density of states functions $g_c(E)$ and $g_v(E)$ will not be exactly symmetrical about the midgap energy. The Fermi level for the intrinsic semiconductor will then shift slightly from the midgap energy in order to obtain equal electron and hole concentrations.

4.1.2 The n_0 and p_0 Equations

We have argued that the Fermi energy for an intrinsic semiconductor is near midgap. In deriving the equations for the thermal-equilibrium concentration of electrons n_0 and the thermal-equilibrium concentration of holes p_0 , we will not be quite so restrictive. We will see later that, in particular situations, the Fermi energy can deviate from this midgap energy. We will assume initially, however, that the Fermi level remains within the bandgap energy.

Thermal-Equilibrium Electron Concentration The equation for the thermal-equilibrium concentration of electrons may be found by integrating Equation (4.1) over the conduction band energy, or

$$n_0 = \int g_c(E)f_F(E) dE \quad (4.3)$$

The lower limit of integration is E_c and the upper limit of integration should be the top of the allowed conduction band energy. However, since the Fermi probability function rapidly approaches zero with increasing energy as indicated in Figure 4.1a, we can take the upper limit of integration to be infinity.

We are assuming that the Fermi energy is within the forbidden-energy bandgap. For electrons in the conduction band, we have $E > E_c$. If $(E_c - E_F) \gg kT$, then $(E - E_F) \gg kT$, so that the Fermi probability function reduces to the Boltzmann approximation,¹ which is

$$f_F(E) = \frac{1}{1 + \exp \frac{(E - E_F)}{kT}} \approx \exp \left[\frac{-(E - E_F)}{kT} \right] \quad (4.4)$$

¹The Maxwell–Boltzmann and Fermi–Dirac distribution functions are within 5 percent of each other when $E - E_F \approx 3kT$ (see Figure 3.35). The \gg notation is then somewhat misleading to indicate when the Boltzmann approximation is valid, although it is commonly used.

Applying the Boltzmann approximation to Equation (4.3), the thermal-equilibrium density of electrons in the conduction band is found from

$$n_0 = \int_{E_c}^{\infty} \frac{4\pi(2m_n^*)^{3/2}}{h^3} \sqrt{E - E_c} \exp\left[\frac{-(E - E_F)}{kT}\right] dE \quad (4.5)$$

The integral of Equation (4.5) may be solved more easily by making a change of variable. If we let

$$\eta = \frac{E - E_c}{kT} \quad (4.6)$$

then Equation (4.5) becomes

$$n_0 = \frac{4\pi(2m_n^* kT)^{3/2}}{h^3} \exp\left[\frac{-(E_c - E_F)}{kT}\right] \int_0^{\infty} \eta^{1/2} \exp(-\eta) d\eta \quad (4.7)$$

The integral is the gamma function, with a value of

$$\int_0^{\infty} \eta^{1/2} \exp(-\eta) d\eta = \frac{1}{2} \sqrt{\pi} \quad (4.8)$$

Then Equation (4.7) becomes

$$n_0 = 2\left(\frac{2\pi m_n^* kT}{h^2}\right)^{3/2} \exp\left[\frac{-(E_c - E_F)}{kT}\right] \quad (4.9)$$

We may define a parameter N_c as

$$N_c = 2\left(\frac{2\pi m_n^* kT}{h^2}\right)^{3/2} \quad (4.10)$$

The parameter m_n^* is the density of states effective mass of the electron. The thermal-equilibrium electron concentration in the conduction band can be written as

$$n_0 = N_c \exp\left[\frac{-(E_c - E_F)}{kT}\right] \quad (4.11)$$

The parameter N_c is called the *effective density of states function in the conduction band*. If we were to assume that $m_n^* = m_0$, then the value of the effective density of states function at $T = 300$ K is $N_c = 2.5 \times 10^{19}$ cm⁻³, which is the order of magnitude of N_c for most semiconductors. If the effective mass of the electron is larger or smaller than m_0 , then the value of the effective density of states function changes accordingly, but is still of the same order of magnitude.

EXAMPLE 4.1

Objective: Calculate the probability that a quantum state in the conduction band at $E = E_c + kT/2$ is occupied by an electron, and calculate the thermal-equilibrium electron concentration in silicon at $T = 300$ K.

Assume the Fermi energy is 0.25 eV below the conduction band. The value of N_c for silicon at $T = 300$ K is $N_c = 2.8 \times 10^{19}$ cm⁻³ (see Appendix B).

■ Solution

The probability that a quantum state at $E = E_c + kT/2$ is occupied by an electron is given by

$$f_F(E) = \frac{1}{1 + \exp\left(\frac{E - E_F}{kT}\right)} \cong \exp\left[\frac{-(E - E_F)}{kT}\right] = \exp\left[\frac{-(E_c + (kT/2) - E_F)}{kT}\right]$$

or

$$f_F(E) = \exp\left[\frac{-(0.25 + (0.0259/2))}{0.0259}\right] = 3.90 \times 10^{-5}$$

The electron concentration is given by

$$n_0 = N_c \exp\left[\frac{-(E_c - E_F)}{kT}\right] = (2.8 \times 10^{19}) \exp\left[\frac{-0.25}{0.0259}\right]$$

or

$$n_0 = 1.80 \times 10^{15} \text{ cm}^{-3}$$

Comment

The probability of a state being occupied can be quite small, but the fact that there are a large number of states means that the electron concentration is a reasonable value.

EXERCISE PROBLEM

Ex 4.1 Determine the probability that a quantum state at energy $E = E_c + kT$ is occupied by an electron, and calculate the electron concentration in GaAs at $T = 300$ K if the Fermi energy level is 0.25 eV below E_c .

$$[\text{Ans. } f_F(E) = 2.36 \times 10^{-5}, n_0 = 3.02 \times 10^{15} \text{ cm}^{-3}]$$

Thermal-Equilibrium Hole Concentration The thermal-equilibrium concentration of holes in the valence band is found by integrating Equation (4.2) over the valence-band energy, or

$$p_0 = \int g_v(E)[1 - f_F(E)] dE \quad (4.12)$$

We may note that

$$1 - f_F(E) = \frac{1}{1 + \exp\left(\frac{E_F - E}{kT}\right)} \quad (4.13a)$$

For energy states in the valence band, $E < E_v$. If $(E_F - E_v) \gg kT$ (the Fermi function is still assumed to be within the bandgap), then we have a slightly different form of the Boltzmann approximation. Equation (4.13a) may be written as

$$1 - f_F(E) = \frac{1}{1 + \exp\left(\frac{E_F - E}{kT}\right)} \approx \exp\left[\frac{-(E_F - E)}{kT}\right] \quad (4.13b)$$

Applying the Boltzmann approximation of Equation (4.13b) to Equation (4.12), we find the thermal-equilibrium concentration of holes in the valence band is

$$p_0 = \int_{-\infty}^{E_v} \frac{4\pi (2m_p^*)^{3/2}}{h^3} \sqrt{E_v - E} \exp\left[\frac{-(E_F - E)}{kT}\right] dE \quad (4.14)$$

where the lower limit of integration is taken as minus infinity instead of the bottom of the valence band. The exponential term decays fast enough so that this approximation is valid.

Equation (4.14) may be solved more easily by again making a change of variable. If we let

$$\eta' = \frac{E_v - E}{kT} \quad (4.15)$$

then Equation (4.14) becomes

$$p_0 = \frac{-4\pi(2m_p^* kT)^{3/2}}{h^3} \exp\left[\frac{-(E_F - E_v)}{kT}\right] \int_{+\infty}^0 (\eta')^{1/2} \exp(-\eta') d\eta' \quad (4.16)$$

where the negative sign comes from the differential $dE = -kT d\eta'$. Note that the lower limit of η' becomes $+\infty$ when $E = -\infty$. If we change the order of integration, we introduce another minus sign. From Equation (4.8), Equation (4.16) becomes

$$p_0 = 2 \left(\frac{2\pi m_p^* kT}{h^2} \right)^{3/2} \exp\left[\frac{-(E_F - E_v)}{kT}\right] \quad (4.17)$$

We may define a parameter N_v as

$$N_v = 2 \left(\frac{2\pi m_p^* kT}{h^2} \right)^{3/2} \quad (4.18)$$

which is called the *effective density of states function in the valence band*. The parameter m_p^* is the density of states effective mass of the hole. The thermal-equilibrium concentration of holes in the valence band may now be written as

$$p_0 = N_v \exp\left[\frac{-(E_F - E_v)}{kT}\right] \quad (4.19)$$

The magnitude of N_v is also on the order of 10^{19} cm $^{-3}$ at $T = 300$ K for most semiconductors.

EXAMPLE 4.2

Objective: Calculate the thermal-equilibrium hole concentration in silicon at $T = 400$ K.

Assume that the Fermi energy is 0.27 eV above the valence-band energy. The value of N_v for silicon at $T = 300$ K is $N_v = 1.04 \times 10^{19}$ cm $^{-3}$. (See Appendix B)

■ Solution

The parameter values at $T = 400$ K are found as:

$$N_v = (1.04 \times 10^{19}) \left(\frac{400}{300} \right)^{3/2} = 1.60 \times 10^{19} \text{ cm}^{-3}$$

and

$$kT = (0.0259) \left(\frac{400}{300} \right) = 0.03453 \text{ eV}$$

The hole concentration is then

$$p_0 = N_v \exp\left[\frac{-(E_F - E_v)}{kT}\right] = (1.60 \times 10^{19}) \exp\left(\frac{-0.27}{0.03453}\right)$$

or

$$p_0 = 6.43 \times 10^{15} \text{ cm}^{-3}$$

Comment

The parameter values at any temperature can easily be found by using the 300 K values and the temperature dependence.

EXERCISE PROBLEM

- Ex 4.2** (a) Repeat Example 4.2 at $T = 250$ K. (b) What is the ratio of p_0 at $T = 250$ K to that at $T = 400$ K? [Ans. (a) $p_0 = 2.92 \times 10^{13} \text{ cm}^{-3}$; (b) 4.54×10^{-3}]

The effective density of states functions, N_c and N_v , are constant for a given semiconductor material at a fixed temperature. Table 4.1 gives the values of the density of states function and of the density of states effective masses for silicon, gallium arsenide, and germanium. Note that the value of N_c for gallium arsenide is smaller than the typical 10^{19} cm^{-3} value. This difference is due to the small electron effective mass in gallium arsenide.

The thermal-equilibrium concentrations of electrons in the conduction band and of holes in the valence band are directly related to the effective density of states constants and to the Fermi energy level.

TEST YOUR UNDERSTANDING

- TYU 4.1** Calculate the thermal equilibrium electron and hole concentration in silicon at $T = 300$ K for the case when the Fermi energy level is 0.22 eV below the conduction-band energy E_c . The value of E_g is given in Appendix B.4.
[Ans. $n_0 = 5.73 \times 10^{15} \text{ cm}^{-3}$, $p_0 = 8.43 \times 10^{15} \text{ cm}^{-3}$)

- TYU 4.2** Determine the thermal equilibrium electron and hole concentration in GaAs at $T = 300$ K for the case when the Fermi energy level is 0.30 eV above the valence-band energy E_v . The value of E_g is given in Appendix B.4.
[Ans. $n_0 = 0.0779 \text{ cm}^{-3}$, $p_0 = 9.53 \times 10^{15} \text{ cm}^{-3}$)

4.1.3 The Intrinsic Carrier Concentration

For an intrinsic semiconductor, the concentration of electrons in the conduction band is equal to the concentration of holes in the valence band. We may denote n_i and p_i as the electron and hole concentrations, respectively, in the intrinsic semiconductor. These parameters are usually referred to as the intrinsic electron concentration and intrinsic hole concentration. However, $n_i = p_i$, so normally we simply use the parameter n_i as the intrinsic carrier concentration, which refers to either the intrinsic electron or hole concentration.

Table 4.1 | Effective density of states function and density of states effective mass values

	$N_c (\text{cm}^{-3})$	$N_v (\text{cm}^{-3})$	m_n^*/m_0	m_p^*/m_0
Silicon	2.8×10^{19}	1.04×10^{19}	1.08	0.56
Gallium arsenide	4.7×10^{17}	7.0×10^{18}	0.067	0.48
Germanium	1.04×10^{19}	6.0×10^{18}	0.55	0.37

The Fermi energy level for the intrinsic semiconductor is called the intrinsic Fermi energy, or $E_F = E_{Fi}$. If we apply Equations (4.11) and (4.19) to the intrinsic semiconductor, then we can write

$$n_0 = n_i = N_c \exp \left[\frac{-(E_c - E_{Fi})}{kT} \right] \quad (4.20)$$

and

$$p_0 = p_i = n_i = N_v \exp \left[\frac{-(E_{Fi} - E_v)}{kT} \right] \quad (4.21)$$

If we take the product of Equations (4.20) and (4.21), we obtain

$$n_i^2 = N_c N_v \exp \left[\frac{-(E_c - E_{Fi})}{kT} \right] \cdot \exp \left[\frac{-(E_{Fi} - E_v)}{kT} \right] \quad (4.22)$$

or

$$n_i^2 = N_c N_v \exp \left[\frac{-(E_c - E_v)}{kT} \right] = N_c N_v \exp \left[\frac{-E_g}{kT} \right] \quad (4.23)$$

where E_g is the bandgap energy. For a given semiconductor material at a constant temperature, the value of n_i is a constant, and independent of the Fermi energy.

The intrinsic carrier concentration for silicon at $T = 300$ K may be calculated by using the effective density of states function values from Table 4.1. The value of n_i calculated from Equation (4.23) for $E_g = 1.12$ eV is $n_i = 6.95 \times 10^9 \text{ cm}^{-3}$. The commonly accepted value² of n_i for silicon at $T = 300$ K is approximately $1.5 \times 10^{10} \text{ cm}^{-3}$. This discrepancy may arise from several sources. First, the values of the effective masses are determined at a low temperature where the cyclotron resonance experiments are performed. Since the effective mass is an experimentally determined parameter, and since the effective mass is a measure of how well a particle moves in a crystal, this parameter may be a slight function of temperature. Next, the density of states function for a semiconductor was obtained by generalizing the model of an electron in a three-dimensional infinite potential well. This theoretical function may also not agree exactly with experiment. However, the difference between the theoretical value and the experimental value of n_i is approximately a factor of 2, which, in many cases, is not significant. Table 4.2 lists the commonly accepted values of n_i for silicon, gallium arsenide, and germanium at $T = 300$ K.

The intrinsic carrier concentration is a very strong function of temperature.

Table 4.2 | Commonly accepted values of n_i at $T \equiv 300$ K

Silicon	$n_i = 1.5 \times 10^{10} \text{ cm}^{-3}$
Gallium arsenide	$n_i = 1.8 \times 10^6 \text{ cm}^{-3}$
Germanium	$n_i = 2.4 \times 10^{13} \text{ cm}^{-3}$

²Various references may list slightly different values of the intrinsic silicon concentration at room temperature. In general, they are all between 1×10^{10} and $1.5 \times 10^{10} \text{ cm}^{-3}$. This difference is, in most cases, not significant.

Objective: Calculate the intrinsic carrier concentration in silicon at $T = 250$ K and at $T = 400$ K.

The values of N_c and N_v for silicon at $T = 300$ K are $2.8 \times 10^{19} \text{ cm}^{-3}$ and $1.04 \times 10^{19} \text{ cm}^{-3}$, respectively. Both N_c and N_v vary as $T^{3/2}$. Assume the bandgap energy of silicon is 1.12 eV and does not vary over this temperature range.

■ Solution

Using Equation (4.23), we find, at $T = 250$ K

$$\begin{aligned} n_i^2 &= (2.8 \times 10^{19})(1.04 \times 10^{19})\left(\frac{250}{300}\right)^3 \exp\left[\frac{-1.12}{(0.0259)(250/300)}\right] \\ &= 4.90 \times 10^{15} \end{aligned}$$

or

$$n_i = 7.0 \times 10^7 \text{ cm}^{-3}$$

At $T = 400$ K, we find

$$\begin{aligned} n_i^2 &= (2.8 \times 10^{19})(1.04 \times 10^{19})\left(\frac{400}{300}\right)^3 \exp\left[\frac{-1.12}{(0.0259)(400/300)}\right] \\ &= 5.67 \times 10^{24} \end{aligned}$$

or

$$n_i = 2.38 \times 10^{12} \text{ cm}^{-3}$$

■ Comment

We may note from this example that the intrinsic carrier concentration increased by over 4 orders of magnitude as the temperature increased by 150°C.

■ EXERCISE PROBLEM

- Ex 4.3** (a) Calculate the intrinsic carrier concentration in GaAs at $T = 400$ K and at $T = 250$ K. Assume that $E_g = 1.42$ eV is constant over this temperature range.
 (b) What is the ratio of n_i at $T = 400$ K to that at $T = 250$ K?

[Ans. (a) $n_i(400) = 3.29 \times 10^9 \text{ cm}^{-3}$, $n_i(250) = 7.13 \times 10^8 \text{ cm}^{-3}$; (b) $4.61 \times 10_5$]

Figure 4.2 is a plot of n_i from Equation (4.23) for silicon, gallium arsenide, and germanium as a function of temperature. As seen in the figure, the value of n_i for these semiconductors may easily vary over several orders of magnitude as the temperature changes over a reasonable range.

TEST YOUR UNDERSTANDING

- TYU 4.3** Calculate the intrinsic concentration in silicon at (a) $T = 200$ K and (b) $T = 450$ K. (c) Determine the ratio of n_i at $T = 450$ K to that at $T = 200$ K.
 [Ans. (a) $n_i = 7.63 \times 10^4 \text{ cm}^{-3}$; (b) $n_i = 1.72 \times 10^{15} \text{ cm}^{-3}$; (c) $2.26 \times 10_8$]

- TYU 4.4** Repeat TYU 4.3 for GaAs.
 [Ans. (a) $n_i = 1.37 \text{ cm}^{-3}$; (b) $n_i = 3.85 \times 10^{10} \text{ cm}^{-3}$; (c) $2.81 \times 10_{10}$]

- TYU 4.5** Repeat TYU 4.3 for Ge.
 [Ans. (a) $n_i = 2.15 \times 10^{10} \text{ cm}^{-3}$; (b) $n_i = 2.97 \times 10^{15} \text{ cm}^{-3}$; (c) $1.38 \times 10_5$]

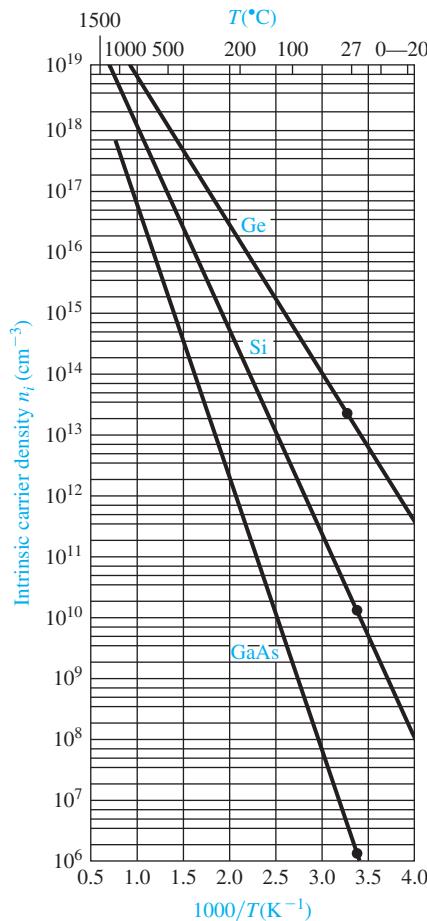


Figure 4.2 | The intrinsic carrier concentration of Ge, Si, and GaAs as a function of temperature.
(From Sze [14].)

4.1.4 The Intrinsic Fermi-Level Position

We have qualitatively argued that the Fermi energy level is located near the center of the forbidden bandgap for the intrinsic semiconductor. We can specifically calculate the intrinsic Fermi-level position. Since the electron and hole concentrations are equal, setting Equations (4.20) and (4.21) equal to each other, we have

$$N_c \exp\left[\frac{-(E_c - E_{Fi})}{kT}\right] = N_v \exp\left[\frac{-(E_{Fi} - E_v)}{kT}\right] \quad (4.24)$$

If we take the natural log of both sides of this equation and solve for E_{Fi} , we obtain

$$E_{Fi} = \frac{1}{2}(E_c + E_v) + \frac{1}{2}kT \ln\left(\frac{N_v}{N_c}\right) \quad (4.25)$$

From the definitions for N_c and N_v given by Equations (4.10) and (4.18), respectively, Equation (4.25) may be written as

$$E_{Fi} = \frac{1}{2}(E_c + E_v) + \frac{3}{4}kT \ln\left(\frac{m_p^*}{m_n^*}\right) \quad (4.26a)$$

The first term, $\frac{1}{2}(E_c + E_v)$, is the energy exactly midway between E_c and E_v , or the midgap energy. We can define

$$\frac{1}{2}(E_c + E_v) = E_{\text{midgap}}$$

so that

$$E_{Fi} - E_{\text{midgap}} = \frac{3}{4}kT \ln\left(\frac{m_p^*}{m_n^*}\right) \quad (4.26b)$$

If the electron and hole effective masses are equal so that $m_p^* = m_n^*$, then the intrinsic Fermi level is exactly in the center of the bandgap. If $m_p^* > m_n^*$, the intrinsic Fermi level is slightly above the center, and if $m_p^* < m_n^*$, it is slightly below the center of the bandgap. The density of states function is directly related to the carrier effective mass; thus, a larger effective mass means a larger density of states function. The intrinsic Fermi level must shift away from the band with the larger density of states in order to maintain equal numbers of electrons and holes.

Objective: Calculate the position of the intrinsic Fermi level with respect to the center of the bandgap in silicon at $T = 300$ K.

EXAMPLE 4.4

The density of states effective carrier masses in silicon are $m_n^* = 1.08m_0$ and $m_p^* = 0.56m_0$.

■ Solution

The intrinsic Fermi level with respect to the center of the bandgap is

$$E_{Fi} - E_{\text{midgap}} = \frac{3}{4}kT \ln\left(\frac{m_p^*}{m_n^*}\right) = \frac{3}{4}(0.0259) \ln\left(\frac{0.56}{1.08}\right)$$

or

$$E_{Fi} - E_{\text{midgap}} = -0.0128 \text{ eV} = -12.8 \text{ meV}$$

■ Comment

The intrinsic Fermi level in silicon is 12.8 meV below the midgap energy. If we compare 12.8 meV to 560 meV, which is one-half of the bandgap energy of silicon, we can, in many applications, simply approximate the intrinsic Fermi level to be in the center of the bandgap.

■ EXERCISE PROBLEM

Ex 4.4 Determine the position of the intrinsic Fermi level at $T = 300$ K with respect to the center of the bandgap for (a) GaAs and (b) Ge. [Ans. (a) $+38.25 \text{ meV}$; (b) -7.70 meV]

TEST YOUR UNDERSTANDING

- TYU 4.6** Determine the position of the intrinsic Fermi level with respect to the center of the bandgap in silicon at (a) $T = 200\text{ K}$ and (b) $T = 400\text{ K}$. Assume the effective masses are constant over this temperature range. [Ans: (a) -17.01 meV ; (b) -17.50 meV]

4.2 | DOPANT ATOMS AND ENERGY LEVELS

The intrinsic semiconductor may be an interesting material, but the real power of semiconductors is realized by adding small, controlled amounts of specific dopant, or impurity, atoms. This doping process, described briefly in Chapter 1, can greatly alter the electrical characteristics of the semiconductor. The doped semiconductor, called an *extrinsic* material, is the primary reason we can fabricate the various semiconductor devices that we will consider in later chapters.

4.2.1 Qualitative Description

In Chapter 3, we discussed the covalent bonding of silicon and considered the simple two-dimensional representation of the single-crystal silicon lattice as shown in Figure 4.3. Now consider adding a group V element, such as phosphorus, as a substitutional impurity. The group V element has five valence electrons. Four of these will contribute to the covalent bonding with the silicon atoms, leaving the fifth more loosely bound to the phosphorus atom. This effect is schematically shown in Figure 4.4. We refer to the fifth valence electron as a donor electron.

The phosphorus atom without the donor electron is positively charged. At very low temperatures, the donor electron is bound to the phosphorus atom. However, by intuition, it should seem clear that the energy required to elevate the donor electron into the conduction band is considerably less than that for the electrons involved in the covalent bonding. Figure 4.5 shows the energy-band diagram that we would expect. The energy level, E_d , is the energy state of the donor electron.

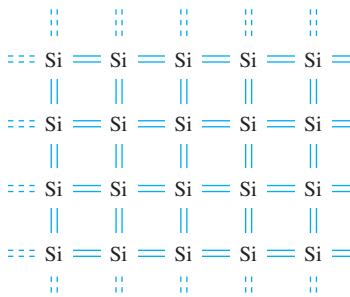


Figure 4.3 | Two-dimensional representation of the intrinsic silicon lattice.

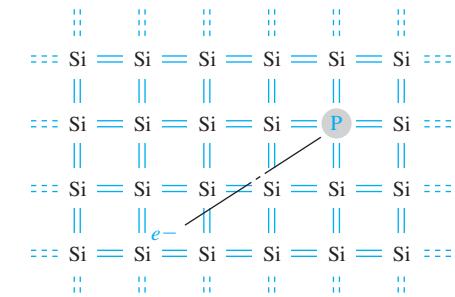


Figure 4.4 | Two-dimensional representation of the silicon lattice doped with a phosphorus atom.

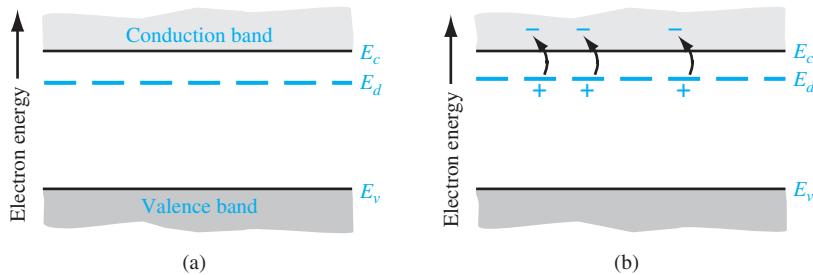


Figure 4.5 | The energy-band diagram showing (a) the discrete donor energy state and (b) the effect of a donor state being ionized.

If a small amount of energy, such as thermal energy, is added to the donor electron, it can be elevated into the conduction band, leaving behind a positively charged phosphorus ion. The electron in the conduction band can now move through the crystal generating a current, while the positively charged ion is fixed in the crystal. This type of impurity atom donates an electron to the conduction band and so is called a *donor impurity atom*. The donor impurity atoms add electrons to the conduction band without creating holes in the valence band. The resulting material is referred to as an *n-type* semiconductor (*n* for the negatively charged electron).

Now consider adding a group III element, such as boron, as a substitutional impurity to silicon. The group III element has three valence electrons, which are all taken up in the covalent bonding. As shown in Figure 4.6a, one covalent bonding position appears to be empty. If an electron were to occupy this “empty” position, its energy would have to be greater than that of the valence electrons, since the net charge state of the boron atom would now be negative. However, the electron occupying this “empty” position does not have sufficient energy to be in the conduction band, so its energy is far smaller than the conduction-band energy. Figure 4.6b shows how valence electrons may gain a small amount of thermal energy and move about in

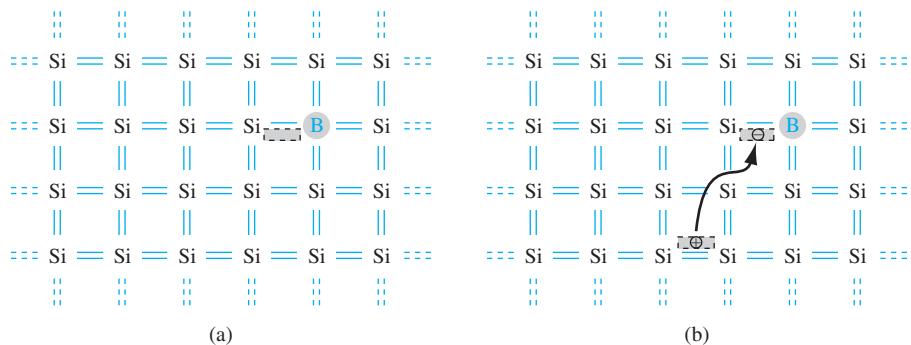


Figure 4.6 | Two-dimensional representation of a silicon lattice (a) doped with a boron atom and (b) showing the ionization of the boron atom resulting in a hole.

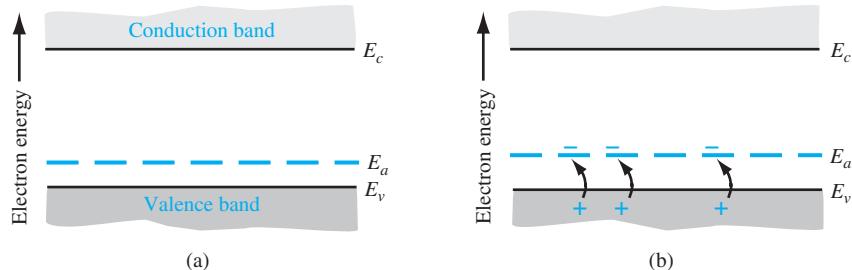


Figure 4.7 | The energy-band diagram showing (a) the discrete acceptor energy state and (b) the effect of an acceptor state being ionized.

the crystal. The “empty” position associated with the boron atom becomes occupied, and other valence electron positions become vacated. These other vacated electron positions can be thought of as holes in the semiconductor material.

Figure 4.7 shows the expected energy state of the “empty” position and also the formation of a hole in the valence band. The hole can move through the crystal generating a current, while the negatively charged boron atom is fixed in the crystal. The group III atom accepts an electron from the valence band and so is referred to as an *acceptor impurity atom*. The acceptor atom can generate holes in the valence band without generating electrons in the conduction band. This type of semiconductor material is referred to as a *p type* material (*p* for the positively charged hole).

The pure single-crystal semiconductor material is called an intrinsic material. Adding controlled amounts of dopant atoms, either donors or acceptors, creates a material called an *extrinsic semiconductor*. An extrinsic semiconductor will have either a preponderance of electrons (*n* type) or a preponderance of holes (*p* type).

4.2.2 Ionization Energy

We can calculate the approximate distance of the donor electron from the donor impurity ion, and also the approximate energy required to elevate the donor electron into the conduction band. This energy is referred to as the ionization energy. We will use the Bohr model of the atom for these calculations. The justification for using this model is that the most probable distance of an electron from the nucleus in a hydrogen atom, determined from quantum mechanics, is the same as the Bohr radius. The energy levels in the hydrogen atom determined from quantum mechanics are also the same as obtained from the Bohr theory.

In the case of the donor impurity atom, we may visualize the donor electron orbiting the donor ion, which is embedded in the semiconductor material. We will need to use the permittivity of the semiconductor material in the calculations rather than the permittivity of free space as is used in the case of the hydrogen atom. We will also use the effective mass of the electron in the calculations.

The analysis begins by setting the coulomb force of attraction between the electron and ion equal to the centripetal force of the orbiting electron. This condition will give a steady orbit. We have

$$\frac{e^2}{4\pi \epsilon r_n^2} = \frac{m^* v^2}{r_n} \quad (4.27)$$

where v is the magnitude of the velocity and r_n is the radius of the orbit. If we assume the angular momentum is also quantized, then we can write

$$m^* r_n v = n \hbar \quad (4.28)$$

where n is a positive integer. Solving for v from Equation (4.28), substituting into Equation (4.27), and solving for the radius, we obtain

$$r_n = \frac{n^2 \hbar^2 4\pi \epsilon}{m^* e^2} \quad (4.29)$$

The assumption of the angular momentum being quantized leads to the radius also being quantized.

The Bohr radius is defined as

$$a_0 = \frac{4\pi\epsilon_0 \hbar^2}{m_0 e^2} = 0.53 \text{ \AA} \quad (4.30)$$

We can normalize the radius of the donor orbital to that of the Bohr radius, which gives

$$\frac{r_n}{a_0} = n^2 \epsilon_r \left(\frac{m_0}{m^*} \right) \quad (4.31)$$

where ϵ_r is the relative dielectric constant of the semiconductor material, m_0 is the rest mass of an electron, and m^* is the conductivity effective mass of the electron in the semiconductor.³

If we consider the lowest energy state in which $n = 1$, and if we consider silicon in which $\epsilon_r = 11.7$ and the conductivity effective mass is $m^*/m_0 = 0.26$, then we have that

$$\frac{r_1}{a_0} = 45 \quad (4.32)$$

or $r_1 = 23.9 \text{ \AA}$. This radius corresponds to approximately four lattice constants of silicon. Recall that one unit cell in silicon effectively contains eight atoms, so the radius of the orbiting donor electron encompasses many silicon atoms. The donor electron is not tightly bound to the donor atom.

The total energy of the orbiting electron is given by

$$E = T + V \quad (4.33)$$

where T is the kinetic energy and V is the potential energy of the electron. The kinetic energy is

$$T = \frac{1}{2} m^* v^2 \quad (4.34)$$

³The conductivity effective mass is used when electrons and holes are in motion. See Appendix F for a discussion of effective mass concepts.

Using the velocity v from Equation (4.28) and the radius r_n from Equation (4.29), the kinetic energy becomes

$$T = \frac{m^* e^4}{2(n\hbar)^2(4\pi\epsilon)^2} \quad (4.35)$$

The potential energy is

$$V = \frac{-e^2}{4\pi\epsilon r_n} = \frac{-m^* e^4}{(n\hbar)^2(4\pi\epsilon)^2} \quad (4.36)$$

The total energy is the sum of the kinetic and potential energies, so that

$$E = T + V = \frac{-m^* e^4}{2(n\hbar)^2(4\pi\epsilon)^2} \quad (4.37)$$

For the hydrogen atom, $m^* = m_0$ and $\epsilon = \epsilon_0$. The ionization energy of the hydrogen atom in the lowest energy state is then $E = -13.6$ eV. If we consider silicon, the ionization energy is $E = -25.8$ meV, much less than the bandgap energy of silicon. This energy is the approximate ionization energy of the donor atom, or the energy required to elevate the donor electron into the conduction band.

For ordinary donor impurities such as phosphorus or arsenic in silicon or germanium, this hydrogenic model works quite well and gives some indication of the magnitudes of the ionization energies involved. Table 4.3 lists the actual experimentally measured ionization energies for a few impurities in silicon and germanium. Germanium and silicon have different relative dielectric constants and effective masses; thus, we expect the ionization energies to differ.

4.2.3 Group III–V Semiconductors

In the previous sections, we have discussed the donor and acceptor impurities in a group IV semiconductor, such as silicon. The situation in the group III–V compound semiconductors, such as gallium arsenide, is more complicated. Group II elements, such as beryllium, zinc, and cadmium, can enter the lattice as substitutional impurities, replacing the group III gallium element to become acceptor impurities. Similarly, group VI elements, such as selenium and tellurium, can enter the lattice substitutionally, replacing the group V arsenic element to become donor impurities. The corresponding ionization energies for these impurities are smaller than those for the impurities in silicon. The ionization energies for the donors in gallium arsenide

Table 4.3 | Impurity ionization energies in silicon and germanium

Impurity	Ionization energy (eV)	
	Si	Ge
<i>Donors</i>		
Phosphorus	0.045	0.012
Arsenic	0.05	0.0127
<i>Acceptors</i>		
Boron	0.045	0.0104
Aluminum	0.06	0.0102

Table 4.4 | Impurity ionization energies in gallium arsenide

Impurity	Ionization energy (eV)
<i>Donors</i>	
Selenium	0.0059
Tellurium	0.0058
Silicon	0.0058
Germanium	0.0061
<i>Acceptors</i>	
Beryllium	0.028
Zinc	0.0307
Cadmium	0.0347
Silicon	0.0345
Germanium	0.0404

are also smaller than those for the acceptors, because of the smaller effective mass of the electron compared to that of the hole.

Group IV elements, such as silicon and germanium, can also be impurity atoms in gallium arsenide. If a silicon atom replaces a gallium atom, the silicon impurity will act as a donor, but if the silicon atom replaces an arsenic atom, then the silicon impurity will act as an acceptor. The same is true for germanium as an impurity atom. Such impurities are called *amphoteric*. Experimentally in gallium arsenide, it is found that germanium is predominantly an acceptor and silicon is predominantly a donor. Table 4.4 lists the ionization energies for the various impurity atoms in gallium arsenide.

TEST YOUR UNDERSTANDING

- TYU 4.7** (a) Calculate the ionization energy and the radius (normalized to the Bohr radius) of a donor electron in its lowest energy state in GaAs. (b) Repeat part (a) for Ge.

$$[\text{Ans. (a)} = 19.5 \text{ eV}, r/a_0 = 13.5; \text{(b)} = 9.7 \text{ eV}, r/a_0 = 11.5]$$

4.3 | THE EXTRINSIC SEMICONDUCTOR

We defined an intrinsic semiconductor as a material with no impurity atoms present in the crystal. An *extrinsic semiconductor* is defined as a semiconductor in which controlled amounts of specific dopant or impurity atoms have been added so that the thermal-equilibrium electron and hole concentrations are different from the intrinsic carrier concentration. One type of carrier will predominate in an extrinsic semiconductor.

4.3.1 Equilibrium Distribution of Electrons and Holes

Adding donor or acceptor impurity atoms to a semiconductor will change the distribution of electrons and holes in the material. Since the Fermi energy is related to the distribution function, the Fermi energy will change as dopant atoms are added.

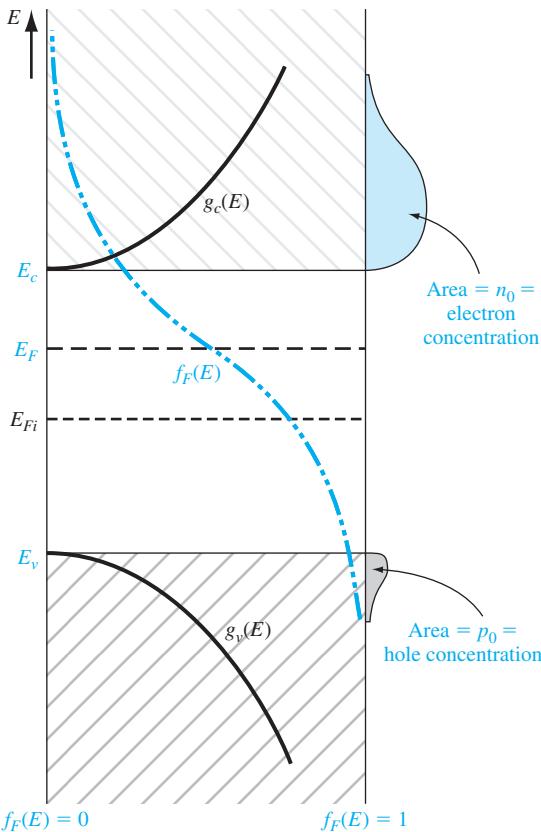


Figure 4.8 | Density of states functions, Fermi–Dirac probability function, and areas representing electron and hole concentrations for the case when E_F is above the intrinsic Fermi energy.

If the Fermi energy changes from near the midgap value, the density of electrons in the conduction band and the density of holes in the valence band will change. These effects are shown in Figures 4.8 and 4.9. Figure 4.8 shows the case for $E_F > E_{Fi}$ and Figure 4.9 shows the case for $E_F < E_{Fi}$. When $E_F > E_{Fi}$, the electron concentration is larger than the hole concentration, and when $E_F < E_{Fi}$, the hole concentration is larger than the electron concentration. When the density of electrons is greater than the density of holes, the semiconductor is n type; donor impurity atoms have been added. When the density of holes is greater than the density of electrons, the semiconductor is p type; acceptor impurity atoms have been added. The Fermi energy level in a semiconductor changes as the electron and hole concentrations change and, again, the Fermi energy changes as donor or acceptor impurities are added. The change in the Fermi level as a function of impurity concentrations is considered in Section 4.6.

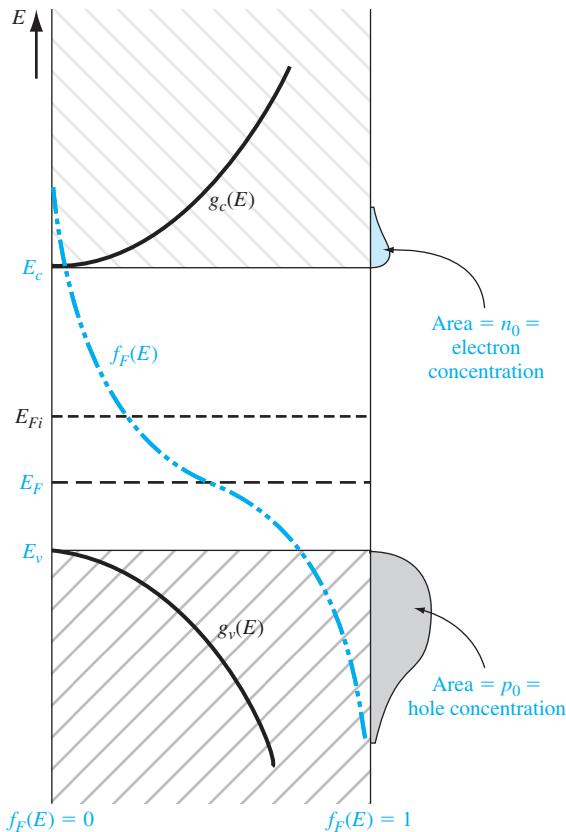


Figure 4.9 | Density of states functions, Fermi–Dirac probability function, and areas representing electron and hole concentrations for the case when E_F is below the intrinsic Fermi energy.

The expressions previously derived for the thermal-equilibrium concentration of electrons and holes, given by Equations (4.11) and (4.19), are general equations for n_0 and p_0 in terms of the Fermi energy. These equations are again given as

$$n_0 = N_c \exp \left[\frac{-(E_c - E_F)}{kT} \right]$$

and

$$p_0 = N_v \exp \left[\frac{-(E_F - E_v)}{kT} \right]$$

As we just discussed, the Fermi energy may vary through the bandgap energy, which will then change the values of n_0 and p_0 .

EXAMPLE 4.5

Objective: Calculate the thermal equilibrium concentrations of electrons and holes for a given Fermi energy.

Consider silicon at $T = 300$ K so that $N_c = 2.8 \times 10^{19} \text{ cm}^{-3}$ and $N_v = 1.04 \times 10^{19} \text{ cm}^{-3}$. Assume that the Fermi energy is 0.25 eV below the conduction band. If we assume that the bandgap energy of silicon is 1.12 eV, then the Fermi energy will be 0.87 eV above the valence band.

■ **Solution**

Using Equation (4.11), we have

$$n_0 = (2.8 \times 10^{19}) \exp\left(\frac{-0.25}{0.0259}\right) = 1.8 \times 10^{15} \text{ cm}^{-3}$$

From Equation (4.19), we can write

$$p_0 = (1.04 \times 10^{19}) \exp\left(\frac{-0.87}{0.0259}\right) = 2.7 \times 10^4 \text{ cm}^{-3}$$

■ **Comment**

The change in the Fermi level is actually a function of the donor or acceptor impurity concentrations that are added to the semiconductor. However, this example shows that electron and hole concentrations change by orders of magnitude from the intrinsic carrier concentration as the Fermi energy changes by a few tenths of an electron-volt.

■ **EXERCISE PROBLEM**

Ex 4.5 Determine the thermal-equilibrium concentrations of electrons and holes in silicon at $T = 300$ K if the Fermi energy level is 0.215 eV above the valence-band energy
 (Ans. $p_0 = 2.58 \times 10^{15} \text{ cm}^{-3}$, $n_0 = 1.87 \times 10^4 \text{ cm}^{-3}$)

In the previous example, since $n_0 > p_0$, the semiconductor is n type. In an n-type semiconductor, electrons are referred to as the majority carrier and holes as the minority carrier. By comparing the relative values of n_0 and p_0 in the example, it is easy to see how this designation came about. Similarly, in a p-type semiconductor where $p_0 > n_0$, holes are the majority carrier and electrons are the minority carrier.

We may derive another form of the equations for the thermal-equilibrium concentrations of electrons and holes. If we add and subtract an intrinsic Fermi energy in the exponent of Equation (4.11), we can write

$$n_0 = N_c \exp\left[\frac{-(E_c - E_{Fi}) + (E_F - E_{Fi})}{kT}\right] \quad (4.38a)$$

or

$$n_0 = N_c \exp\left[\frac{-(E_c - E_{Fi})}{kT}\right] \exp\left[\frac{(E_F - E_{Fi})}{kT}\right] \quad (4.38b)$$

The intrinsic carrier concentration is given by Equation (4.20) as

$$n_i = N_c \exp\left[\frac{-(E_c - E_{Fi})}{kT}\right]$$

so that the thermal-equilibrium electron concentration can be written as

$$n_0 = n_i \exp\left[\frac{E_F - E_{Fi}}{kT}\right] \quad (4.39)$$

Similarly, if we add and subtract an intrinsic Fermi energy in the exponent of Equation (4.19), we will obtain

$$p_0 = n_i \exp\left[\frac{-(E_F - E_{Fi})}{kT}\right] \quad (4.40)$$

As we will see, the Fermi level changes when donors and acceptors are added, but Equations (4.39) and (4.40) show that, as the Fermi level changes from the intrinsic Fermi level, n_0 and p_0 change from the n_i value. If $E_F > E_{Fi}$, then we will have $n_0 > n_i$ and $p_0 < n_i$. One characteristic of an n-type semiconductor is that $E_F > E_{Fi}$ so that $n_0 > p_0$. Similarly, in a p-type semiconductor, $E_F < E_{Fi}$ so that $p_0 > n_i$ and $n_0 < n_i$; thus, $p_0 > n_0$.

We can see the functional dependence of n_0 and p_0 with E_F in Figures 4.8 and 4.9. As E_F moves above or below E_{Fi} , the overlapping probability function with the density of states functions in the conduction band and valence band changes. As E_F moves above E_{Fi} , the probability function in the conduction band increases, while the probability, $1 - f_F(E)$, of an empty state (hole) in the valence band decreases. As E_F moves below E_{Fi} , the opposite occurs.

4.3.2 The $n_0 p_0$ Product

We may take the product of the general expressions for n_0 and p_0 as given in Equations (4.11) and (4.19), respectively. The result is

$$n_0 p_0 = N_c N_v \exp\left[\frac{-(E_c - E_F)}{kT}\right] \exp\left[\frac{-(E_F - E_v)}{kT}\right] \quad (4.41)$$

which may be written as

$$n_0 p_0 = N_c N_v \exp\left[\frac{-E_g}{kT}\right] \quad (4.42)$$

As Equation (4.42) was derived for a general value of Fermi energy, the values of n_0 and p_0 are not necessarily equal. However, Equation (4.42) is exactly the same as Equation (4.23), which we derived for the case of an intrinsic semiconductor. We then have that, for the semiconductor in thermal equilibrium,

$$n_0 p_0 = n_i^2 \quad (4.43)$$

Equation (4.43) states that the product of n_0 and p_0 is always a constant for a given semiconductor material at a given temperature. Although this equation seems

very simple, it is one of the fundamental principles of semiconductors in thermal equilibrium. The significance of this relation will become more apparent in the chapters that follow. It is important to keep in mind that Equation (4.43) was derived using the Boltzmann approximation. If the Boltzmann approximation is not valid, then likewise, Equation (4.43) is not valid.

An extrinsic semiconductor in thermal equilibrium does not, strictly speaking, contain an intrinsic carrier concentration, although some thermally generated carriers are present. The intrinsic electron and hole carrier concentrations are modified by the donor or acceptor impurities. However, we may think of the intrinsic concentration n_i in Equation (4.43) simply as a parameter of the semiconductor material.

*4.3.3 The Fermi–Dirac Integral

In the derivation of the Equations (4.11) and (4.19) for the thermal equilibrium electron and hole concentrations, we assumed that the Boltzmann approximation was valid. If the Boltzmann approximation does not hold, the thermal equilibrium electron concentration is written from Equation (4.3) as

$$n_0 = \frac{4\pi}{h^3} (2m_n^*)^{3/2} \int_{E_c}^{\infty} \frac{(E - E_c)^{1/2} dE}{1 + \exp\left(\frac{E - E_F}{kT}\right)} \quad (4.44)$$

If we again make a change of variable and let

$$\eta = \frac{E - E_c}{kT} \quad (4.45a)$$

and also define

$$\eta_F = \frac{E_F - E_c}{kT} \quad (4.45b)$$

then we can rewrite Equation (4.44) as

$$n_0 = 4\pi \left(\frac{2m_n^* kT}{h^2} \right)^{3/2} \int_0^{\infty} \frac{\eta^{1/2} d\eta}{1 + \exp(\eta - \eta_F)} \quad (4.46)$$

The integral is defined as

$$F_{1/2}(\eta_F) = \int_0^{\infty} \frac{\eta^{1/2} d\eta}{1 + \exp(\eta - \eta_F)} \quad (4.47)$$

This function, called the Fermi–Dirac integral, is a tabulated function of the variable η_F . Figure 4.10 is a plot of the Fermi–Dirac integral. Note that if $\eta_F > 0$, then $E_F > E_c$; thus, the Fermi energy is actually in the conduction band.

EXAMPLE 4.6

Objective: Calculate the electron concentration using the Fermi–Dirac integral.

Let $\eta_F = 2$ so that the Fermi energy is above the conduction band by approximately 52 meV at $T = 300$ K.

■ Solution

Equation (4.46) can be written as

$$n_0 = \frac{2}{\sqrt{\pi}} N_c F_{1/2}(\eta_F)$$

For silicon at $T = 300$ K, $N_c = 2.8 \times 10^{19}$ cm $^{-3}$ and, from Figure 4.10, the Fermi–Dirac integral has a value or $F_{1/2}(2) = 2.7$. Then

$$n_0 = \frac{2}{\sqrt{\pi}} (2.8 \times 10^{19})(2.7) = 8.53 \times 10^{19} \text{ cm}^{-3}$$

Comment

Note that if we had used Equation (4.11), the thermal equilibrium value of n_0 would be $n_0 = 2.08 \times 10^{20}$ cm $^{-3}$, which is incorrect since the Boltzmann approximation is not valid for this case.

EXERCISE PROBLEM

- Ex 4.6** If $n_0 = 1.5 \times 10^{20}$ cm $^{-3}$ in silicon at $T = 300$ K, determine the position of the Fermi level relative to the conduction-band energy E_c . (Ans. $E_F - E_c \equiv 0.0828 \text{ eV}$)

We may use the same general method to calculate the thermal equilibrium concentration of holes. We obtain

$$p_0 = 4\pi \left(\frac{2m_p^* kT}{h^2} \right)^{3/2} \int_0^\infty \frac{(\eta')^{1/2} d\eta'}{1 + \exp(\eta' - \eta'_F)} \quad (4.48)$$

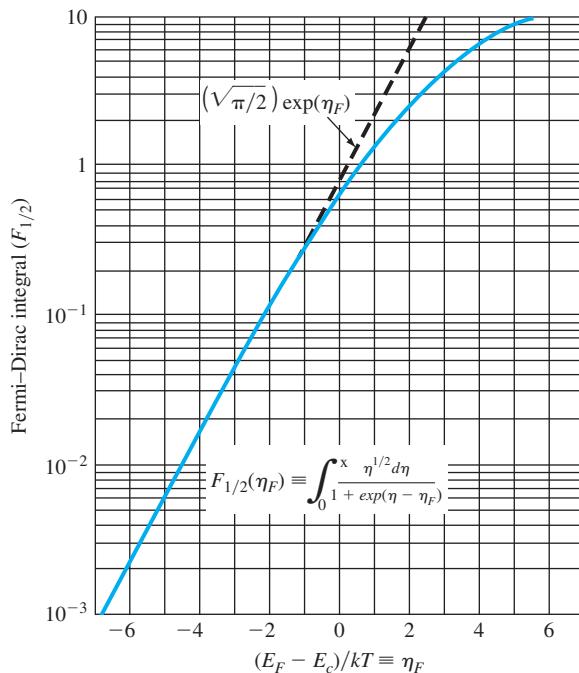


Figure 4.10 | The Fermi–Dirac integral $F_{1/2}$ as a function of the Fermi energy.
(From Sze [14].)

where

$$\eta' = \frac{E_v - E}{kT} \quad (4.49a)$$

and

$$\eta'_F = \frac{E_v - E_F}{kT} \quad (4.49b)$$

The integral in Equation (4.48) is the same Fermi–Dirac integral defined by Equation (4.47), although the variables have slightly different definitions. We may note that if $\eta'_F > 0$, then the Fermi level is in the valence band.

TEST YOUR UNDERSTANDING

- TYU 4.8** (a) Calculate the thermal-equilibrium electron concentration in silicon at $T = 300\text{ K}$ for the case when $E_F = E_c$. (b) Calculate the thermal-equilibrium hole concentration in silicon at $T = 300\text{ K}$ for the case when $E_F = E_v$.

[Ans. (a) $n_e = 2.05 \times 10^{19}\text{ cm}^{-3}$; (b) $N_h = 7.63 \times 10^{18}\text{ cm}^{-3}$]

4.3.4 Degenerate and Nondegenerate Semiconductors

In our discussion of adding dopant atoms to a semiconductor, we have implicitly assumed that the concentration of dopant atoms added is small when compared to the density of host or semiconductor atoms. The small number of impurity atoms are spread far enough apart so that there is no interaction between donor electrons, for example, in an n-type material. We have assumed that the impurities introduce discrete, noninteracting donor energy states in the n-type semiconductor and discrete, noninteracting acceptor states in the p-type semiconductor. These types of semiconductors are referred to as nondegenerate semiconductors.

If the impurity concentration increases, the distance between the impurity atoms decreases and a point will be reached when donor electrons, for example, will begin to interact with each other. When this occurs, the single discrete donor energy will split into a band of energies. As the donor concentration further increases, the band of donor states widens and may overlap the bottom of the conduction band. This overlap occurs when the donor concentration becomes comparable with the effective density of states. When the concentration of electrons in the conduction band exceeds the density of states N_c , the Fermi energy lies within the conduction band. This type of semiconductor is called a degenerate n-type semiconductor.

In a similar way, as the acceptor doping concentration increases in a p-type semiconductor, the discrete acceptor energy states will split into a band of energies and may overlap the top of the valence band. The Fermi energy will lie in the valence band when the concentration of holes exceeds the density of states N_v . This type of semiconductor is called a degenerate p-type semiconductor.

Schematic models of the energy-band diagrams for a degenerate n-type and degenerate p-type semiconductor are shown in Figure 4.11. The energy states below E_F

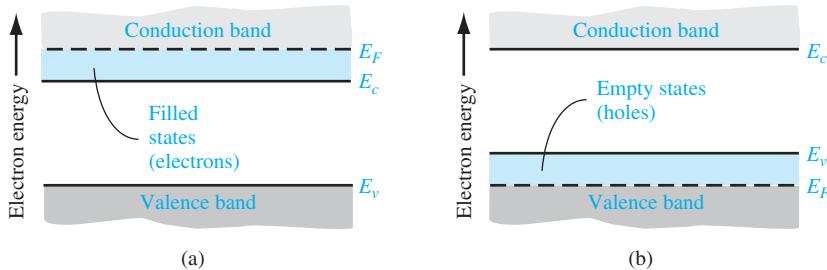


Figure 4.11 | Simplified energy-band diagrams for degenerately doped (a) n-type and (b) p-type semiconductors.

are mostly filled with electrons and the energy states above E_F are mostly empty. In the degenerate n-type semiconductor, the states between E_F and E_c are mostly filled with electrons; thus, the electron concentration in the conduction band is very large. Similarly, in the degenerate p-type semiconductor, the energy states between E_v and E_F are mostly empty; thus, the hole concentration in the valence band is very large.

4.4 | STATISTICS OF DONORS AND ACCEPTORS

In the previous chapter, we discussed the Fermi–Dirac distribution function, which gives the probability that a particular energy state will be occupied by an electron. We need to reconsider this function and apply the probability statistics to the donor and acceptor energy states.

4.4.1 Probability Function

One postulate used in the derivation of the Fermi–Dirac probability function was the Pauli exclusion principle, which states that only one particle is permitted in each quantum state. The Pauli exclusion principle also applies to the donor and acceptor states.

Suppose we have N_i electrons and g_i quantum states, where the subscript i indicates the i th energy level. There are g_i ways of choosing where to put the first particle. Each donor level has two possible spin orientations for the donor electron; thus, each donor level has two quantum states. The insertion of an electron into one quantum state, however, precludes putting an electron into the second quantum state. By adding one electron, the vacancy requirement of the atom is satisfied, and the addition of a second electron in the donor level is not possible. The distribution function of donor electrons in the donor energy states is then slightly different than the Fermi–Dirac function.

The probability function of electrons occupying the donor state is

$$n_d = \frac{N_d}{1 + \frac{1}{2} \exp\left(\frac{E_d - E_F}{kT}\right)} \quad (4.50)$$

where n_d is the density of electrons occupying the donor level and E_d is the energy of the donor level. The factor $\frac{1}{2}$ in this equation is a direct result of the spin factor just mentioned. The $\frac{1}{2}$ factor is sometimes written as $1/g$, where g is called a degeneracy factor.

Equation (4.50) can also be written in the form

$$n_d = N_d - N_d^+ \quad (4.51)$$

where N_d^+ is the concentration of ionized donors. In many applications, we will be interested more in the concentration of ionized donors than in the concentration of electrons remaining in the donor states.

If we do the same type of analysis for acceptor atoms, we obtain the expression

$$p_a = \frac{N_a^-}{1 + \frac{1}{g} \exp\left(\frac{E_F - E_a}{kT}\right)} = N_a^- - N_a \quad (4.52)$$

where N_a is the concentration of acceptor atoms, E_a is the acceptor energy level, p_a is the concentration of holes in the acceptor states, and N_a^- is the concentration of ionized acceptors. A hole in an acceptor state corresponds to an acceptor atom that is neutrally charged and still has an “empty” bonding position as we have discussed in Section 4.2.1. The parameter g is, again, a degeneracy factor. The ground state degeneracy factor g is normally taken as 4 for the acceptor level in silicon and gallium arsenide because of the detailed band structure.

4.4.2 Complete Ionization and Freeze-Out

The probability function for electrons in the donor energy state was just given by Equation (4.50). If we assume that $(E_d - E_F) \gg kT$, then

$$n_d \approx \frac{N_d}{\frac{1}{2} \exp\left(\frac{E_d - E_F}{kT}\right)} = 2N_d \exp\left[\frac{-(E_d - E_F)}{kT}\right] \quad (4.53)$$

If $(E_d - E_F) \gg kT$, then the Boltzmann approximation is also valid for the electrons in the conduction band so that, from Equation (4.11),

$$n_0 = N_c \exp\left[\frac{-(E_c - E_F)}{kT}\right]$$

We can determine the relative number of electrons in the donor state compared with the total number of electrons; therefore, we can consider the ratio of electrons in the donor state to the total number of electrons in the conduction band plus donor state. Using the expressions of Equations (4.53) and (4.11), we write

$$\frac{n_d}{n_d + n_0} = \frac{2N_d \exp\left[\frac{-(E_d - E_F)}{kT}\right]}{2N_d \exp\left[\frac{-(E_d - E_F)}{kT}\right] + N_c \exp\left[\frac{-(E_c - E_F)}{kT}\right]} \quad (4.54)$$

The Fermi energy cancels out of this expression. Dividing by the numerator term, we obtain

$$\frac{n_d}{n_d + n_0} = \frac{1}{1 + \frac{N_c}{2N_d} \exp\left[\frac{-(E_c - E_d)}{kT}\right]} \quad (4.55)$$

The factor $(E_c - E_d)$ is just the ionization energy of the donor electrons.

Objective: Determine the fraction of total electrons still in the donor states at $T = 300$ K.

Consider phosphorus doping in silicon, for $T = 300$ K, at a concentration of $N_d = 10^{16}$ cm $^{-3}$.

EXAMPLE 4.7

■ Solution

Using Equation (4.55), we find

$$\frac{n_d}{n_0 + n_d} = \frac{1}{1 + \frac{2.8 \times 10^{19}}{2(10^{16})} \exp\left(\frac{-0.045}{0.0259}\right)} = 0.0041 = 0.41\%$$

■ Comment

This example shows that there are very few electrons in the donor state compared with the conduction band. Essentially all of the electrons from the donor states are in the conduction band and, since only about 0.4 percent of the donor states contain electrons, the donor states are said to be completely ionized.

■ EXERCISE PROBLEM

Ex 4.7 Repeat Example 4.7 for (a) $T = 250$ K and (b) $T = 200$ K. (c) What can be said about the fraction as the temperature decreases?

[Ans. (a) 7.50×10^{-3} ; (b) 1.75×10^{-2} ; (c) Fraction increases as temperature decreases.]

At room temperature, then, the donor states are essentially completely ionized and, for a typical doping of 10^{16} cm $^{-3}$, almost all donor impurity atoms have donated an electron to the conduction band.

At room temperature, there is also essentially *complete ionization* of the acceptor atoms. This means that each acceptor atom has accepted an electron from the valence band so that p_a is zero. At typical acceptor doping concentrations, a hole is created in the valence band for each acceptor atom. This ionization effect and the creation of electrons and holes in the conduction band and valence band, respectively, are shown in Figure 4.12.

The opposite of complete ionization occurs at $T = 0$ K. At absolute zero degrees, all electrons are in their lowest possible energy state; that is, for an n-type semiconductor, each donor state must contain an electron, therefore $n_d = N_d$ or $N_d^+ = 0$. We must have, then, from Equation (4.50) that $\exp[(E_d - E_F)/kT] = 0$. Since $T = 0$ K, this will occur for $\exp(-\infty) = 0$, which means that $E_F > E_d$. The Fermi energy level must be above the donor energy level at absolute zero. In the case of a p-type semiconductor at absolute zero temperature, the impurity atoms will not contain any electrons, so that the Fermi energy level must be below the acceptor energy state. The distribution of electrons among the various energy states, and hence the Fermi energy, is a function of temperature.

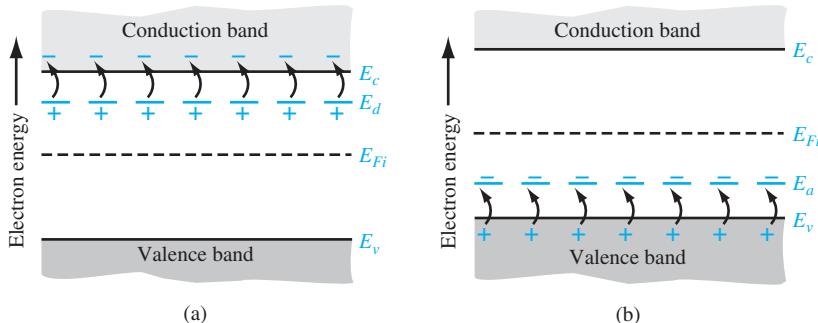


Figure 4.12 | Energy-band diagrams showing complete ionization of (a) donor states and (b) acceptor states.

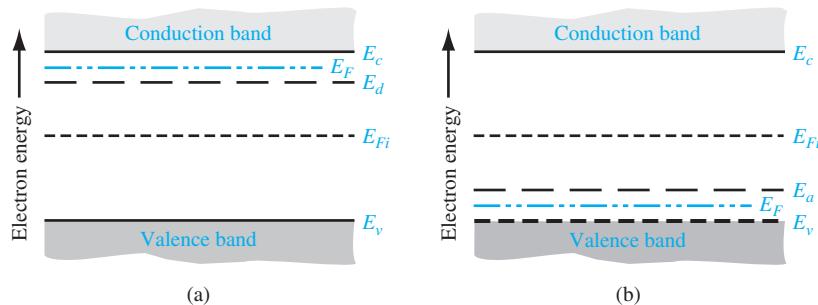


Figure 4.13 | Energy-band diagram at $T = 0$ K for (a) n-type and (b) p-type semiconductors.

A detailed analysis, not given in this text, shows that at $T = 0$ K, the Fermi energy is halfway between E_c and E_d for the n-type material and halfway between E_a and E_v for the p-type material. Figure 4.13 shows these effects. No electrons from the donor state are thermally elevated into the conduction band; this effect is called *freeze-out*. Similarly, when no electrons from the valence band are elevated into the acceptor states, the effect is also called freeze-out.

Between $T = 0$ K, freeze-out, and $T = 300$ K, complete ionization, we have partial ionization of donor or acceptor atoms.

EXAMPLE 4.8

Objective: Determine the temperature at which 90 percent of acceptor atoms are ionized. Consider p-type silicon doped with boron at a concentration of $N_a = 10^{16} \text{ cm}^{-3}$.

■ Solution

Find the ratio of holes in the acceptor state to the total number of holes in the valence band plus acceptor state. Taking into account the Boltzmann approximation and assuming the degeneracy factor is $g = 4$, we write

$$\frac{p_a}{p_0 + p_a} = \frac{1}{1 + \frac{N_v}{4N_a} \cdot \exp\left[\frac{-(E_a - E_v)}{kT}\right]}$$

For 90 percent ionization,

$$\frac{p_a}{p_0 + p_a} = 0.10 = \frac{1}{1 + \frac{(1.04 \times 10^{19}) \left(\frac{T}{300}\right)^{3/2}}{4(10^{16})} \cdot \exp\left[\frac{-0.045}{0.0259 \left(\frac{T}{300}\right)}\right]}$$

Using trial and error, we find that $T = 193$ K.

Comment

This example shows that at approximately 100°C below room temperature, we still have 90 percent of the acceptor atoms ionized; in other words, 90 percent of the acceptor atoms have “donated” a hole to the valence band.

EXERCISE PROBLEM

Ex 4.8 Determine the fraction of total holes still in the acceptor states in silicon for

$$N_a = 10^{16} \text{ cm}^{-3} \text{ at (a) } T = 250 \text{ K and (b) } T = 200 \text{ K.}$$

[Ans. (a) 3.91 × 10⁻⁷; (b) 8.76 × 10⁻⁶] [Ans. (a) 3.91 × 10⁻⁷; (b) 8.76 × 10⁻⁶]

TEST YOUR UNDERSTANDING

TYU 4.9 Determine the fraction of total holes still in the acceptor states in silicon at $T = 300$ K for a boron impurity concentration of $N_a = 10^{17} \text{ cm}^{-3}$. [Ans. 0.176]

TYU 4.10 Consider silicon with a phosphorus impurity concentration of $N_d = 10^{15} \text{ cm}^{-3}$. Determine the percent of ionized phosphorus atoms at (a) $T = 100$ K, (b) $T = 200$ K, (c) $T = 300$ K, and (d) $T = 400$ K.

[Ans. (a) 93.62%; (b) 99.82%; (c) 99.96%; (d) 99.98%]

4.5 | CHARGE NEUTRALITY

In thermal equilibrium, the semiconductor crystal is electrically neutral. The electrons are distributed among the various energy states, creating negative and positive charges, but the net charge density is zero. This charge-neutrality condition is used to determine the thermal-equilibrium electron and hole concentrations as a function of the impurity doping concentration. We will define a compensated semiconductor and then determine the electron and hole concentrations as a function of the donor and acceptor concentrations.

4.5.1 Compensated Semiconductors

A *compensated semiconductor* is one that contains both donor and acceptor impurity atoms in the same region. A compensated semiconductor can be formed, for example, by diffusing acceptor impurities into an n-type material or by diffusing donor impurities into a p-type material. An n-type compensated semiconductor occurs when $N_d > N_a$, and a p-type compensated semiconductor occurs when $N_a > N_d$. If $N_a = N_d$, we have a completely compensated semiconductor that has, as we will show, the characteristics of an intrinsic material. Compensated semiconductors are created quite naturally during device fabrication as we will see later.

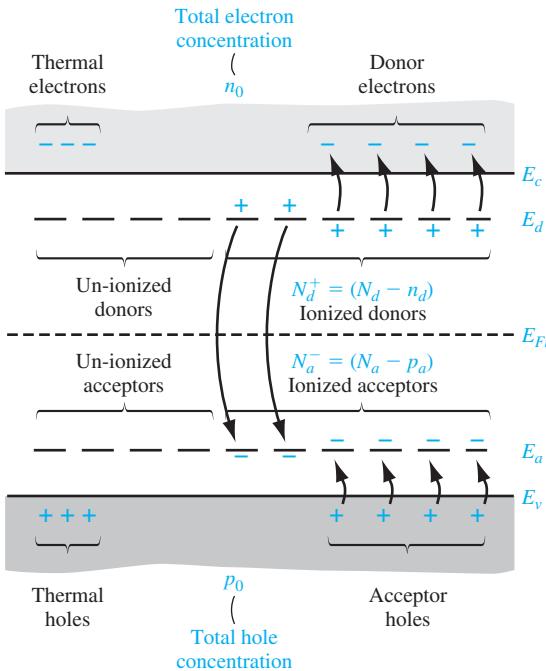


Figure 4.14 | Energy-band diagram of a compensated semiconductor showing ionized and un-ionized donors and acceptors.

4.5.2 Equilibrium Electron and Hole Concentrations

Figure 4.14 shows the energy-band diagram of a semiconductor when both donor and acceptor impurity atoms are added to the same region to form a compensated semiconductor. The figure shows how the electrons and holes can be distributed among the various states.

The charge neutrality condition is expressed by equating the density of negative charges to the density of positive charges. We then have

$$n_0 + N_a^- = p_0 + N_d^+ \quad (4.56)$$

or

$$n_0 + (N_a - p_a) = p_0 + (N_d - n_d) \quad (4.57)$$

where n_0 and p_0 are the thermal-equilibrium concentrations of electrons and holes in the conduction band and valence band, respectively. The parameter n_d is the concentration of electrons in the donor energy states, so $N_d^+ = N_d - n_d$ is the concentration of positively charged donor states. Similarly, p_a is the concentration of holes in the acceptor states, so $N_a^- = N_a - p_a$ is the concentration of negatively charged acceptor states. We have expressions for n_0 , p_0 , n_d , and p_a in terms of the Fermi energy and temperature.

Thermal-Equilibrium Electron Concentration If we assume complete ionization, n_d and p_a are both zero, and Equation (4.57) becomes

$$n_0 + N_a = p_0 + n_d \quad (4.58)$$

If we express p_0 as n_i^2/n_0 , then Equation (4.58) can be written as

$$n_0 + N_a = \frac{n_i^2}{n_0} + N_d \quad (4.59a)$$

which in turn can be written as

$$n_0^2 - (N_d - N_a)n_0 - n_i^2 = 0 \quad (4.59b)$$

The electron concentration n_0 can be determined using the quadratic formula, or

$$n_0 = \frac{(N_d - N_a)}{2} + \sqrt{\left(\frac{N_d - N_a}{2}\right)^2 + n_i^2} \quad (4.60)$$

The positive sign in the quadratic formula must be used, since, in the limit of an intrinsic semiconductor when $N_a = N_d = 0$, the electron concentration must be a positive quantity, or $n_0 = n_i$.

Equation (4.60) is used to calculate the electron concentration in an n-type semiconductor, or when $N_d > N_a$. Although Equation (4.60) was derived for a compensated semiconductor, the equation is also valid for $N_a = 0$.

Objective: Determine the thermal-equilibrium electron and hole concentrations in silicon at $T = 300$ K for given doping concentrations. (a) Let $N_d = 10^{16}$ cm $^{-3}$ and $N_a = 0$. (b) Let $N_d = 5 \times 10^{15}$ cm $^{-3}$ and $N_a = 2 \times 10^{15}$ cm $^{-3}$.

Recall that $n_i = 1.5 \times 10^{10}$ cm $^{-3}$ in silicon at $T = 300$ K.

EXAMPLE 4.9

■ Solution

(a) From Equation (4.60), the majority carrier electron concentration is

$$n_0 = \frac{10^{16}}{2} + \sqrt{\left(\frac{10^{16}}{2}\right)^2 + (1.5 \times 10^{10})^2} \cong 10^{16} \text{ cm}^{-3}$$

The minority carrier hole concentration is found to be

$$p_0 = \frac{n_i^2}{n_0} = \frac{(1.5 \times 10^{10})^2}{10^{16}} = 2.25 \times 10^4 \text{ cm}^{-3}$$

(b) Again, from Equation (4.60), the majority carrier electron concentration is

$$n_0 = \frac{5 \times 10^{15} - 2 \times 10^{15}}{2} + \sqrt{\left(\frac{5 \times 10^{15} - 2 \times 10^{15}}{2}\right)^2 + (1.5 \times 10^{10})^2} \cong 3 \times 10^{15} \text{ cm}^{-3}$$

The minority carrier hole concentration is

$$p_0 = \frac{n_i^2}{n_0} = \frac{(1.5 \times 10^{10})^2}{3 \times 10^{15}} = 7.5 \times 10^4 \text{ cm}^{-3}$$

Comment

In both parts of this example, $(N_d - N_a) \gg n_i$, so the thermal-equilibrium majority carrier electron concentration is essentially equal to the difference between the donor and acceptor concentrations. Also, in both cases, the majority carrier electron concentration is orders of magnitude larger than the minority carrier hole concentration.

EXERCISE PROBLEM

- Ex 4.9** Find the thermal-equilibrium electron and hole concentrations in silicon with doping concentrations of $N_d = 7 \times 10^{15} \text{ cm}^{-3}$ and $N_a = 3 \times 10^{15} \text{ cm}^{-3}$ for (a) $T = 250 \text{ K}$ and (b) $T = 400 \text{ K}$.

$$\begin{aligned} & [\text{Ans. (a)} n_0 = 4 \times 10^{15} \text{ cm}^{-3}, p_0 = 1.225 \text{ cm}^{-3}; \text{(b)} n_0 = 4 \times 10^{15} \text{ cm}^{-3}, p_0 = \\ & \quad 1.416 \times 10^9 \text{ cm}^{-3}] \end{aligned}$$

We have argued in our discussion and we may note from the results of Example 4.9 that the concentration of electrons in the conduction band increases above the intrinsic carrier concentration as we add donor impurity atoms. At the same time, the minority carrier hole concentration decreases below the intrinsic carrier concentration as we add donor atoms. We must keep in mind that as we add donor impurity atoms and the corresponding donor electrons, there is a redistribution of electrons among available energy states. Figure 4.15 shows a schematic of this physical redistribution. A few of the donor electrons will fall into the empty states in the valence band and, in doing so, will annihilate some of the intrinsic holes. The minority carrier hole concentration will therefore decrease as we have seen in Example 4.9. At the

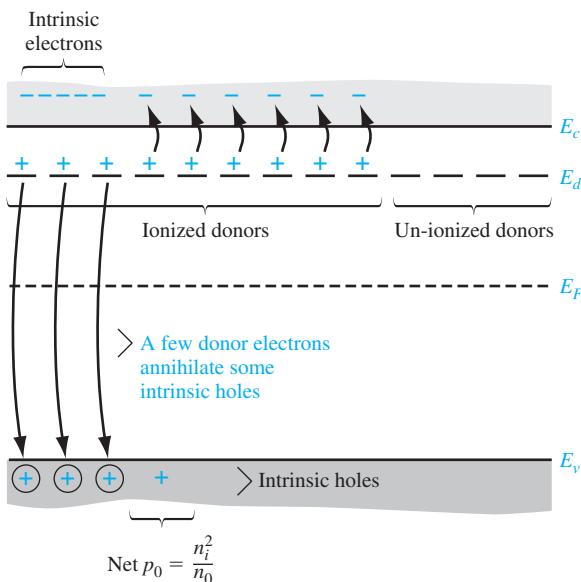


Figure 4.15 | Energy-band diagram showing the redistribution of electrons when donors are added.

same time, because of this redistribution, the net electron concentration in the conduction band is *not* simply equal to the donor concentration plus the intrinsic electron concentration.

Objective: Calculate the thermal-equilibrium electron and hole concentrations in germanium for a given doping concentration.

EXAMPLE 4.10

Consider a germanium sample at $T = 300$ K in which $N_d = 2 \times 10^{14} \text{ cm}^{-3}$ and $N_a = 0$. Assume that $n_i = 2.4 \times 10^{13} \text{ cm}^{-3}$.

■ Solution

Again, from Equation (4.60), the majority carrier electron concentration is

$$n_0 = \frac{2 \times 10^{14}}{2} + \sqrt{\left(\frac{2 \times 10^{14}}{2}\right)^2 + (2.4 \times 10^{13})^2} \cong 2.028 \times 10^{14} \text{ cm}^{-3}$$

The minority carrier hole concentration is

$$p_0 = \frac{n_i^2}{n_0} = \frac{(2.4 \times 10^{13})^2}{2.028 \times 10^{14}} = 2.84 \times 10^{12} \text{ cm}^{-3}$$

■ Comment

If the donor impurity concentration is not too different in magnitude from the intrinsic carrier concentration, then the thermal-equilibrium majority carrier electron concentration is influenced by the intrinsic concentration.

■ EXERCISE PROBLEM

Ex 4.10 Repeat Example 4.10 for (a) $T = 250$ K and (b) $T = 350$ K. (c) What can be said about a very low-doped material as the temperature increases?

[Ans. (a) $n_0 \equiv 2 \times 10^{14} \text{ cm}^{-3}$, $p_0 = 9.47 \times 10^9 \text{ cm}^{-3}$; (b) $n_0 = 3.059 \times 10^{14} \text{ cm}^{-3}$; (c) Material approaches an intrinsic semiconductor]

[Ans. (a) $n_0 \equiv 2 \times 10^{14} \text{ cm}^{-3}$, $p_0 = 9.47 \times 10^9 \text{ cm}^{-3}$; (b) $n_0 = 3.059 \times 10^{14} \text{ cm}^{-3}$]

We have seen that the intrinsic carrier concentration n_i is a very strong function of temperature. As the temperature increases, additional electron–hole pairs are thermally generated so that the n_i^2 term in Equation (4.60) may begin to dominate. The semiconductor will eventually lose its extrinsic characteristics. Figure 4.16 shows the electron concentration versus temperature in silicon doped with 5×10^{14} donors per cm^3 . As the temperature increases, we can see where the intrinsic concentration begins to dominate. Also shown is the partial ionization, or the onset of freeze-out, at the low temperature.

Thermal-Equilibrium Hole Concentration If we reconsider Equation (4.58) and express n_0 as n_i^2/p_0 , then we have

$$\frac{n_i^2}{p_0} + N_a = p_0 + N_d \quad (4.61\text{a})$$

which we can write as

$$p_0^2 - (N_a - N_d)p_0 - n_i^2 = 0 \quad (4.61\text{b})$$

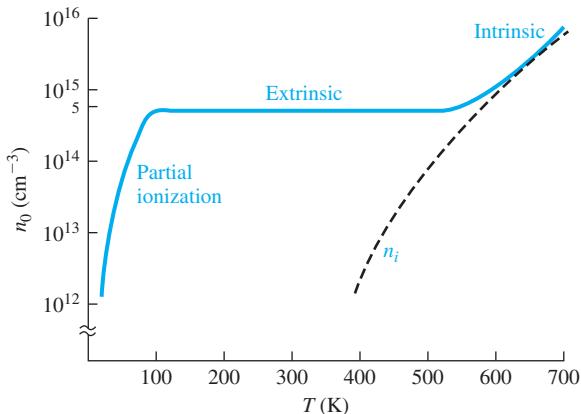


Figure 4.16 | Electron concentration versus temperature showing the three regions: partial ionization, extrinsic, and intrinsic.

Using the quadratic formula, the hole concentration is given by

$$p_0 = \frac{N_a - N_d}{2} + \sqrt{\left(\frac{N_a - N_d}{2}\right)^2 + n_i^2} \quad (4.62)$$

where the positive sign, again, must be used. Equation (4.62) is used to calculate the thermal-equilibrium majority carrier hole concentration in a p-type semiconductor, or when \$N_a > N_d\$. This equation also applies for \$N_d = 0\$.

EXAMPLE 4.11

Objective: Calculate the thermal-equilibrium electron and hole concentrations in a compensated p-type semiconductor.

Consider a silicon semiconductor at \$T = 300\$ K in which \$N_a = 10^{16} \text{ cm}^{-3}\$ and \$N_d = 3 \times 10^{15} \text{ cm}^{-3}\$. Assume \$n_i = 1.5 \times 10^{10} \text{ cm}^{-3}\$.

■ Solution

Since \$N_a > N_d\$, the compensated semiconductor is p-type and the thermal-equilibrium majority carrier hole concentration is given by Equation (4.62) as

$$p_0 = \frac{10^{16} - 3 \times 10^{15}}{2} + \sqrt{\left(\frac{10^{16} - 3 \times 10^{15}}{2}\right)^2 + (1.5 \times 10^{10})^2}$$

so that

$$p_0 \approx 7 \times 10^{15} \text{ cm}^{-3}$$

The minority carrier electron concentration is

$$n_0 = \frac{n_i^2}{p_0} = \frac{(1.5 \times 10^{10})^2}{7 \times 10^{15}} = 3.21 \times 10^4 \text{ cm}^{-3}$$

■ Comment

If we assume complete ionization and if \$(N_a - N_d) \gg n_i\$, then the majority carrier hole concentration is, to a very good approximation, just the difference between the acceptor and donor concentrations.

EXERCISE PROBLEM

- Ex 4.11** Consider silicon at $T = 300$ K. Calculate the thermal-equilibrium electron and hole concentrations for impurity concentrations of (a) $N_a = 4 \times 10^{16} \text{ cm}^{-3}$, $N_d = 8 \times 10^{15} \text{ cm}^{-3}$ and (b) $N_a = N_d = 3 \times 10^{15} \text{ cm}^{-3}$.

$$[\text{Ans. (a)} d = 3.2 \times 10^{16} \text{ cm}^{-3}, n_0 = 1.5 \times 10^{16} \text{ cm}^{-3}; \text{Ans. (b)} d = 7.03 \times 10^5 \text{ cm}^{-3}, n_0 = 1.5 \times 10^{16} \text{ cm}^{-3}]$$

We may note that, for a compensated p-type semiconductor, the minority carrier electron concentration is determined from

$$n_0 = \frac{n_i^2}{p_0} = \frac{n_i^2}{(N_a - N_d)}$$

Equations (4.60) and (4.62) are used to calculate the majority carrier electron concentration in an n-type semiconductor and majority carrier hole concentration in a p-type semiconductor, respectively. The minority carrier hole concentration in an n-type semiconductor could, theoretically, be calculated from Equation (4.62). However, we would be subtracting two numbers on the order of 10^{16} cm^{-3} , for example, to obtain a number on the order of 10^4 cm^{-3} , which from a practical point of view is not possible. The minority carrier concentrations are calculated from $n_0 p_0 = n_i^2$ once the majority carrier concentration has been determined.

TEST YOUR UNDERSTANDING

- TYU 4.11** Consider a compensated GaAs semiconductor at $T = 300$ K doped at $N_d = 5 \times 10^{15} \text{ cm}^{-3}$ and $N_a = 2 \times 10^{16} \text{ cm}^{-3}$. Calculate the thermal equilibrium electron and hole concentrations. ($d = 2.16 \times 10^{16} \text{ cm}^{-3}, n_0 = 1.5 \times 10^{16} \text{ cm}^{-3}$)
- TYU 4.12** Silicon is doped at $N_d = 10^{15} \text{ cm}^{-3}$ and $N_a = 0$. (a) Plot the concentration of electrons versus temperature over the range $300 \leq T \leq 600$ K. (b) Calculate the temperature at which the electron concentration is equal to $1.1 \times 10^{15} \text{ cm}^{-3}$. ($\text{Ans. } T \approx 552 \text{ K}$)
- TYU 4.13** A silicon device with n-type material is to be operated at $T = 550$ K. At this temperature, the intrinsic carrier concentration must contribute no more than 5 percent of the total electron concentration. Determine the minimum donor concentration required to meet this specification. ($\text{Ans. } N_d = 1.40 \times 10^{15} \text{ cm}^{-3}$)

4.6 | POSITION OF FERMI ENERGY LEVEL

We have discussed qualitatively in Section 4.3.1 how the electron and hole concentrations change as the Fermi energy level moves through the bandgap energy. Then, in Section 4.5, we calculated the electron and hole concentrations as a function of donor and acceptor impurity concentrations. We can now determine the position of the Fermi energy level as a function of the doping concentrations and as a function of temperature. The relevance of the Fermi energy level will be further discussed after the mathematical derivations.

4.6.1 Mathematical Derivation

The position of the Fermi energy level within the bandgap can be determined by using the equations already developed for the thermal-equilibrium electron and hole concentrations. If we assume the Boltzmann approximation to be valid, then from Equation (4.11) we have $n_0 = N_c \exp [-(E_c - E_F)/kT]$. We can solve for $E_c - E_F$ from this equation and obtain

$$E_c - E_F = kT \ln \left(\frac{N_c}{n_0} \right) \quad (4.63)$$

where n_0 is given by Equation (4.60). If we consider an n-type semiconductor in which $N_d \gg n_i$, then $n_0 \approx N_d$, so that

$$E_c - E_F = kT \ln \left(\frac{N_c}{N_d} \right) \quad (4.64)$$

The distance between the bottom of the conduction band and the Fermi energy is a logarithmic function of the donor concentration. As the donor concentration increases, the Fermi level moves closer to the conduction band. Conversely, if the Fermi level moves closer to the conduction band, then the electron concentration in the conduction band is increasing. We may note that if we have a compensated semiconductor, then the N_d term in Equation (4.64) is simply replaced by $N_d - N_a$, or- the net effective donor concentration.

DESIGN EXAMPLE 4.12

Objective: Determine the required donor impurity concentration to obtain a specified Fermi energy.

Silicon at $T = 300$ K contains an acceptor impurity concentration of $N_a = 10^{16}$ cm⁻³. Determine the concentration of donor impurity atoms that must be added so that the silicon is n type and the Fermi energy is 0.20 eV below the conduction-band edge.

■ Solution

From Equation (4.64), we have

$$E_c - E_F = kT \ln \left(\frac{N_c}{N_d - N_a} \right)$$

which can be rewritten as

$$N_d - N_a = N_c \exp \left[\frac{-(E_c - E_F)}{kT} \right]$$

Then

$$N_d - N_a = 2.8 \times 10^{19} \exp \left[\frac{-0.20}{0.0259} \right] = 1.24 \times 10^{16} \text{ cm}^{-3}$$

or

$$N_d = 1.24 \times 10^{16} + N_a = 2.24 \times 10^{16} \text{ cm}^{-3}$$

■ Comment

A compensated semiconductor can be fabricated to provide a specific Fermi energy level.

EXERCISE PROBLEM

Ex 4.12 Consider silicon at $T = 300$ K with doping concentrations of $N_d = 8 \times 10^{15} \text{ cm}^{-3}$ and $N_a = 5 \times 10^{15} \text{ cm}^{-3}$. Determine the position of the Fermi energy level with respect to E_c . (Ans. $E_F - E_c = 0.23\text{eV}$)

We may develop a slightly different expression for the position of the Fermi level. We had from Equation (4.39) that $n_0 = n_i \exp [-(E_F - E_{Fi})/kT]$. We can solve for $E_F - E_{Fi}$ as

$$E_F - E_{Fi} = kT \ln \left(\frac{n_0}{n_i} \right) \quad (4.65)$$

Equation (4.65) can be used specifically for an n-type semiconductor, where n_0 is given by Equation (4.60), to find the difference between the Fermi level and the intrinsic Fermi level as a function of the donor concentration. We may note that, if the net effective donor concentration is zero, that is, $N_d - N_a = 0$, then $n_0 = n_i$ and $E_F = E_{Fi}$. A completely compensated semiconductor has the characteristics of an intrinsic material in terms of carrier concentration and Fermi-level position.

We can derive the same types of equations for a p-type semiconductor. From Equation (4.19), we have $p_0 = N_v \exp [-(E_F - E_v)/kT]$, so that

$$E_F - E_v = kT \ln \left(\frac{N_v}{p_0} \right) \quad (4.66)$$

If we assume that $N_a \gg n_i$, then Equation (4.66) can be written as

$$E_F - E_v = kT \ln \left(\frac{N_v}{N_a} \right) \quad (4.67)$$

The distance between the Fermi level and the top of the valence-band energy for a p-type semiconductor is a logarithmic function of the acceptor concentration: as the acceptor concentration increases, the Fermi level moves closer to the valence band. Equation (4.67) still assumes that the Boltzmann approximation is valid. Again, if we have a compensated p-type semiconductor, then the N_a term in Equation (4.67) is replaced by $N_a - N_d$, or the net effective acceptor concentration.

We can also derive an expression for the relationship between the Fermi level and the intrinsic Fermi level in terms of the hole concentration. We have from Equation (4.40) that $p_0 = n_i \exp [-(E_F - E_{Fi})/kT]$, which yields

$$E_{Fi} - E_F = kT \ln \left(\frac{p_0}{n_i} \right) \quad (4.68)$$

Equation (4.68) can be used to find the difference between the intrinsic Fermi level and the Fermi energy in terms of the acceptor concentration. The hole concentration p_0 in Equation (4.68) is given by Equation (4.62).

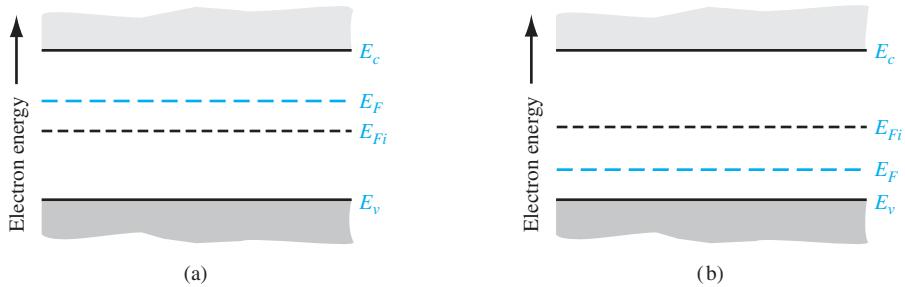


Figure 4.17 | Position of Fermi level for an (a) n-type ($N_d > N_a$) and (b) p-type ($N_d > N_a$) semiconductor.

We may again note from Equation (4.65) that, for an n-type semiconductor, $n_0 > n_i$ and $E_F > E_{Fi}$. The Fermi level for an n-type semiconductor is above E_{Fi} . For a p-type semiconductor, $p_0 > n_i$, and from Equation (4.68) we see that $E_{Fi} > E_F$. The Fermi level for a p-type semiconductor is below E_{Fi} . These results are shown in Figure 4.17.

4.6.2 Variation of E_F with Doping Concentration and Temperature

We may plot the position of the Fermi energy level as a function of the doping concentration. Figure 4.18 shows the Fermi energy level as a function of donor concentration (n type) and as a function of acceptor concentration (p type) for silicon at $T = 300$ K. As the doping levels increase, the Fermi energy level moves closer to the conduction band for the n-type material and closer to the valence band for the p-type material. Keep in mind that the equations for the Fermi energy level that we have derived assume that the Boltzmann approximation is valid.

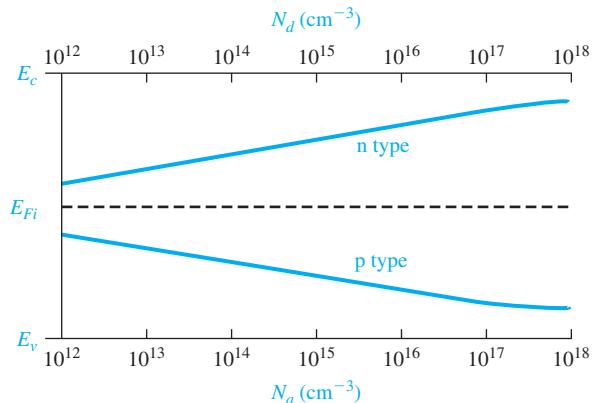


Figure 4.18 | Position of Fermi level as a function of donor concentration (n type) and acceptor concentration (p type).

Objective: Determine the Fermi energy level and the maximum doping concentration at which the Boltzmann approximation is still valid.

Consider p-type silicon, at $T = 300$ K, doped with boron. We may assume that the limit of the Boltzmann approximation occurs when $E_F - E_a = 3kT$. (See Section 4.1.2.)

■ Solution

From Table 4.3, we find the ionization energy is $E_a - E_v = 0.045$ eV for boron in silicon. If we assume that $E_{Fi} \approx E_{\text{midgap}}$, then from Equation (4.68), the position of the Fermi level at the maximum doping is given by

$$E_{Fi} - E_F = \frac{E_g}{2} - (E_a - E_v) - (E_F - E_a) = kT \ln \left(\frac{N_a}{n_i} \right)$$

or

$$0.56 - 0.045 - 3(0.0259) = 0.437 = (0.0259) \ln \left(\frac{N_a}{n_i} \right)$$

We can then solve for the doping as

$$N_a = n_i \exp \left(\frac{0.437}{0.0259} \right) = 3.2 \times 10^{17} \text{ cm}^{-3}$$

■ Comment

If the acceptor (or donor) concentration in silicon is greater than approximately $3 \times 10^{17} \text{ cm}^{-3}$, then the Boltzmann approximation of the distribution function becomes less valid and the equations for the Fermi-level position are no longer quite as accurate.

■ EXERCISE PROBLEM

Ex 4.13 Consider n-type silicon at $T = 300$ K doped with arsenic. Determine the maximum doping at which the Boltzmann approximation is still valid. Assume the limit is such that $E_d - E_F = 3kT$. (Ans. $n_i = 2.02 \times 10^{17} \text{ cm}^{-3}$)

The intrinsic carrier concentration n_i , in Equations (4.65) and (4.68), is a strong function of temperature, so that E_F is a function of temperature also. Figure 4.19 shows the variation of the Fermi energy level in silicon with temperature for several donor and acceptor concentrations. As the temperature increases, n_i increases, and E_F moves closer to the intrinsic Fermi level. At high temperature, the semiconductor material begins to lose its extrinsic characteristics and begins to behave more like an intrinsic semiconductor. At the very low temperature, freeze-out occurs; the Boltzmann approximation is no longer valid and the equations we derived for the Fermi-level position no longer apply. At the low temperature where freeze-out occurs, the Fermi level goes above E_d for the n-type material and below E_a for the p-type material. At absolute zero degrees, all energy states below E_F are full and all energy states above E_F are empty.

4.6.3 Relevance of the Fermi Energy

We have been calculating the position of the Fermi energy level as a function of doping concentrations and temperature. This analysis may seem somewhat arbitrary and fictitious. However, these relations do become significant later in our discussion of pn junctions and the other semiconductor devices we consider. An important

EXAMPLE 4.13

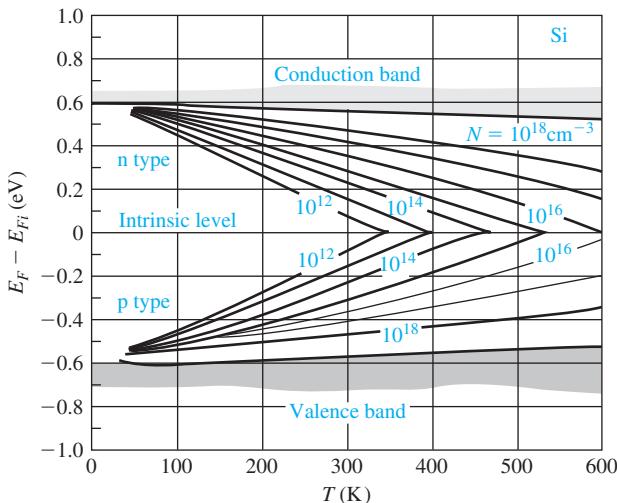


Figure 4.19 | Position of Fermi level as a function of temperature for various doping concentrations.
(From Sze [14].)

point is that, in thermal equilibrium, the Fermi energy level is a constant throughout a system. We will not prove this statement, but we can intuitively see its validity by considering the following example.

Suppose we have a particular material, A, whose electrons are distributed in the energy states of an allowed band as shown in Figure 4.20a. Most of the energy states below E_{FA} contain electrons and most of the energy states above E_{FA} are empty of electrons. Consider another material, B, whose electrons are distributed in the energy states of an allowed band as shown in Figure 4.20b. The energy states below E_{FB} are mostly full and the energy states above E_{FB} are mostly empty. If these two materials are brought into intimate contact, the electrons in the entire system will tend to seek the lowest possible energy. Electrons from material A will flow into the lower energy states of material B, as indicated in Figure 4.20c, until thermal equilibrium is reached. Thermal equilibrium occurs when the distribution of electrons, as a function of energy, is the same in the two materials. This equilibrium state occurs when the Fermi energy is the same in the two materials as shown in Figure 4.20d. The Fermi energy, important in the physics of the semiconductor, also provides a good pictorial representation of the characteristics of the semiconductor materials and devices.

TEST YOUR UNDERSTANDING

TYU 4.14 Determine the position of the Fermi level with respect to the valence-band energy in p-type GaAs at $T = 300$ K. The doping concentrations are $N_a = 5 \times 10^{16} \text{ cm}^{-3}$ and $N_d = 4 \times 10^{15} \text{ cm}^{-3}$.
(Ans. $E_F - E_v = 0.130 \text{ eV}$)

TYU 4.15 Calculate the position of the Fermi energy level in n-type silicon at $T = 300$ K with respect to the intrinsic Fermi energy level. The doping concentrations are $N_d = 2 \times 10^{17} \text{ cm}^{-3}$ and $N_a = 3 \times 10^{16} \text{ cm}^{-3}$.
(Ans. $E_F - E_Fi = 0.421 \text{ eV}$)

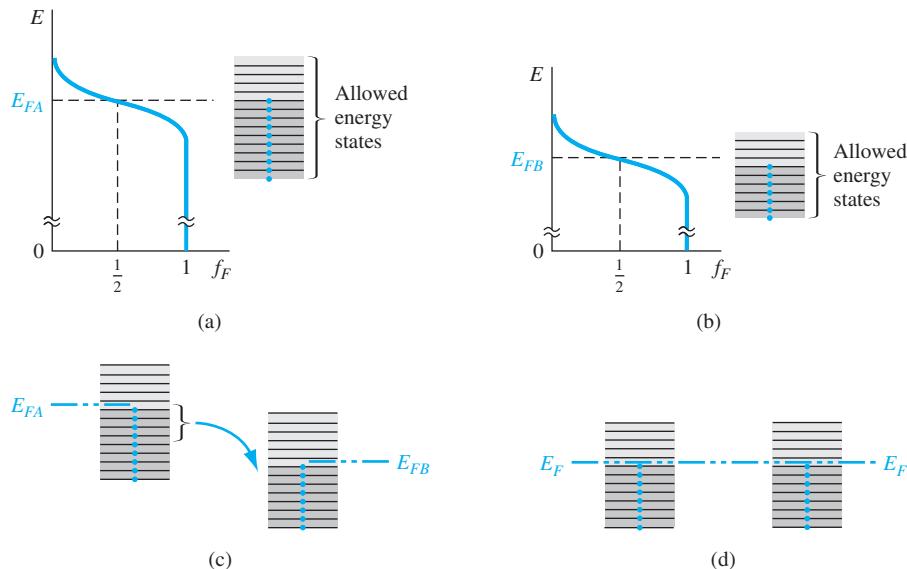


Figure 4.20 | The Fermi energy of (a) material A in thermal equilibrium, (b) material B in thermal equilibrium, (c) materials A and B at the instant they are placed in contact, and (d) materials A and B in contact at thermal equilibrium.

4.7 | SUMMARY

- The concentration of electrons in the conduction band is the integral over the conduction-band energy of the product of the density of states function in the conduction band and the Fermi-Dirac probability function.
- The concentration of holes in the valence band is the integral over the valence-band energy of the product of the density of states function in the valence band and the probability of a state being empty, which is $[1 - f_F(E)]$.
- Using the Maxwell-Boltzmann approximation, the thermal-equilibrium concentration of electrons in the conduction band is given by

$$n_0 = N_c \exp\left[\frac{-(E_c - E_F)}{kT}\right]$$

where N_c is the effective density of states in the conduction band.

- Using the Maxwell-Boltzmann approximation, the thermal-equilibrium concentration of holes in the valence band is given by

$$p_0 = N_v \exp\left[\frac{-(E_F - E_v)}{kT}\right]$$

where N_v is the effective density of states in the valence band.

- The intrinsic carrier concentration is found from

$$n_i^2 = N_c N_v \exp\left[\frac{-E_g}{kT}\right]$$

- The concept of doping the semiconductor with donor (group V elements) impurities and acceptor (group III elements) impurities to form n-type and p-type extrinsic semiconductors was discussed.
- The fundamental relationship of $n_0 p_0 = n_i^2$ was derived.
- Using the concepts of complete ionization and charge neutrality, equations for the electron and hole concentrations as a function of impurity doping concentrations were derived.
- The position of the Fermi energy level as a function of impurity doping concentrations was derived.
- The relevance of the Fermi energy was discussed. The Fermi energy is a constant throughout a semiconductor that is in thermal equilibrium.

GLOSSARY OF IMPORTANT TERMS

acceptor atoms Impurity atoms added to a semiconductor to create a p-type material.

charge carrier The electron and/or hole that moves inside the semiconductor and gives rise to electrical currents.

compensated semiconductor A semiconductor that contains both donors and acceptors in the same semiconductor region.

complete ionization The condition when all donor atoms are positively charged by giving up their donor electrons and all acceptor atoms are negatively charged by accepting electrons.

degenerate semiconductor A semiconductor whose electron concentration or hole concentration is greater than the effective density of states, so that the Fermi level is in the conduction band (n type) or in the valence band (p type).

donor atoms Impurity atoms added to a semiconductor to create an n-type material.

effective density of states The parameter N_c , which results from integrating the density of quantum states $g_c(E)$ times the Fermi function $f_F(E)$ over the conduction-band energy, and the parameter N_v , which results from integrating the density of quantum states $g_v(E)$ times $[1 - f_F(E)]$ over the valence-band energy.

extrinsic semiconductor A semiconductor in which controlled amounts of donors and/or acceptors have been added so that the electron and hole concentrations change from the intrinsic carrier concentration and a preponderance of either electrons (n type) or holes (p type) is created.

freeze-out The condition that occurs in a semiconductor when the temperature is lowered and the donors and acceptors become neutrally charged. The electron and hole concentrations become very small.

CHECKPOINT

After studying this chapter, the reader should have the ability to:

- Derive the equations for the thermal equilibrium concentrations of electrons and holes in terms of the Fermi energy.
- Derive the equation for the intrinsic carrier concentration.
- Discuss what is meant by the effective density of states for electrons and holes.
- Describe the effect of adding donor and acceptor impurity atoms to a semiconductor.

- Understand the concept of complete ionization.
- Derive the fundamental relationship $n_0 p_0 = n_i^2$.
- Describe the meaning of degenerate and nondegenerate semiconductor materials.
- Discuss the concept of charge neutrality.
- Derive the equations for n_0 and p_0 in terms of impurity doping concentrations.
- Derive the equations for the Fermi energy in terms of the impurity doping concentrations.
- Discuss the variation of the Fermi energy with doping concentration and temperature.

REVIEW QUESTIONS

1. How does the electron concentration in the conduction band change with energy E for $E > E_c$?
2. In deriving the equation for n_0 in terms of the Fermi function, the upper limit of the integral should be the energy at the top of the conduction band. Justify using infinity instead.
3. Assuming the Boltzmann approximation applies, write the equations for n_0 and p_0 in terms of the Fermi energy.
4. What is the source of electrons and holes in an intrinsic semiconductor?
5. Under what condition would the intrinsic Fermi level be at the midgap energy?
6. What is a donor impurity? What is an acceptor impurity?
7. What is meant by complete ionization? What is meant by freeze-out?
8. What is the product of n_0 and p_0 equal to?
9. Write the equation for charge neutrality for the condition of complete ionization.
10. Sketch a graph of n_0 versus temperature for an n-type material.
11. Sketch graphs of the Fermi energy versus donor impurity concentration and versus temperature.
12. What is the relevance of the Fermi energy?

PROBLEMS

Section 4.1 Charge Carriers in Semiconductors

- 4.1 Calculate the intrinsic carrier concentration, n_i , at $T = 200, 400$, and 600 K for (a) silicon, (b) germanium, and (c) gallium arsenide.
- 4.2 Plot the intrinsic carrier concentration, n_i , for a temperature range of $200 \leq T \leq 600$ K for (a) silicon, (b) germanium, and (c) gallium arsenide. (Use a log scale for n_i .)
- 4.3 (a) The maximum intrinsic carrier concentration in a silicon device must be limited to $5 \times 10^{11} \text{ cm}^{-3}$. Assume $E_g = 1.12 \text{ eV}$. Determine the maximum temperature allowed for the device. (b) Repeat part (a) if the maximum intrinsic carrier concentration is limited to $5 \times 10^{12} \text{ cm}^{-3}$.
- 4.4 In a particular semiconductor material, the effective density of states functions are given by $N_c = N_{c0} \cdot (T/300)^{3/2}$ and $N_v = N_{v0} \cdot (T/300)^{3/2}$ where N_{c0} and N_{v0} are constants independent of temperature. Experimentally determined intrinsic carrier concentrations are found to be $n_i = 1.40 \times 10^2 \text{ cm}^{-3}$ at $T = 200$ K and $n_i = 7.70 \times 10^{10} \text{ cm}^{-2}$ at $T = 400$ K. Determine the product $N_{c0} \cdot N_{v0}$ and the bandgap energy E_g . (Assume E_g is constant over this temperature range.)

- 4.5** Two semiconductor materials have exactly the same properties except material A has a bandgap energy of 0.90 eV and material B has a bandgap energy of 1.10 eV. Determine the ratio of n_i of material B to that of material A for (a) $T = 200$ K, (b) $T = 300$ K, and (c) $T = 400$ K.
- 4.6** (a) The magnitude of the product $g_c(E)f_F(E)$ in the conduction band is a function of energy as shown in Figure 4.1. Assume the Boltzmann approximation is valid. Determine the energy with respect to E_c at which the maximum occurs. (b) Repeat part (a) for the magnitude of the product $g_v(E)[1 - f_F(E)]$ in the valence band.
- 4.7** Assume the Boltzmann approximation in a semiconductor is valid. Determine the ratio of $n(E) = g_c(E)f_F(E)$ at $E = E_c + 4kT$ to that at $E = E_c + kT/2$.
- 4.8** Assume that $E_c - E_F = 0.20$ eV in silicon. Plot $n(E) = g_c(E)f_F(E)$ over the range $E_c \leq E \leq E_c + 0.10$ eV for (a) $T = 200$ K and (b) $T = 400$ K.
- 4.9** (a) Consider silicon at $T = 300$ K. Plot the thermal equilibrium electron concentration n_0 (on a log scale) over the energy range $0.2 \leq E_c - E_F \leq 0.4$ eV. (b) Repeat part (a) for the hole concentration over the range $0.2 \leq E_F - E_v \leq 0.4$ eV.
- 4.10** Given the effective masses of electrons and holes in silicon, germanium, and gallium arsenide, calculate the position of the intrinsic Fermi energy level with respect to the center of the bandgap for each semiconductor at $T = 300$ K.
- 4.11** Calculate E_{Fi} with respect to the center of the bandgap in silicon for $T = 200$, 400, and 600 K.
- 4.12** (a) The carrier effective masses in a semiconductor are $m_n^* = 1.21 m_0$ and $m_p^* = 0.70 m_0$. Determine the position of the intrinsic Fermi level with respect to the center of the bandgap at $T = 300$ K. (b) Repeat part (a) if $m_n^* = 0.080 m_0$ and $m_p^* = 0.75 m_0$.
- 4.13** If the density of states function in the conduction band of a particular semiconductor is a constant equal to K, derive the expression for the thermal-equilibrium concentration of electrons in the conduction band, assuming Fermi-Dirac statistics and assuming the Boltzmann approximation is valid.
- 4.14** Repeat Problem 4.13 if the density of states function is given by $g_c(E) = C_1(E - E_c)$ for $E \geq E_c$ where C_1 is a constant.

Section 4.2 Dopant Atoms and Energy Levels

- 4.15** Calculate the ionization energy and radius of the donor electron in germanium using the Bohr theory.
- 4.16** Repeat Problem 4.15 for gallium arsenide.

Section 4.3 The Extrinsic Semiconductor

- 4.17** Silicon at $T = 300$ K is doped with arsenic atoms such that the concentration of electrons is $n_0 = 7 \times 10^{15} \text{ cm}^{-3}$. (a) Find $E_c - E_F$. (b) Determine $E_F - E_v$. (c) Calculate p_0 . (d) Which carrier is the minority carrier? (e) Find $E_F - E_{Fi}$.
- 4.18** The value of p_0 in silicon at $T = 300$ K is $2 \times 10^{16} \text{ cm}^{-3}$. (a) Determine $E_F - E_v$. (b) Calculate the value of $E_c - E_F$. (c) What is the value of n_0 ? (d) Determine $E_{Fi} - E_F$.
- 4.19** The electron concentration in silicon at $T = 300$ K is $n_0 = 2 \times 10^5 \text{ cm}^{-3}$. (a) Determine the position of the Fermi level with respect to the valence band energy level. (b) Determine p_0 . (c) Is this n- or p-type material?

- 4.20** (a) If $E_c - E_F = 0.28$ eV in gallium arsenide at $T = 375$ K, calculate the values of n_0 and p_0 . (b) Assuming the value of n_0 in part (a) remains constant, determine $E_c - E_F$ and p_0 at $T = 300$ K.
- 4.21** Repeat Problem 4.20 for silicon.
- 4.22** The Fermi energy level in silicon at $T = 300$ K is as close to the top of the valence band as to the midgap energy. (a) Is the material n type or p type? (b) Calculate the values of n_0 and p_0 .
- 4.23** (a) The Fermi energy level in silicon at $T = 300$ K is 0.22 eV above the intrinsic Fermi level. Determine n_0 and p_0 . (b) Repeat part (a) for GaAs.
- 4.24** Silicon at $T = 300$ K is doped with boron atoms such that the concentration of holes is $p_0 = 5 \times 10^{15} \text{ cm}^{-3}$. (a) Find $E_F - E_v$. (b) Determine $E_c - E_F$. (c) Determine n_0 . (d) Which carrier is the majority carrier? (e) Determine $E_{Fi} - E_F$.
- 4.25** Repeat Problem 4.24 for $T = 400$ K, assuming the hole concentration remains constant.
- 4.26** (a) Determine the values of n_0 and p_0 in GaAs at $T = 300$ K if $E_F - E_v = 0.25$ eV. (b) Assuming the value of p_0 in part (a) remains constant, determine the values of $E_F - E_v$ and n_0 at $T = 400$ K.
- 4.27** Repeat Problem 4.26 for silicon.
- 4.28** (a) Assume that $E_F = E_c + kT/2$ at $T = 300$ K in silicon. Determine n_0 . (b) Repeat part (a) for GaAs.
- 4.29** Consider silicon at $T = 300$ K in which the hole concentration is $p_0 = 5 \times 10^{19} \text{ cm}^{-3}$. Determine $E_v - E_F$.
- 4.30** (a) In silicon at $T = 300$ K, we find that $E_F - E_c = 4 kT$. Determine the electron concentration. (b) Repeat part (a) for GaAs.

Section 4.4 Statistics of Donors and Acceptors

- ***4.31** The electron and hole concentrations as a function of energy in the conduction band and valence band peak at a particular energy as shown in Figure 4.8. Consider silicon and assume $E_c - E_F = 0.20$ eV. Determine the energy, relative to the band edges, at which the concentrations peak.
- ***4.32** For the Boltzmann approximation to be valid for a semiconductor, the Fermi level must be at least $3 kT$ below the donor level in an n-type material and at least $3 kT$ above the acceptor level in a p-type material. If $T = 300$ K, determine the maximum electron concentration in an n-type semiconductor and the maximum hole concentration in a p-type semiconductor for the Boltzmann approximation to be valid in (a) silicon and (b) gallium arsenide.
- 4.33** Plot the ratio of un-ionized donor atoms to the total electron concentration versus temperature for silicon over the range $50 \leq T \leq 200$ K.

Section 4.5 Charge Neutrality

- 4.34** Determine the equilibrium electron and hole concentrations in silicon for the following conditions:
- (a) $T = 300$ K, $N_d = 10^{15} \text{ cm}^{-3}$, $N_a = 4 \times 10^{15} \text{ cm}^{-3}$
- (b) $T = 300$ K, $N_d = 3 \times 10^{16} \text{ cm}^{-3}$, $N_a = 0$

*Asterisks next to problems indicate problems that are more difficult.

- (c) $T = 300 \text{ K}$, $N_d = N_a = 2 \times 10^{15} \text{ cm}^{-3}$
 (d) $T = 375 \text{ K}$, $N_d = 0$, $N_a = 4 \times 10^{15} \text{ cm}^{-3}$
 (e) $T = 450 \text{ K}$, $N_d = 10^{14} \text{ cm}^{-3}$, $N_a = 0$
- 4.35** Repeat Problem 4.34 for GaAs.
- 4.36** (a) Consider a germanium semiconductor at $T = 300 \text{ K}$. Calculate the thermal equilibrium electron and hole concentrations for (i) $N_d = 2 \times 10^{15} \text{ cm}^{-3}$, $N_a = 0$, and (ii) $N_a = 10^{16} \text{ cm}^{-3}$, $N_d = 7 \times 10^{15} \text{ cm}^{-3}$. (b) Repeat part (a) for GaAs. (c) For the case of GaAs in part (b), the minority carrier concentrations are on the order of 10^{-3} cm^{-3} . What does this result mean physically?
- ***4.37** The Fermi level in n-type silicon at $T = 300 \text{ K}$ is 245 meV below the conduction band and 200 meV below the donor level. Determine the probability of finding an electron (a) in the donor level and (b) in a state in the conduction-band kT above the conduction-band edge.
- 4.38** Assume that silicon, germanium, and gallium arsenide each have dopant concentrations of $N_d = 1 \times 10^{13} \text{ cm}^{-3}$ and $N_a = 2.5 \times 10^{13} \text{ cm}^{-3}$ at $T = 300 \text{ K}$. For each of the three materials: (a) Is this material n type or p type? (b) Calculate n_0 and p_0 .
- 4.39** A silicon semiconductor material at $T = 300 \text{ K}$ is doped with arsenic atoms to a concentration of $2 \times 10^{15} \text{ cm}^{-3}$ and with boron atoms to a concentration of $1.2 \times 10^{15} \text{ cm}^{-3}$. (a) Is the material n type or p type? (b) Determine n_0 and p_0 . (c) Additional boron atoms are to be added such that the hole concentration is $4 \times 10^{15} \text{ cm}^{-3}$. What concentration of boron atoms must be added and what is the new value of n_0 ?
- 4.40** The thermal equilibrium hole concentration in silicon at $T = 300 \text{ K}$ is $p_0 = 2 \times 10^5 \text{ cm}^{-3}$. Determine the thermal-equilibrium electron concentration. Is the material n type or p type?
- 4.41** In a germanium sample at $T = 250 \text{ K}$, it is found that $p_0 = 4n_0$ and that $N_d = 0$. Determine p_0 , n_0 , and N_a .
- 4.42** Consider a sample of silicon doped at $N_d = 0$ and $N_a = 10^{14} \text{ cm}^{-3}$. Plot the majority carrier concentration versus temperature over the range $200 \leq T \leq 500 \text{ K}$.
- 4.43** The temperature of a sample of silicon is $T = 300 \text{ K}$ and the acceptor doping concentration is $N_a = 0$. Plot the minority carrier concentration (on a log–log plot) versus N_d over the range $10^{15} \leq N_d \leq 10^{18} \text{ cm}^{-3}$.
- 4.44** Repeat problem 4.43 for GaAs.
- 4.45** A particular semiconductor material is doped at $N_d = 2 \times 10^{14} \text{ cm}^{-3}$ and $N_a = 1.2 \times 10^{14} \text{ cm}^{-3}$. The thermal equilibrium electron concentration is found to be $n_0 = 1.1 \times 10^{14} \text{ cm}^{-3}$. Assuming complete ionization, determine the intrinsic carrier concentration and the thermal equilibrium hole concentration.
- 4.46** (a) Silicon at $T = 300 \text{ K}$ is uniformly doped with boron atoms to a concentration of $3 \times 10^{16} \text{ cm}^{-3}$ and with arsenic atoms to a concentration of $1.5 \times 10^{16} \text{ cm}^{-3}$. Is the material n type or p type? Calculate the thermal equilibrium concentrations of majority and minority carriers. (b) Additional impurity atoms are added such that holes are the majority carrier and the thermal equilibrium concentration is $p_0 = 5 \times 10^{16} \text{ cm}^{-3}$. What type and concentration of impurity atoms must be added? What is the new value of n_0 ?
- 4.47** In silicon at $T = 300 \text{ K}$, it is found that $N_a = 7 \times 10^{15} \text{ cm}^{-3}$ and $p_0 = 2 \times 10^4 \text{ cm}^{-3}$. (a) Is the material n type or p type? (b) What are the majority and minority carrier concentrations? (c) What must be the concentration of donor impurities?

Section 4.6 Position of Fermi Energy Level

- 4.48** Consider germanium with an acceptor concentration of $N_a = 10^{15} \text{ cm}^{-3}$ and a donor concentration of $N_d = 0$. Consider temperatures of $T = 200, 400$, and 600 K . Calculate the position of the Fermi energy with respect to the intrinsic Fermi level at these temperatures.
- 4.49** Consider silicon at $T = 300 \text{ K}$ with donor concentrations of $N_d = 10^{14}, 10^{15}, 10^{16}$, and 10^{17} cm^{-3} . Assume $N_a = 0$. (a) Calculate the position of the Fermi energy level with respect to the conduction band for these donor concentrations. (b) Determine the position of the Fermi energy level with respect to the intrinsic Fermi energy level for the donor concentrations given in part (a).
- 4.50** A silicon device is doped with donor impurity atoms at a concentration of 10^{15} cm^{-3} . For the device to operate properly, the intrinsic carriers must contribute no more than 5 percent to the total electron concentration. (a) What is the maximum temperature that the device may operate? (b) What is the change in $E_c - E_F$ from the $T = 300 \text{ K}$ value to the maximum temperature value determined in part (a). (c) Is the Fermi level closer or further from the intrinsic value at the higher temperature?
- 4.51** Silicon is doped with acceptor impurity atoms at a concentration of $N_a = 3 \times 10^{15} \text{ cm}^{-3}$. Assume $N_d = 0$. Plot the position of the Fermi energy level with respect to the intrinsic Fermi energy level over the temperature range of $200 \leq T \leq 600 \text{ K}$.
- 4.52** Consider GaAs at $T = 300 \text{ K}$ with $N_d = 0$. (a) Plot the position of the Fermi energy level with respect to the intrinsic Fermi energy level as a function of the acceptor impurity concentration over the range of $10^{14} \leq N_a \leq 10^{17} \text{ cm}^{-3}$. (b) Plot the position of the Fermi energy level with respect to the valence-band energy over the same acceptor impurity concentration as given in part (a).
- 4.53** For a particular semiconductor, $E_g = 1.50 \text{ eV}$, $m_p^* = 10 m_n^*$, $T = 300 \text{ K}$, and $n_i = 1 \times 10^5 \text{ cm}^{-3}$. (a) Determine the position of the intrinsic Fermi energy level with respect to the center of the bandgap. (b) Impurity atoms are added so that the Fermi energy level is 0.45 eV below the center of the bandgap. (i) Are acceptor or donor atoms added? (ii) What is the concentration of impurity atoms added?
- 4.54** Silicon at $T = 300 \text{ K}$ contains acceptor atoms at a concentration of $N_a = 5 \times 10^{15} \text{ cm}^{-3}$. Donor atoms are added forming an n-type compensated semiconductor such that the Fermi level is 0.215 eV below the conduction-band edge. What concentration of donor atoms are added?
- 4.55** (a) Silicon at $T = 300 \text{ K}$ is doped with donor impurity atoms at a concentration of $N_d = 6 \times 10^{15} \text{ cm}^{-3}$. (i) Determine $E_c - E_F$. (ii) Calculate the concentration of additional donor impurity atoms that must be added to move the Fermi energy level a distance kT closer to the conduction band edge. (b) Repeat part (a) for GaAs if the original donor impurity concentration is $N_d = 1 \times 10^{15} \text{ cm}^{-3}$.
- 4.56** (a) Determine the position of the Fermi energy level with respect to the intrinsic Fermi level in silicon at $T = 300 \text{ K}$ that is doped with boron atoms at a concentration of $N_a = 2 \times 10^{16} \text{ cm}^{-3}$. (b) Repeat part (a) if the silicon is doped with phosphorus atoms at a concentration of $N_d = 2 \times 10^{16} \text{ cm}^{-3}$. (c) Calculate n_0 and p_0 in parts (a) and (b).
- 4.57** GaAs at $T = 300 \text{ K}$ is doped with donor impurity atoms at a concentration of $7 \times 10^{15} \text{ cm}^{-3}$. Additional impurity atoms are to be added such that the Fermi level is 0.55 eV above the intrinsic Fermi level. Determine the type (donor or acceptor) and concentration of impurity atoms to be added.

- 4.58** Determine the Fermi energy level with respect to the intrinsic Fermi level for each condition given in Problem 4.34.
- 4.59** Find the Fermi energy level with respect to the valence-band energy for the conditions given in Problem 4.35.
- 4.60** Calculate the position of the Fermi energy level with respect to the intrinsic Fermi for the conditions given in Problem 4.47.

Summary and Review

- 4.61** A new semiconductor material is to be “designed.” The semiconductor is to be p type and doped with $5 \times 10^{15} \text{ cm}^{-3}$ acceptor atoms. Assume complete ionization and assume $N_d = 0$. The effective density of states functions are $N_c = 1.2 \times 10^{19} \text{ cm}^{-3}$ and $N_v = 1.8 \times 10^{19} \text{ cm}^{-3}$ at $T = 300 \text{ K}$ and vary as T^2 . A special semiconductor device fabricated with this material requires that the hole concentration be no greater than $5.08 \times 10^{15} \text{ cm}^{-3}$ at $T = 350 \text{ K}$. What is the minimum bandgap energy required in this new material?
- 4.62** Silicon atoms, at a concentration of $7 \times 10^{15} \text{ cm}^{-3}$, are added to gallium arsenide. Assume that the silicon atoms act as fully ionized dopant atoms and that 5 percent of the concentration added replace gallium atoms and 95 percent replace arsenic atoms. Let $T = 300 \text{ K}$. (a) Determine the donor and acceptor concentrations. (b) Is the material n type or p type? (c) Calculate the electron and hole concentrations. (d) Determine the position of the Fermi level with respect to E_{Fi} .
- 4.63** Defects in a semiconductor material introduce allowed energy states within the forbidden bandgap. Assume that a particular defect in silicon introduces two discrete levels: a donor level 0.25 eV above the top of the valence band and an acceptor level 0.65 eV above the top of the valence band. The charge state of each defect is a function of the position of the Fermi level. (a) Sketch the charge density of each defect as the Fermi level moves from E_v to E_c . Which defect level dominates in (i) heavily doped n type material and (ii) in heavily doped p-type material? (b) Determine the electron and hole concentrations and the location of the Fermi level in (i) an n type sample doped at $N_d = 10^{17} \text{ cm}^{-3}$ and (ii) in a p-type sample doped at $N_a = 10^{17} \text{ cm}^{-3}$. (c) Determine the Fermi-level position if no dopant atoms are added. Is the material n type, p type, or intrinsic?

READING LIST

- *1. Hess, K. *Advanced Theory of Semiconductor Devices*. Englewood Cliffs, NJ: Prentice Hall, 1988.
- 2. Hu, C. C. *Modern Semiconductor Devices for Integrated Circuits*. Upper Saddle River, NJ: Pearson Prentice Hall, 2010.
- 3. Kano, K. *Semiconductor Devices*. Upper Saddle River, NJ: Prentice Hall, 1998.
- *4. Li, S. S. *Semiconductor Physical Electronics*. New York: Plenum Press, 1993.
- 5. McKelvey, J. P. *Solid State Physics for Engineering and Materials Science*. Malabar, FL.: Krieger Publishing, 1993.
- 6. Navon, D. H. *Semiconductor Microdevices and Materials*. New York: Holt, Rinehart & Winston, 1986.

7. Pierret, R. F. *Semiconductor Device Fundamentals*. Reading, MA: Addison-Wesley, 1996.
8. Shur, M. *Introduction to Electronic Devices*. New York: John Wiley and Sons, 1996.
- *9. Shur, M. *Physics of Semiconductor Devices*. Englewood Cliffs, NJ: Prentice Hall, 1990.
10. Singh, J. *Semiconductor Devices: An Introduction*. New York: McGraw-Hill, 1994.
11. Singh, J. *Semiconductor Devices: Basic Principles*. New York: John Wiley and Sons, 2001.
- *12. Smith, R. A. *Semiconductors*. 2nd ed. New York: Cambridge University Press, 1978.
13. Streetman, B. G., and S. Banerjee. *Solid State Electronic Devices*, 6th ed. Upper Saddle River, NJ: Pearson Prentice Hall, 2006.
14. Sze, S. M., and K. K. Ng. *Physics of Semiconductor Devices*. 3rd ed. Hoboken, NJ: John Wiley and Sons, 2007.
- *15. Wang, S. *Fundamentals of Semiconductor Theory and Device Physics*. Englewood Cliffs, NJ: Prentice Hall, 1989.
- *16. Wolfe, C. M., N. Holonyak, Jr., and G. E. Stillman. *Physical Properties of Semiconductors*. Englewood Cliffs, NJ: Prentice Hall, 1989.
17. Yang, E. S. *Microelectronic Devices*. New York: McGraw-Hill, 1988.

*Indicates references that are at an advanced level compared to this text.

Carrier Transport Phenomena

In the previous chapter, we considered the semiconductor in equilibrium and determined electron and hole concentrations in the conduction and valence bands, respectively. A knowledge of the densities of these charged particles is important toward an understanding of the electrical properties of a semiconductor material. The net flow of the electrons and holes in a semiconductor will generate currents. The process by which these charged particles move is called *transport*. In this chapter we consider two basic transport mechanisms in a semiconductor crystal: drift—the movement of charge due to electric fields, and diffusion—the flow of charge due to density gradients.

The carrier transport phenomena are the foundation for finally determining the current–voltage characteristics of semiconductor devices. We will implicitly assume in this chapter that, although there will be a net flow of electrons and holes due to the transport processes, thermal equilibrium will not be substantially disturbed. Non-equilibrium processes are considered in the next chapter. ■

5.0 | PREVIEW

In this chapter, we will:

- Describe the mechanism of carrier drift and induced drift current due to an applied electric field.
- Define and describe the characteristics of carrier mobility.
- Describe the mechanism of carrier diffusion and induced diffusion current due to a gradient in the carrier concentration.
- Define the carrier diffusion coefficient.
- Describe the effects of a nonuniform impurity doping concentration in a semiconductor material.
- Discuss and analyze the Hall effect in a semiconductor material.

5.1 | CARRIER DRIFT

An electric field applied to a semiconductor will produce a force on electrons and holes so that they will experience a net acceleration and net movement, provided there are available energy states in the conduction and valence bands. This **net movement of charge due to an electric field is called *drift***. The net drift of charge gives rise to a *drift current*.

5.1.1 Drift Current Density

If we have a positive volume charge density ρ moving at an average drift velocity v_d , the drift current density is given by

$$J_{drf} = \rho v_d \quad (5.1a)$$

In terms of units, we have

$$J_{drf} = \left(\frac{\text{Coul}}{\text{cm}^3} \right) \cdot \left(\frac{\text{cm}}{\text{s}} \right) = \frac{\text{Coul}}{\text{cm}^2 \cdot \text{s}} = \frac{\text{A}}{\text{cm}^2} \quad (5.1b)$$

If the volume charge density is due to positively charged holes, then

$$J_{p|drf} = (ep)v_{dp} \quad (5.2)$$

where $J_{p|drf}$ is the drift current density due to holes and v_{dp} is the average drift velocity of the holes.

The equation of motion of a positively charged hole in the presence of an electric field is

$$F = m_{cp}^* a = eE \quad (5.3)$$

where e is the magnitude of the electronic charge, a is the acceleration, E is the electric field, and m_{cp}^* is the conductivity effective mass of the hole.¹ If the electric field is constant, then we expect the velocity to increase linearly with time. However, charged particles in a semiconductor are involved in collisions with ionized impurity atoms and with thermally vibrating lattice atoms. These collisions, or scattering events, alter the velocity characteristics of the particle.

As the hole accelerates in a crystal due to the electric field, the velocity increases. When the charged particle collides with an atom in the crystal, for example, the particle loses most, or all, of its energy. The particle will again begin to accelerate and gain energy until it is again involved in a scattering process. This continues over and over again. Throughout this process, the particle will gain an average drift velocity which, for low electric fields, is directly proportional to the electric field. We may then write

$$v_{dp} = \mu_p E \quad (5.4)$$

where μ_p is the proportionality factor and is called the **hole mobility**. The mobility is an important parameter of the semiconductor since it describes how well a particle

¹The conductivity effective mass is used when carriers are in motion. See Appendix F for further discussion of effective mass concepts.

Table 5.1 | Typical mobility values at $T = 300$ K and low doping concentrations

	μ_n (cm ² /V-s)	μ_p (cm ² /V-s)
Silicon	1350	480
Gallium arsenide	8500	400
Germanium	3900	1900

will move due to an electric field. The unit of mobility is usually expressed in terms of cm²/V-s.

By combining Equations (5.2) and (5.4), we may write the drift current density due to holes as

$$J_{p|drf} = (ep)v_{dp} = e\mu_p p E \quad (5.5)$$

The drift current due to holes is in the same direction as the applied electric field.

The same discussion of drift applies to electrons. We may write

$$J_{n|drf} = \rho v_{dn} = (-en)v_{dn} \quad (5.6)$$

where $J_{n|drf}$ is the drift current density due to electrons and v_{dn} is the average drift velocity of electrons. The net charge density of electrons is negative.

The average drift velocity of an electron is also proportional to the electric field for small fields. However, since the electron is negatively charged, the net motion of the electron is opposite to the electric field direction. We can then write

$$v_{dn} = -\mu_n E \quad (5.7)$$

where μ_n is the *electron mobility* and is a positive quantity. Equation (5.6) may now be written as

$$J_{n|drf} = (-en)(-\mu_n E) = e\mu_n n E \quad (5.8)$$

The conventional drift current due to electrons is also in the same direction as the applied electric field even though the electron movement is in the opposite direction.

Electron and hole mobilities are functions of temperature and doping concentrations, as we will see in the next section. Table 5.1 shows some typical mobility values at $T = 300$ K for low doping concentrations.

Since both electrons and holes contribute to the drift current, the total *drift current density* is the sum of the individual electron and hole drift current densities, so we may write

$$J_{drf} = e(\mu_n n + \mu_p p)E \quad (5.9)$$

EXAMPLE 5.1

Objective: Calculate the drift current density in a semiconductor for a given electric field.

Consider a gallium arsenide sample at $T = 300$ K with doping concentrations of $N_a = 0$ and $N_d = 10^{16}$ cm⁻³. Assume complete ionization and assume electron and hole mobilities given in Table 5.1. Calculate the drift current density if the applied electric field is $E = 10$ V/cm.

■ Solution

Since $N_d > N_a$, the semiconductor is n type and the majority carrier electron concentration, from Chapter 4 is given by

$$n = \frac{N_d - N_a}{2} + \sqrt{\left(\frac{N_d - N_a}{2}\right)^2 + n_i^2} \approx 10^{16} \text{ cm}^{-3}$$

The minority carrier hole concentration is

$$p = \frac{n_i^2}{n} = \frac{(1.8 \times 10^6)^2}{10^{16}} = 3.24 \times 10^{-4} \text{ cm}^{-3}$$

For this extrinsic n-type semiconductor, the drift current density is

$$J_{drf} = e(\mu_n n + \mu_p p)E \approx e\mu_n N_d E$$

Then

$$J_{drf} = (1.6 \times 10^{-19})(8500)(10^{16})(10) = 136 \text{ A/cm}^2$$

■ Comment

Significant drift current densities can be obtained in a semiconductor applying relatively small electric fields. We may note from this example that the drift current will usually be due primarily to the majority carrier in an extrinsic semiconductor.

■ EXERCISE PROBLEM

Ex 5.1 A drift current density of $J_{drf} = 75 \text{ A/cm}^2$ is required in a device using p-type silicon when an electric field of $E = 120 \text{ V/cm}$ is applied. Determine the required impurity doping concentration to achieve this specification. Assume that electron and hole mobilities given in Table 5.1 apply. (Ans. $N_a = 8.14 \times 10^{15} \text{ cm}^{-3}$)

5.1.2 Mobility Effects

In the previous section, we defined mobility, which relates the average drift velocity of a carrier to the electric field. Electron and hole mobilities are important semiconductor parameters in the characterization of carrier drift, as seen in Equation (5.9).

Equation (5.3) related the acceleration of a hole to a force such as an electric field. We may write this equation as

$$F = m_{cp}^* \frac{dv}{dt} = eE \quad (5.10)$$

where v is the velocity of the particle due to the electric field and does not include the random thermal velocity. If we assume that the conductivity effective mass and electric field are constants, then we may integrate Equation (5.10) and obtain

$$v = \frac{eEt}{m_{cp}^*} \quad (5.11)$$

where we have assumed the initial drift velocity to be zero.

Figure 5.1a shows a schematic model of the random thermal velocity and motion of a hole in a semiconductor with zero electric field. There is a mean time

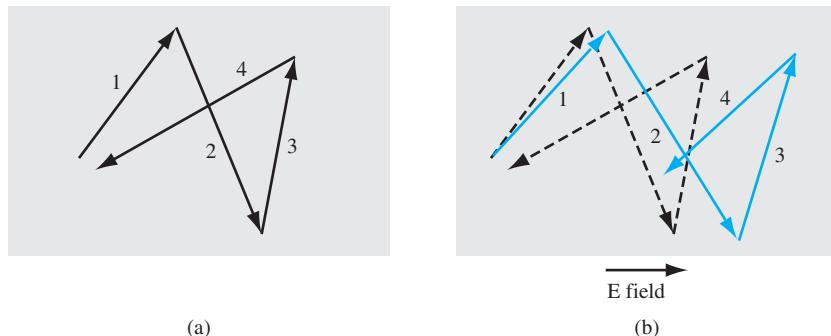


Figure 5.1 | Typical random behavior of a hole in a semiconductor (a) without an electric field and (b) with an electric field.

between collisions which may be denoted by τ_{cp} . If a small electric field (E-field) is applied as indicated in Figure 5.1b, there will be a net drift of the hole in the direction of the E-field, and the net drift velocity will be a small perturbation on the random thermal velocity, so the time between collisions will not be altered appreciably. If we use the mean time between collisions τ_{cp} in place of the time t in Equation (5.11), then the mean peak velocity just prior to a collision or scattering event is

$$v_{d|\text{peak}} = \left(\frac{e\tau_{cp}}{m_{cp}^*} \right) E \quad (5.12a)$$

The average drift velocity is one half the peak value so that we can write

$$\langle v_d \rangle = \frac{1}{2} \left(\frac{e\tau_{cp}}{m_{cp}^*} \right) E \quad (5.12b)$$

However, the collision process is not as simple as this model, but is statistical in nature. In a more accurate model including the effect of a statistical distribution, the factor $\frac{1}{2}$ in Equation (5.12b) does not appear. The hole mobility is then given by

$$\mu_p = \frac{v_{dp}}{E} = \frac{e\tau_{cp}}{m^*_m} \quad (5.13)$$

The same analysis applies to electrons; thus, we can write the electron mobility as

$$\mu_n = \frac{e\tau_{cn}}{m_{cn}^*} \quad (5.14)$$

where τ_{cn} is the mean time between collisions for an electron.

There are two collision or scattering mechanisms that dominate in a semiconductor and affect the carrier mobility: phonon or lattice scattering, and ionized impurity scattering.

The atoms in a semiconductor crystal have a certain amount of thermal energy at temperatures above absolute zero that causes the atoms to randomly vibrate about their lattice position within the crystal. The lattice vibrations cause a disruption in the perfect periodic potential function. A perfect periodic potential in a solid allows

electrons to move unimpeded, or with no scattering, through the crystal. But the thermal vibrations cause a disruption of the potential function, resulting in an interaction between the electrons or holes and the vibrating lattice atoms. This *lattice scattering* is also referred to as *phonon scattering*.

Since lattice scattering is related to the thermal motion of atoms, the rate at which the scattering occurs is a function of temperature. If we denote μ_L as the mobility that would be observed if only lattice scattering existed, then the scattering theory states that to first order

$$\mu_L \propto T^{-3/2} \quad (5.15)$$

Mobility that is due to lattice scattering increases as the temperature decreases. Intuitively, we expect the lattice vibrations to decrease as the temperature decreases, which implies that the probability of a scattering event also decreases, thus increasing mobility.

Figure 5.2 shows the temperature dependence of electron and hole mobilities in silicon. In lightly doped semiconductors, lattice scattering dominates and the carrier mobility decreases with temperature as we have discussed. The temperature dependence of mobility is proportional to T^{-n} . The inserts in the figure show that the parameter n is not equal to $\frac{3}{2}$ as the first-order scattering theory predicted. However, mobility does increase as the temperature decreases.

The second interaction mechanism affecting carrier mobility is called *ionized impurity scattering*. We have seen that impurity atoms are added to the semiconductor to control or alter its characteristics. These impurities are ionized at room temperature so that a coulomb interaction exists between the electrons or holes and the ionized impurities. This coulomb interaction produces scattering or collisions and also alters the velocity characteristics of the charge carrier. If we denote μ_I as the mobility that would be observed if only ionized impurity scattering existed, then to first order we have

$$\mu_I \propto \frac{T^{+3/2}}{N_I} \quad (5.16)$$

where $N_I = N_d^+ + N_a^-$ is the total ionized impurity concentration in the semiconductor. If temperature increases, the random thermal velocity of a carrier increases, reducing the time the carrier spends in the vicinity of the ionized impurity center. The less time spent in the vicinity of a coulomb force, the smaller the scattering effect and the larger the expected value of μ_I . If the number of ionized impurity centers increases, then the probability of a carrier encountering an ionized impurity center increases, implying a smaller value of μ_I .

Figure 5.3 is a plot of electron and hole mobilities in germanium, silicon, and gallium arsenide at $T = 300$ K as a function of impurity concentration. More accurately, these curves are of mobility versus ionized impurity concentration N_I . As the impurity concentration increases, the number of impurity scattering centers increases, thus reducing mobility.

If τ_L is the mean time between collisions due to lattice scattering, then dt/τ_L is the probability of a lattice scattering event occurring in a differential time dt . Likewise, if τ_I is the mean time between collisions due to ionized impurity scattering, then dt/τ_I is the probability of an ionized impurity scattering event occurring in the differential time dt .

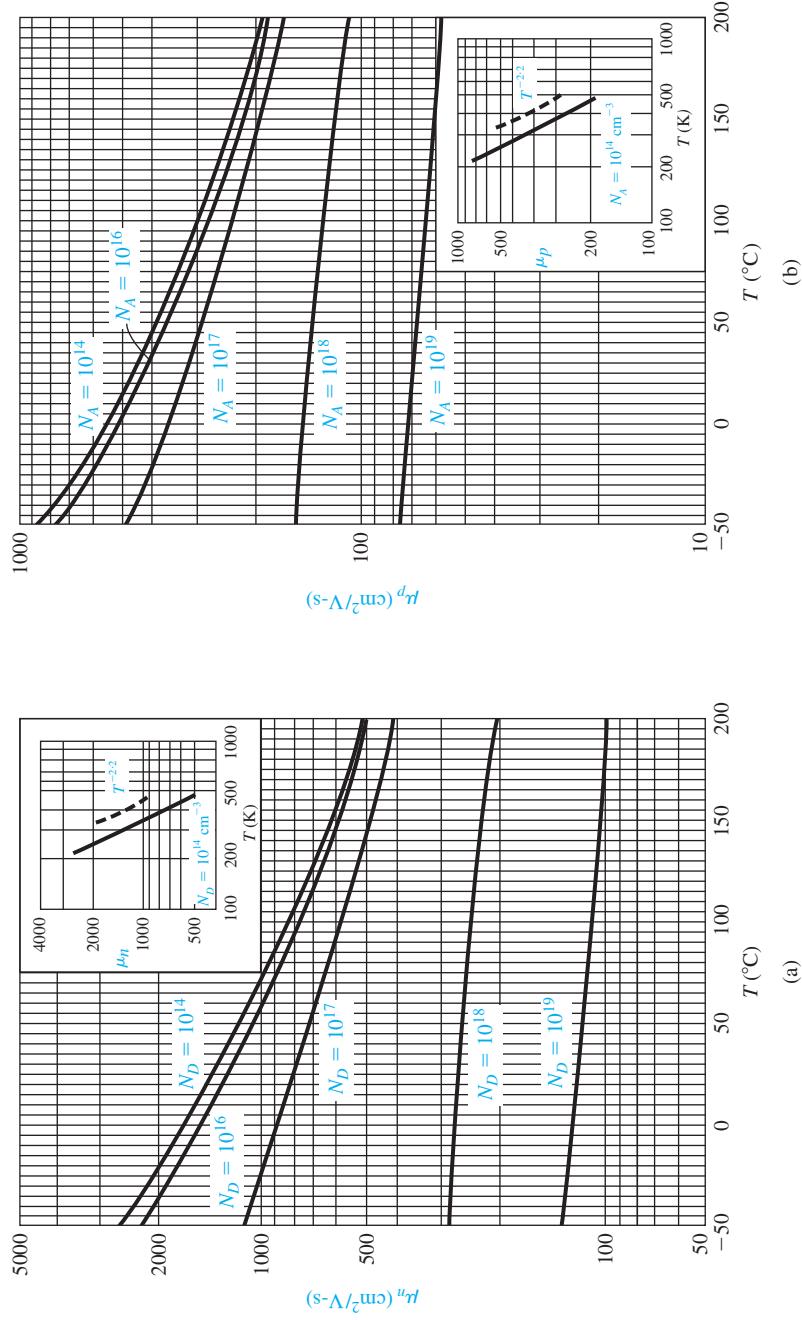


Figure 5.2 | (a) Electron and (b) hole mobilities in silicon versus temperature for various doping concentrations. Insets show temperature dependence for “almost” intrinsic silicon. (From Pierret [8].)

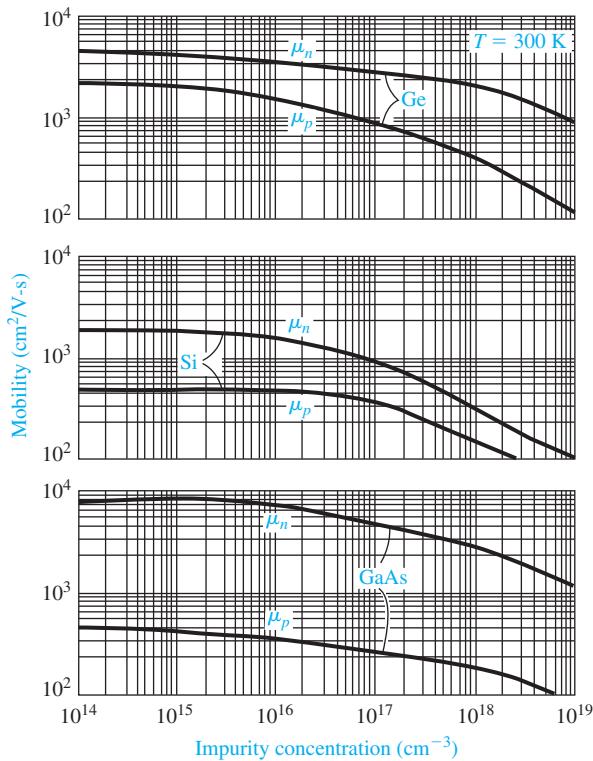


Figure 5.3 | Electron and hole mobilities versus impurity concentrations for germanium, silicon, and gallium arsenide at $T = 300$ K.
(From Sze [14].)

If these two scattering processes are independent, then the total probability of a scattering event occurring in the differential time dt is the sum of the individual events, or

$$\frac{dt}{\tau} = \frac{dt}{\tau_I} + \frac{dt}{\tau_L} \quad (5.17)$$

where τ is the mean time between any scattering event.

Comparing Equation (5.17) with the definitions of mobility given by Equation (5.13) or (5.14), we can write

$$\frac{1}{\mu} = \frac{1}{\mu_I} + \frac{1}{\mu_L} \quad (5.18)$$

where μ_I is the mobility due to the ionized impurity scattering process and μ_L is the mobility due to the lattice scattering process. The parameter μ is the net mobility. With two or more independent scattering mechanisms, the inverse mobilities add, which means that the net mobility decreases.

EXAMPLE 5.2

Objective: Determine the electron mobility in silicon at various doping concentrations and various temperatures.

Using Figure 5.2, find the electron mobility in silicon for:

- (a) $T = 25^\circ\text{C}$ for (i) $N_d = 10^{16} \text{ cm}^{-3}$ and (ii) $N_d = 10^{17} \text{ cm}^{-3}$.
- (b) $N_d = 10^{16} \text{ cm}^{-3}$ for (i) $T = 0^\circ\text{C}$ and (ii) $T = 100^\circ\text{C}$.

■ Solution:

From Figure 5.2, we find the following:

- (a) $T = 25^\circ\text{C}$; (i) $N_d = 10^{16} \text{ cm}^{-3} \Rightarrow \mu_n \approx 1200 \text{ cm}^2/\text{V}\cdot\text{s}$.
 (ii) $N_d = 10^{17} \text{ cm}^{-3} \Rightarrow \mu_n \approx 800 \text{ cm}^2/\text{V}\cdot\text{s}$.
- (b) $N_d = 10^{16} \text{ cm}^{-3}$; (i) $T = 0^\circ\text{C} \Rightarrow \mu_n \approx 1400 \text{ cm}^2/\text{V}\cdot\text{s}$.
 (ii) $T = 100^\circ\text{C} \Rightarrow \mu_n \approx 780 \text{ cm}^2/\text{V}\cdot\text{s}$.

■ Comment

The results of this example show that the mobility values are strong functions of the doping concentration and temperature. These variations must be taken into account in the design of semiconductor devices.

■ EXERCISE PROBLEM

Ex 5.2 Using Figure 5.2, find the hole mobility in silicon for:

- (a) $T = 25^\circ\text{C}$ for (i) $N_a = 10^{16} \text{ cm}^{-3}$ and (ii) $N_a = 10^{18} \text{ cm}^{-3}$, and
 - (b) $N_a = 10^{14} \text{ cm}^{-3}$ for (i) $T = 0^\circ\text{C}$ and (ii) $T = 100^\circ\text{C}$.
- [Ans. (a) (i) $\mu_p = 550 \text{ cm}^2/\text{V}\cdot\text{s}$, (ii) $\mu_p = 300 \text{ cm}^2/\text{V}\cdot\text{s}$;
 (b) (i) $\mu_p = 410 \text{ cm}^2/\text{V}\cdot\text{s}$, (ii) $\mu_p = 130 \text{ cm}^2/\text{V}\cdot\text{s}$]

5.1.3 Conductivity

The drift current density, given by Equation (5.9), may be written as

$$J_{drf} = e(\mu_n n + \mu_p p)E = \sigma E \quad (5.19)$$

where σ is the *conductivity* of the semiconductor material. The conductivity is given in units of $(\Omega\cdot\text{cm})^{-1}$ and is a function of the electron and hole concentrations and mobilities. We have just seen that the mobilities are functions of impurity concentrations; conductivity, then is a somewhat complicated function of impurity concentration.

The reciprocal of conductivity is *resistivity*, which is denoted by ρ and is given in units of ohm-cm. We can write the formula for resistivity as²

$$\rho = \frac{1}{\sigma} = \frac{1}{e(\mu_n n + \mu_p p)} \quad (5.20)$$

Figure 5.4 is a plot of resistivity as a function of impurity concentration in silicon, germanium, gallium arsenide, and gallium phosphide at $T = 300 \text{ K}$. Obviously, the curves are not linear functions of N_d or N_a because of mobility effects.

²The symbol ρ is also used for volume charge density. The context in which ρ is used should make it clear whether it stands for charge density or resistivity.

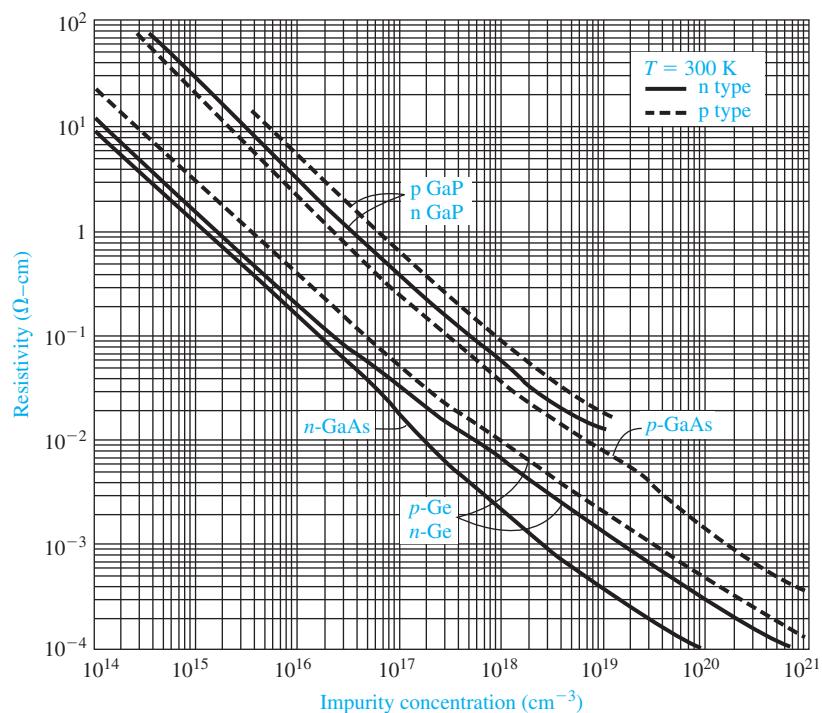
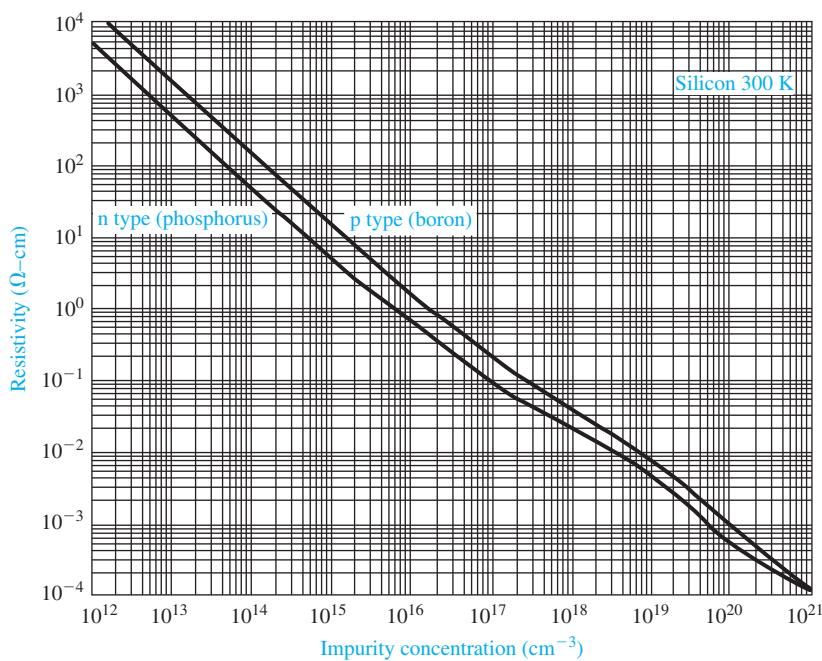


Figure 5.4 | Resistivity versus impurity concentration at $T = 300 \text{ K}$ in (a) silicon and (b) germanium, gallium arsenide, and gallium phosphide.
(From Sze [14].)

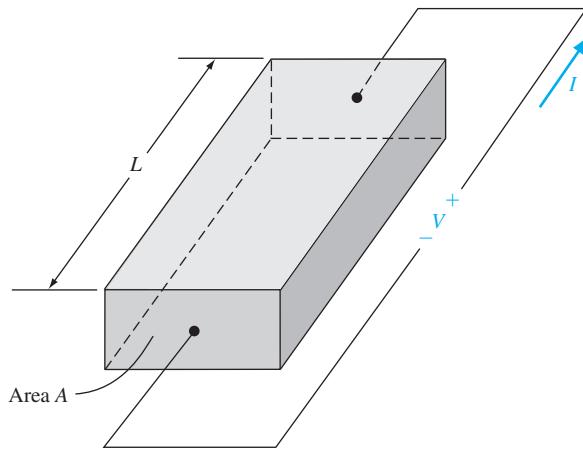


Figure 5.5 | Bar of semiconductor material as a resistor.

If we have a bar of semiconductor material as shown in Figure 5.5 with a voltage applied that produces a current I , then we can write

$$J = \frac{I}{A} \quad (5.21a)$$

and

$$E = \frac{V}{L} \quad (5.21b)$$

We can now rewrite Equation (5.19) as

$$\frac{I}{A} = \sigma \left(\frac{V}{L} \right) \quad (5.22a)$$

or

$$V = \left(\frac{L}{\sigma A} \right) I = \left(\frac{\rho L}{A} \right) I = IR \quad (5.22b)$$

Equation (5.22b) is Ohm's law for a semiconductor. The resistance is a function of resistivity, or conductivity, as well as the geometry of the semiconductor.

If we consider, for example, a **p-type semiconductor** with an acceptor doping $N_a (N_d = 0)$ in which $N_a \gg n_i$, and if we assume that the electron and hole mobilities are of the same order of magnitude, then the conductivity becomes

$$\sigma = e(\mu_n n + \mu_p p) \approx e\mu_p p \quad (5.23)$$

If we also assume complete ionization, then Equation (5.23) becomes

$$\sigma \approx e\mu_p N_a \approx \frac{1}{\rho} \quad (5.24)$$

The conductivity and resistivity of an extrinsic semiconductor are a function primarily of the majority carrier parameters.

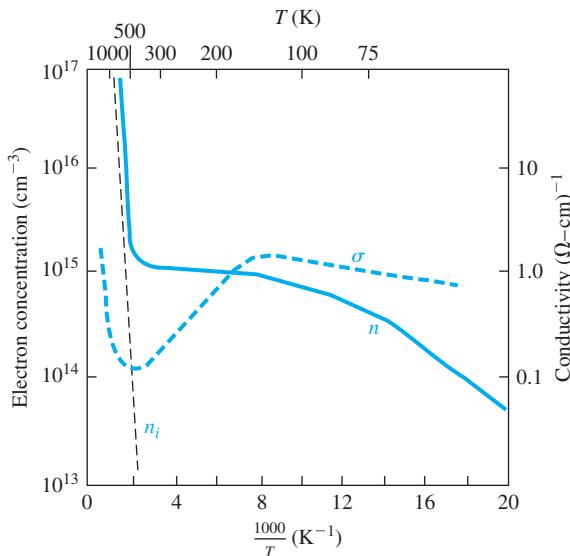


Figure 5.6 | Electron concentration and conductivity versus inverse temperature for silicon.
(After Sze [14].)

We may plot the carrier concentration and conductivity of a semiconductor as a function of temperature for a particular doping concentration. Figure 5.6 shows the electron concentration and conductivity of silicon as a function of inverse temperature for the case when $N_d = 10^{15} \text{ cm}^{-3}$. In the midtemperature range, or extrinsic range, as shown, we have complete ionization—the electron concentration remains essentially constant. However, the mobility is a function of temperature so the conductivity varies with temperature in this range. At higher temperatures, the intrinsic carrier concentration increases and begins to dominate the electron concentration as well as the conductivity. In the lower temperature range, freeze-out begins to occur; the electron concentration and conductivity decrease with decreasing temperature.

Objective: Determine the doping concentration and majority carrier mobility given the type and conductivity of a compensated semiconductor.

EXAMPLE 5.3

Consider compensated n-type silicon at $T = 300 \text{ K}$, with a conductivity of $\sigma = 16 (\Omega \cdot \text{cm})^{-1}$ and an acceptor doping concentration of 10^{17} cm^{-3} . Determine the donor concentration and the electron mobility.

■ Solution

For n-type silicon at $T = 300 \text{ K}$, we can assume complete ionization; therefore the conductivity, assuming $N_d - N_a \gg n_i$, is given by

$$\sigma \approx e\mu_n n = e\mu_n (N_d - N_a)$$

We have that

$$16 = (1.6 \times 10^{-19})\mu_n(N_d - 10^{17})$$

Since mobility is a function of the ionized impurity concentration, we can use Figure 5.3 along with trial and error to determine μ_n and N_d . For example, if we choose $N_d = 2 \times 10^{17}$, then $N_l = N_d^+ + N_a^- = 3 \times 10^{17}$ so that $\mu_n \approx 510 \text{ cm}^2/\text{V-s}$ which gives $\sigma = 8.16 (\Omega\text{-cm})^{-1}$. If we choose $N_d = 5 \times 10^{17}$, then $N_l = 6 \times 10^{17}$ so that $\mu_n \approx 325 \text{ cm}^2/\text{V-s}$, which gives $\sigma = 20.8 (\Omega\text{-cm})^{-1}$. The doping is bounded between these two values. Further trial and error yields

$$N_d \approx 3.5 \times 10^{17} \text{ cm}^{-3} \quad \text{and} \quad \mu_n \approx 400 \text{ cm}^2/\text{V-s}$$

which gives

$$\sigma \approx 16 (\Omega\text{-cm})^{-1}$$

Comment

We can see from this example that, in a high-conductivity semiconductor material, mobility is a strong function of carrier concentration.

EXERCISE PROBLEM

Ex 5.3 A compensated p-type silicon material at $T = 300 \text{ K}$ has impurity doping concentrations of $N_a = 2.8 \times 10^{17} \text{ cm}^{-3}$ and $N_d = 8 \times 10^{16} \text{ cm}^{-3}$. Determine the (a) hole mobility, (b) conductivity, and (c) resistivity.

[Ans. (a) $\mu_h \approx 200 \text{ cm}^2/\text{V-s}$; (b) $\sigma = 6.4 (\Omega\text{-cm})^{-1}$; (c) $\rho = 0.156 (\Omega\text{-cm})$]

DESIGN EXAMPLE 5.4

Objective: Design a semiconductor resistor with a specified resistance to handle a given current density.

A silicon semiconductor at $T = 300 \text{ K}$ is initially doped with donors at a concentration of $N_d = 5 \times 10^{15} \text{ cm}^{-3}$. Acceptors are to be added to form a compensated p-type material. The resistor is to have a resistance of $10 \text{ k}\Omega$ and handle a current density of 50 A/cm^2 when 5 V is applied.

Solution

For 5 V applied to a $10\text{-k}\Omega$ resistor, the total current is

$$I = \frac{V}{R} = \frac{5}{10} = 0.5 \text{ mA}$$

If the current density is limited to 50 A/cm^2 , then the cross-sectional area is

$$A = \frac{I}{J} = \frac{0.5 \times 10^{-3}}{50} = 10^{-5} \text{ cm}^2$$

If we, somewhat arbitrarily at this point, limit the electric field to $E = 100 \text{ V/cm}$, then the length of the resistor is

$$L = \frac{V}{E} = \frac{5}{100} = 5 \times 10^{-2} \text{ cm}$$

From Equation (5.22b), the conductivity of the semiconductor is

$$\sigma = \frac{L}{RA} = \frac{5 \times 10^{-2}}{(10^3)(10^{-5})} = 0.50 (\Omega\text{-cm})^{-1}$$

The conductivity of a compensated p-type semiconductor is

$$\sigma \approx e\mu_p p = e\mu_p (N_a - N_d)$$

where the mobility is a function of the total ionized impurity concentration $N_a + N_d$.

Using trial and error, if $N_a = 1.25 \times 10^{16} \text{ cm}^{-3}$, then $N_a + N_d = 1.75 \times 10^{16} \text{ cm}^{-3}$, and the hole mobility, from Figure 5.3, is approximately $\mu_p = 410 \text{ cm}^2/\text{V}\cdot\text{s}$. The conductivity is then

$$\sigma = e\mu_p(N_a - N_d) = (1.6 \times 10^{-19})(410)(1.25 \times 10^{16} - 5 \times 10^{15}) = 0.492$$

which is very close to the value we need.

Comment

Since the mobility is related to the total ionized impurity concentration, the determination of the impurity concentration to achieve a particular conductivity is not straightforward.

EXERCISE PROBLEM

Ex 5.4 A bar of p-type silicon, such as shown in Figure 5.5, has a cross-sectional area

$A = 10^{-6} \text{ cm}^2$ and a length $L = 1.2 \times 10^{-3} \text{ cm}$. For an applied voltage of 5 V, a current of 2 mA is required. What is the required (a) resistance, (b) resistivity, and

(c) impurity doping concentration? (d) What is the resulting hole mobility?

[Ans. (a) 2.5 kΩ; (b) 2.083 (Ω·cm); (c) $N_a \equiv 7.3 \times 10^{15} \text{ cm}^{-3}$; (d) $\mu_p \equiv 410 \text{ cm}^2/\text{V}\cdot\text{s}$]

For an intrinsic material, the conductivity can be written as

$$\sigma_i = e(\mu_n + \mu_p) n_i \quad (5.25)$$

The concentrations of electrons and holes are equal in an intrinsic semiconductor, so the intrinsic conductivity includes both the electron and hole mobility. Since, in general, the electron and hole mobilities are not equal, the intrinsic conductivity is not the minimum value possible at a given temperature.

5.1.4 Velocity Saturation

So far in our discussion of drift velocity, we have assumed that mobility is not a function of electric field, meaning that the drift velocity will increase linearly with applied electric field. The total velocity of a particle is the sum of the random thermal velocity and drift velocity. At $T = 300 \text{ K}$, the average random thermal energy is given by

$$\frac{1}{2}mv_{th}^2 = \frac{3}{2}kT = \frac{3}{2}(0.0259) = 0.03885 \text{ eV} \quad (5.26)$$

This energy translates into a mean thermal velocity of approximately 10^7 cm/s for an electron in silicon. If we assume an electron mobility of $\mu_n = 1350 \text{ cm}^2/\text{V}\cdot\text{s}$ in low-doped silicon, a drift velocity of 10^5 cm/s , or 1 percent of the thermal velocity, is achieved if the applied electric field is approximately 75 V/cm . This applied electric field does not appreciably alter the energy of the electron.

Figure 5.7 is a plot of average drift velocity as a function of applied electric field for electrons and holes in silicon, gallium arsenide, and germanium. At low electric fields, where there is a linear variation of velocity with electric field, the slope of

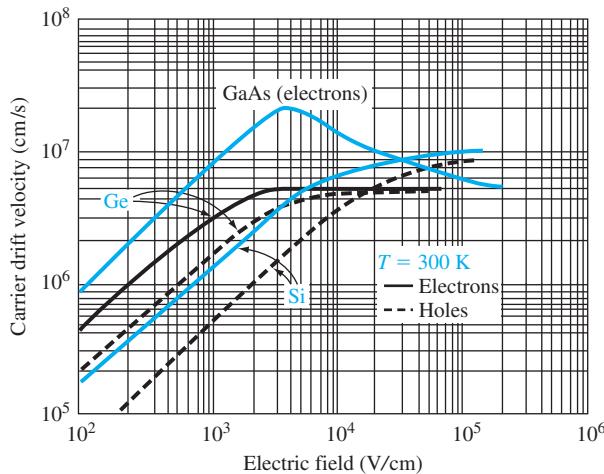


Figure 5.7 | Carrier drift velocity versus electric field for high-purity silicon, germanium, and gallium arsenide.
(From Sze [14].)

the drift velocity versus electric field curve is the mobility. The behavior of the drift velocity of carriers at high electric fields deviates substantially from the linear relationship observed at low fields. The drift velocity of electrons in silicon, for example, saturates at approximately 10^7 cm/s at an electric field of approximately 30 kV/cm. If the drift velocity of a charge carrier saturates, then the drift current density also saturates and becomes independent of the applied electric field.

The experimental carrier drift velocity versus electric field in silicon can be approximated for electrons by [2]

$$v_n = \frac{v_s}{\left[1 + \left(\frac{E_{on}}{E} \right)^2 \right]^{1/2}} \quad (5.27a)$$

and for holes by

$$v_p = \frac{v_s}{\left[1 + \left(\frac{E_{op}}{E} \right) \right]} \quad (5.27b)$$

The variables are $v_s = 10^7$ cm/s at $T = 300$ K, $E_{on} = 7 \times 10^3$ V/cm, and $E_{op} = 2 \times 10^4$ V/cm.

We may note that for small electric fields, the drift velocities reduce to

$$v_n \cong \left(\frac{E}{E_{on}} \right) \cdot v_s \quad (5.28a)$$

and

$$v_p \cong \left(\frac{E}{E_{op}} \right) \cdot v_s \quad (5.28b)$$

At low electric fields, the drift velocities are linear functions of the electric field as we have discussed. However, for large electric fields, the drift velocities approach the saturation value.

The drift velocity versus electric field characteristic of gallium arsenide is more complicated than for silicon or germanium. At low fields, the slope of the drift velocity versus E-field is constant and is the low-field electron mobility, which is approximately $8500 \text{ cm}^2/\text{V}\cdot\text{s}$ for gallium arsenide. The low-field electron mobility in gallium arsenide is much larger than in silicon. As the field increases, the electron drift velocity in gallium arsenide reaches a peak and then decreases. A differential mobility is the slope of the v_d versus E curve at a particular point on the curve and the negative slope of the drift velocity versus electric field represents a negative differential mobility. The negative differential mobility produces a negative differential resistance; this characteristic is used in the design of oscillators.

The negative differential mobility can be understood by considering the E versus k diagram for gallium arsenide, which is shown again in Figure 5.8. The density of states effective mass of the electron in the lower valley is $m_n^* = 0.067m_0$. The small effective mass leads to a large mobility. As the E-field increases, the energy of the electron increases and the electron can be scattered into the upper valley, where the density of states effective mass is $0.55m_0$. The larger effective mass in the upper valley yields a smaller mobility. This intervalley transfer mechanism results in a decreasing average drift velocity of electrons with electric field, or the negative differential mobility characteristic.

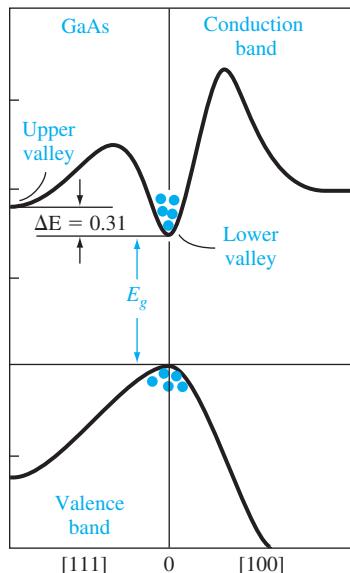


Figure 5.8 | Energy-band structure for gallium arsenide showing the upper valley and lower valley in the conduction band.
(From Sze [15].)

TEST YOUR UNDERSTANDING

TYU 5.1 Consider a sample of silicon at $T = 300\text{ K}$ doped at an impurity concentration of $N_d = 10^{15}\text{ cm}^{-3}$ and $N_a = 10^{14}\text{ cm}^{-3}$. Assume electron and hole mobilities given in Table 5.1. Calculate the drift current density if the applied electric field is $E = 35\text{ V/cm}$. (Ans. 6.08 A/cm^2)

TYU 5.2 Silicon at $T = 300\text{ K}$ is doped with impurity concentrations of $N_d = 5 \times 10^{16}\text{ cm}^{-3}$ and $N_a = 2 \times 10^{16}\text{ cm}^{-3}$. (a) What are the electron and hole mobilities? (b) Determine the conductivity and resistivity of the material.

[Ans. (a) $\mu_n = 1000\text{ cm}^2/\text{V-s}$, $\mu_p = 350\text{ cm}^2/\text{V-s}$; (b) $\sigma = 4.8\text{ (}\Omega\text{-cm)}^{-1}$, $\rho = 0.208\text{ }\Omega\text{-cm}$]

TYU 5.3 For a particular silicon semiconductor device at $T = 300\text{ K}$, the required material is n type with a resistivity of $0.10\text{ }\Omega\text{-cm}$. (a) Determine the required impurity doping concentration and (b) the resulting electron mobility. [Ans. (a) From Figure 5.4, $N_d \approx 9 \times 10^{16}\text{ cm}^{-3}$; (b) $\mu_n \approx 695\text{ cm}^2/\text{V-s}$]

5.2 | CARRIER DIFFUSION

There is a second mechanism, in addition to drift, that can induce a current in a semiconductor. We may consider a classic physics example in which a container, as shown in Figure 5.9, is divided into two parts by a membrane. The left side contains gas molecules at a particular temperature and the right side is initially empty. The gas molecules are in continual random thermal motion so that, when the membrane is broken, the gas molecules flow into the right side of the container. *Diffusion* is the process whereby particles flow from a region of **high concentration toward** a region of **low concentration**. If the gas molecules were electrically charged, the net flow of charge would result in a *diffusion current*.

5.2.1 Diffusion Current Density

To begin to understand the diffusion process in a semiconductor, we will consider a simplified analysis. Assume that an electron concentration varies in one dimension as shown in Figure 5.10. The temperature is assumed to be uniform so that the average thermal velocity of electrons is independent of x . To calculate the current, we will determine the net flow of electrons per unit time per unit area crossing the plane at $x = 0$. If the distance l shown in Figure 5.10 is less than the mean-free path of an electron, that is, the average

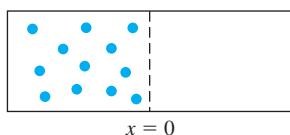


Figure 5.9 | Container divided by a membrane with gas molecules on one side.

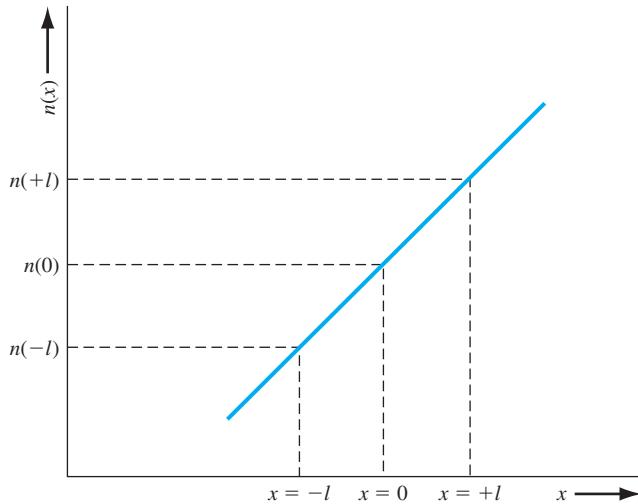


Figure 5.10 | Electron concentration versus distance.

distance an electron travels between collisions ($l < v_{th} \tau_{cn}$), then on the average, electrons moving to the right at $x = -l$ and electrons moving to the left at $x = +l$ will cross the $x = 0$ plane. One half of the electrons at $x = -l$ will be traveling to the right at any instant of time and one half of the electrons at $x = +l$ will be traveling to the left at any given time. The net rate of electron flow, F_n , in the $+x$ direction at $x = 0$ is given by

$$F_n = \frac{1}{2}n(-l)v_{th} - \frac{1}{2}n(+l)v_{th} = \frac{1}{2}v_{th}[n(-l) - n(+l)] \quad (5.29)$$

If we expand the electron concentration in a Taylor series about $x = 0$ keeping only the first two terms, then we can write Equation (5.29) as

$$F_n = \frac{1}{2}v_{th}\left\{\left[n(0) - l\frac{dn}{dx}\right] - \left[n(0) + l\frac{dn}{dx}\right]\right\} \quad (5.30)$$

which becomes

$$F_n = -v_{th}l\frac{dn}{dx} \quad (5.31)$$

Each electron has a charge ($-e$), so the current is

$$J = -eF_n = +ev_{th}l\frac{dn}{dx} \quad (5.32)$$

The current described by Equation (5.32) is the electron diffusion current and is proportional to the spatial derivative, or density gradient, of the electron concentration.

The diffusion of electrons from a region of high concentration to a region of low concentration produces a flux of electrons flowing in the negative x direction for this example. Since electrons have a negative charge, the conventional current direction is in the positive x direction. Figure 5.11a shows these one-dimensional flux and

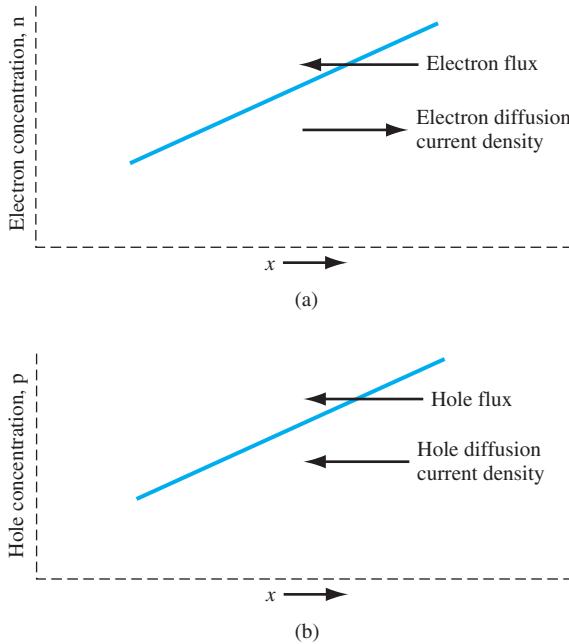


Figure 5.11 | (a) Diffusion of electrons due to a density gradient. (b) Diffusion of holes due to a density gradient.

current directions. We may write the electron diffusion current density for this one-dimensional case, in the form

$$J_{nx|dif} = eD_n \frac{dn}{dx} \quad (5.33)$$

where D_n is called the *electron diffusion coefficient*, has units of cm^2/s , and is a positive quantity. If the electron density gradient becomes negative, the electron diffusion current density will be in the negative x direction.

Figure 5.11b shows an example of a hole concentration as a function of distance in a semiconductor. The diffusion of holes, from a region of high concentration to a region of low concentration, produces a flux of holes in the negative x direction. Since holes are positively charged particles, the conventional diffusion current density is also in the negative x direction. The hole diffusion current density is proportional to the hole density gradient and to the electronic charge, so we may write

$$J_{px|dif} = -eD_p \frac{dp}{dx} \quad (5.34)$$

for the one-dimensional case. The parameter D_p is called the *hole diffusion coefficient*, has units of cm^2/s , and is a positive quantity. If the hole density gradient becomes negative, the hole diffusion current density will be in the positive x direction.

Objective: Calculate the diffusion current density given a density gradient.

EXAMPLE 5.5

Assume that, in an n-type gallium arsenide semiconductor at $T = 300$ K, the electron concentration varies linearly from 1×10^{18} to 7×10^{17} cm $^{-3}$ over a distance of 0.10 cm. Calculate the diffusion current density if the electron diffusion coefficient is $D_n = 225$ cm 2 /s.

■ Solution

The diffusion current density is given by

$$\begin{aligned} J_{n\text{dif}} &= eD_n \frac{dn}{dx} \approx eD_n \frac{\Delta n}{\Delta x} \\ &= (1.6 \times 10^{-19})(225) \left(\frac{1 \times 10^{18} - 7 \times 10^{17}}{0.10} \right) = 108 \text{ A/cm}^2 \end{aligned}$$

■ Comment

A significant diffusion current density can be generated in a semiconductor material with only a modest density gradient.

■ EXERCISE PROBLEM

Ex 5.5 The hole density in silicon is given by $p(x) = 10^{16} e^{-(x/L_p)}$ ($x \geq 0$) where $L_p = 2 \times 10^{-4}$ cm.

Assume the hole diffusion coefficient is $D_p = 8$ cm 2 /s. Determine the hole diffusion current density at (a) $x = 0$, (b) $x = 2 \times 10^{-4}$ cm, and (c) $x = 10^{-3}$ cm.

[Ans. (a) $J_p = 64$ A/cm 2 ; (b) $J_p = 23.54$ A/cm 2 ; (c) $J_p = 0.431$ A/cm 2]

5.2.2 Total Current Density

We now have four possible independent current mechanisms in a semiconductor. These components are electron drift and diffusion currents and hole drift and diffusion currents. The total current density is the sum of these four components, or, for the one-dimensional case,

$$J = en\mu_n E_x + ep\mu_p E_x + eD_n \frac{dn}{dx} - eD_p \frac{dp}{dx} \quad (5.35)$$

This equation may be generalized to three dimensions as

$$J = en\mu_n E + ep\mu_p E + eD_n \nabla n - eD_p \nabla p \quad (5.36)$$

The electron mobility gives an indication of how well an electron moves in a semiconductor as a result of the force of an electric field. The electron diffusion coefficient gives an indication of how well an electron moves in a semiconductor as a result of a density gradient. The electron mobility and diffusion coefficient are not independent parameters. Similarly, the hole mobility and diffusion coefficient are not independent parameters. The relationship between mobility and the diffusion coefficient is developed in the next section.

The expression for the total current in a semiconductor contains four terms. Fortunately in most situations, we will only need to consider one term at any one time at a particular point in a semiconductor.

TEST YOUR UNDERSTANDING

TYU 5.4 The electron concentration in silicon is given by $n(x) = 10^{15} e^{-(x/L_n)} \text{ cm}^{-3}$ ($x \geq 0$) where $L_n = 10^{-4} \text{ cm}$. The electron diffusion coefficient is $D_n = 25 \text{ cm}^2/\text{s}$. Determine the electron diffusion current density at (a) $x = 0$, (b) $x = 10^{-4} \text{ cm}$, and (c) $x \rightarrow \infty$. [Ans. (a) -40 A/cm^2 ; (b) -14.7 A/cm^2 ; (c) 0]

TYU 5.5 The hole concentration in silicon varies linearly from $x = 0$ to $x = 0.01 \text{ cm}$. The hole diffusion coefficient is $D_p = 10 \text{ cm}^2/\text{s}$, the hole diffusion current density is 20 A/cm^2 , and the hole concentration at $x = 0$ is $p = 4 \times 10^{17} \text{ cm}^{-3}$. What is the value of the hole concentration at $x = 0.01 \text{ cm}$? [Ans. $2.75 \times 10^{17} \text{ cm}^{-3}$]

5.3 | GRADED IMPURITY DISTRIBUTION

In most cases so far, we have assumed that the semiconductor is uniformly doped. In many semiconductor devices, however, there may be regions that are nonuniformly doped. We will investigate how a nonuniformly doped semiconductor reaches thermal equilibrium and, from this analysis, we will derive the Einstein relation, which relates mobility and the diffusion coefficient.

5.3.1 Induced Electric Field

Consider a semiconductor that is nonuniformly doped with donor impurity atoms. If the semiconductor is in thermal equilibrium, the Fermi energy level is constant through the crystal so the energy-band diagram may qualitatively look like that shown in Figure 5.12. The doping concentration decreases as x increases in this case. There will be a diffusion of majority carrier electrons from the region of high concentration to the region of low concentration, which is in the $+x$ direction. The flow of negative electrons leaves behind positively charged donor ions. The separation of

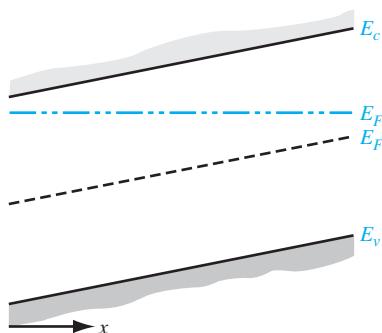


Figure 5.12 | Energy-band diagram for a semiconductor in thermal equilibrium with a nonuniform donor concentration.

positive and negative charge induces an electric field that is in a direction to oppose the diffusion process. When equilibrium is reached, the mobile carrier concentration is not exactly equal to the fixed impurity concentration and the induced electric field prevents any further separation of charge. In most cases of interest, the space charge induced by this diffusion process is a small fraction of the impurity concentration, thus the mobile carrier concentration is not too different from the impurity dopant density.

The electric potential ϕ is related to electron potential energy by the charge ($-e$), so we can write

$$\phi = +\frac{1}{e} (E_F - E_{Fi}) \quad (5.37)$$

The electric field for the one-dimensional situation is defined as

$$E_x = -\frac{d\phi}{dx} = \frac{1}{e} \frac{dE_{Fi}}{dx} \quad (5.38)$$

If the intrinsic Fermi-level changes as a function of distance through a semiconductor in thermal equilibrium, an electric field exists in the semiconductor.

If we assume a quasi-neutrality condition in which the electron concentration is almost equal to the donor impurity concentration, then we can still write

$$n_0 = n_i \exp \left[\frac{E_F - E_{Fi}}{kT} \right] \approx N_d(x) \quad (5.39)$$

Solving for $E_F - E_{Fi}$, we obtain

$$E_F - E_{Fi} = kT \ln \left(\frac{N_d(x)}{n_i} \right) \quad (5.40)$$

The **Fermi level is constant for thermal equilibrium** so when we take the derivative with respect to x we obtain

$$-\frac{dE_{Fi}}{dx} = \frac{kT}{N_d(x)} \frac{dN_d(x)}{dx} \quad (5.41)$$

The electric field can then be written, combining Equations (5.41) and (5.38), as

$$E_x = -\left(\frac{kT}{e}\right) \frac{1}{N_d(x)} \frac{dN_d(x)}{dx} \quad (5.42)$$

Since we have an electric field, there will be a potential difference through the semiconductor due to the nonuniform doping.

Objective: Determine the induced electric field in a semiconductor in thermal equilibrium, given a linear variation in doping concentration.

EXAMPLE 5.6

Assume that the donor concentration in an n-type semiconductor at $T = 300$ K is given by

$$N_d(x) = 10^{16} - 10^{19}x \quad (\text{cm}^{-3})$$

where x is given in cm and ranges between $0 \leq x \leq 1 \mu\text{m}$

■ Solution

Taking the derivative of the donor concentration, we have

$$\frac{dN_d(x)}{dx} = -10^{19} \text{ cm}^{-4}$$

The electric field is given by Equation (5.42), so we have

$$E_x = \frac{-(0.0259)(-10^{19})}{(10^{16} - 10^{19}x)}$$

At $x = 0$, for example, we find

$$E_x = 25.9 \text{ V/cm}$$

■ Comment

We may recall from our previous discussion of drift current that fairly small electric fields can produce significant drift current densities, so that an induced electric field from nonuniform doping can significantly influence semiconductor device characteristics.

■ EXERCISE PROBLEM

- Ex 5.6** Assume the donor concentration in an n-type semiconductor at $T = 300 \text{ K}$ is given by $N_d(x) = 10^{16} e^{-x/L}$ where $L = 2 \times 10^{-2} \text{ cm}$. Determine the induced electric field in the semiconductor at (a) $x = 0$ and (b) $x = 10^{-4} \text{ cm}$.

[Ans. E = 1.295 V/cm for (a) and (b)]

5.3.2 The Einstein Relation

If we consider the nonuniformly doped semiconductor represented by the energy-band diagram shown in Figure 5.12 and assume there are no electrical connections so that the semiconductor is in thermal equilibrium, then the individual electron and hole currents must be zero. We can write

$$J_n = 0 = en\mu_n E_x + eD_n \frac{dn}{dx} \quad (5.43)$$

If we assume quasi-neutrality so that $n \approx N_d(x)$, then we can rewrite Equation (5.43) as

$$J_n = 0 = e\mu_n N_d(x)E_x + eD_n \frac{dN_d(x)}{dx} \quad (5.44)$$

Substituting the expression for the electric field from Equation (5.42) into Equation (5.44), we obtain

$$0 = -e\mu_n N_d(x) \left(\frac{kT}{e}\right) \frac{1}{N_d(x)} \frac{dN_d(x)}{dx} + eD_n \frac{dN_d(x)}{dx} \quad (5.45)$$

Equation (5.45) is valid for the condition

$$\frac{D_n}{\mu_n} = \frac{kT}{e} \quad (5.46a)$$

The hole current must also be zero in the semiconductor. From this condition, we can show that

$$\frac{D_p}{\mu_p} = \frac{kT}{e} \quad (5.46b)$$

Combining Equations (5.46a) and (5.46b) gives

$$\frac{D_n}{\mu_n} = \frac{D_p}{\mu_p} = \frac{kT}{e} \quad (5.47)$$

The diffusion coefficient and mobility are not independent parameters. This relation between the mobility and diffusion coefficient, given by Equation (5.47), is known as the *Einstein relation*.

Objective: Determine the diffusion coefficient given the carrier mobility.

EXAMPLE 5.7

Assume that the mobility of a particular carrier is $1000 \text{ cm}^2/\text{V}\cdot\text{s}$ at $T = 300 \text{ K}$.

■ Solution

Using the Einstein relation, we have that

$$D = \left(\frac{kT}{e}\right)\mu = (0.0259)(1000) = 25.9 \text{ cm}^2/\text{s}$$

■ Comment

Although this example is fairly simple and straightforward, it is important to keep in mind the relative orders of magnitude of the mobility and diffusion coefficient. The diffusion coefficient is approximately 40 times smaller than the mobility at room temperature.

■ EXERCISE PROBLEM

Ex 5.7 Assume the electron diffusion coefficient of a semiconductor at $T = 300 \text{ K}$ is

$$D_n = 215 \text{ cm}^2/\text{s}. \text{ Determine the electron mobility. (s}\cdot\text{m}^2/\text{V}\cdot\text{A})$$

Table 5.2 shows the diffusion coefficient values at $T = 300 \text{ K}$ corresponding to the mobilities listed in Table 5.1 for silicon, gallium arsenide, and germanium.

The relation between the mobility and diffusion coefficient given by Equation (5.47) contains temperature. It is important to keep in mind that the major temperature effects are a result of lattice scattering and ionized impurity scattering processes, as discussed in Section 5.1.2. As the mobilities are strong functions of temperature because of the scattering processes, the diffusion coefficients are also strong functions of temperature. The specific temperature dependence given in Equation (5.47) is a small fraction of the real temperature characteristic.

Table 5.2 | Typical mobility and diffusion coefficient values at $T = 300 \text{ K}$ ($\mu = \text{cm}^2/\text{V}\cdot\text{s}$ and $D = \text{cm}^2/\text{s}$)

	μ_m	D_n	μ_p	D_p
Silicon	1350	35	480	12.4
Gallium arsenide	8500	220	400	10.4
Germanium	3900	101	1900	49.2

*5.4 | THE HALL EFFECT

The Hall effect is a consequence of the forces that are exerted on moving charges by electric and magnetic fields. The Hall effect is used to distinguish whether a semiconductor is n type or p type³ and to measure the majority carrier concentration and majority carrier mobility. The Hall effect device, as discussed in this section, is used to experimentally measure semiconductor parameters. However, it is also used extensively in engineering applications as a magnetic probe and in other circuit applications.

The force on a particle having a charge q and moving in a magnetic field is given by

$$F = qv \times B \quad (5.48)$$

where the cross product is taken between velocity and magnetic field so that the force vector is perpendicular to both the velocity and magnetic field.

Figure 5.13 illustrates the Hall effect. A semiconductor with a current I_x is placed in a magnetic field perpendicular to the current. In this case, the magnetic field is in the z direction. Electrons and holes flowing in the semiconductor will experience a force as indicated in the figure. The force on both electrons and holes is in the ($-y$) direction. In a p-type semiconductor ($p_0 > n_0$), there will be a buildup of positive charge on the $y = 0$ surface of the semiconductor and, in an n-type semiconductor ($n_0 > p_0$), there will be a buildup of negative charge on the $y = 0$ surface. This net charge induces an electric field in the y direction as shown in the figure. In steady

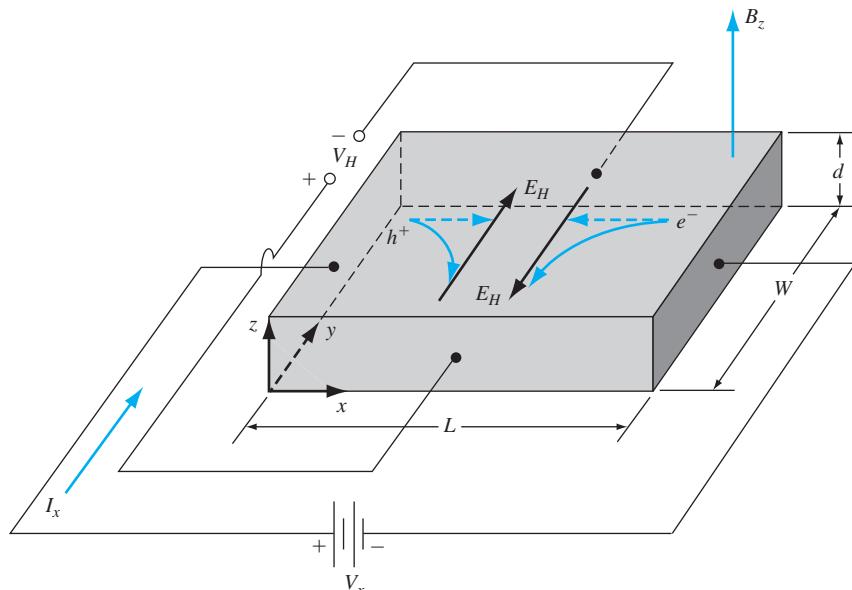


Figure 5.13 | Geometry for measuring the Hall effect.

*Indicates sections that will aid in the total summation of understanding of semiconductor devices, but may be skipped the first time through the text without loss of continuity.

³We will assume an extrinsic semiconductor material in which the majority carrier concentration is much larger than the minority carrier concentration.

state, the magnetic field force will be exactly balanced by the induced electric field force. This balance may be written as

$$F = q[\mathbf{E} + \mathbf{v} \times \mathbf{B}] = 0 \quad (5.49a)$$

which becomes

$$qE_y = qv_x B_z \quad (5.49b)$$

The induced electric field in the y direction is called the *Hall field*. The Hall field produces a voltage across the semiconductor which is called the *Hall voltage*. We can write

$$V_H = +E_H W \quad (5.50)$$

where E_H is assumed positive in the $+y$ direction and V_H is positive with the polarity shown.

In a p-type semiconductor, in which holes are the majority carrier, the Hall voltage will be positive as defined in Figure 5.13. In an n-type semiconductor, in which electrons are the majority carrier, the Hall voltage will have the opposite polarity. The polarity of the Hall voltage is used to determine whether an extrinsic semiconductor is n type or p type.

Substituting Equation (5.50) into Equation (5.49) gives

$$V_H = v_x WB_z \quad (5.51)$$

For a p-type semiconductor, the drift velocity of holes can be written as

$$v_{dx} = \frac{J_x}{ep} = \frac{I_x}{(ep)(Wd)} \quad (5.52)$$

where e is the magnitude of the electronic charge. Combining Equations (5.52) and (5.50), we have

$$V_H = \frac{I_x B_z}{epd} \quad (5.53)$$

or, solving for the hole concentration, we obtain

$$p = \frac{I_x B_z}{edV_H} \quad (5.54)$$

The majority carrier hole concentration is determined from the current, magnetic field, and Hall voltage.

For an n-type semiconductor, the Hall voltage is given by

$$V_H = -\frac{I_x B_z}{ned} \quad (5.55)$$

so that the electron concentration is

$$n = -\frac{I_x B_z}{edV_H} \quad (5.56)$$

Note that the Hall voltage is negative for the n-type semiconductor; therefore, the electron concentration determined from Equation (5.56) is actually a positive quantity.

Once the majority carrier concentration has been determined, we can calculate the low-field majority carrier mobility. For a p-type semiconductor, we can write

$$J_x = ep\mu_p E_x \quad (5.57)$$

The current density and electric field can be converted to current and voltage so that Equation (5.57) becomes

$$\frac{I_x}{Wd} = \frac{ep\mu_p V_x}{L} \quad (5.58)$$

The hole mobility is then given by

$$\mu_p = \frac{I_x L}{epV_x Wd} \quad (5.59)$$

Similarly for an n-type semiconductor, the low-field electron mobility is determined from

$$\mu_n = \frac{I_x L}{enV_x Wd} \quad (5.60)$$

EXAMPLE 5.8

Objective: Determine the majority carrier concentration and mobility, given Hall effect parameters.

Consider the geometry shown in Figure 5.13. Let $L = 10^{-1}$ cm, $W = 10^{-2}$ cm, and $d = 10^{-3}$ cm. Also assume that $I_x = 1.0$ mA, $V_x = 12.5$ V, $B_z = 500$ gauss = 5×10^{-2} tesla, and $V_H = -6.25$ mV.

■ Solution

A negative Hall voltage for this geometry implies that we have an n-type semiconductor. Using Equation (5.56), we can calculate the electron concentration as

$$n = \frac{-(10^{-3})(5 \times 10^{-2})}{(1.6 \times 10^{-19})(10^{-5})(-6.25 \times 10^{-3})} = 5 \times 10^{21} \text{ m}^{-3} = 5 \times 10^{15} \text{ cm}^{-3}$$

The electron mobility is then determined from Equation (5.60) as

$$\mu_n = \frac{(10^{-3})(10^{-3})}{(1.6 \times 10^{-19})(5 \times 10^{21})(12.5)(10^{-4})(10^{-5})} = 0.10 \text{ m}^2/\text{V-s}$$

or

$$\mu_n = 1000 \text{ cm}^2/\text{V-s}$$

■ Comment

It is important to note that the MKS units must be used consistently in the Hall effect equations to yield correct results.

■ EXERCISE PROBLEM

Ex 5.8 A p-type silicon sample with the geometry shown in Figure 5.13 has parameters $L = 0.2$ cm, $W = 10^{-2}$ cm, and $d = 8 \times 10^{-4}$ cm. The semiconductor parameters are $p = 10^{16} \text{ cm}^{-3}$ and $\mu_p = 320 \text{ cm}^2/\text{V-s}$. For $V_x = 10$ V and $B_z = 500$ gauss = 5×10^{-2} tesla, determine I_x and V_H . (Ans. $I_x = 0.2048$ mA, $V_H = 0.080$ mV)

5.5 | SUMMARY

- The two basic transport mechanisms are drift, due to an applied electric field, and diffusion, due to a density gradient.
- Carriers reach an average drift velocity in the presence of an applied electric field, due to scattering events. Two scattering processes within a semiconductor are lattice scattering and impurity scattering.
- The average drift velocity is a linear function of the applied electric field for small values of electric field, but the drift velocity reaches a saturation limit that is on the order of 10^7 cm/s at high electric fields.
- Carrier mobility is the ratio of the average drift velocity and applied electric field. The electron and hole mobilities are functions of temperature and of the ionized impurity concentration.
- The drift current density is the product of conductivity and electric field (a form of Ohm's law). Conductivity is a function of the carrier concentrations and mobilities. Resistivity is the inverse of conductivity.
- The diffusion current density is proportional to the carrier diffusion coefficient and the carrier density gradient.
- The diffusion coefficient and mobility are related through the Einstein relation.
- The Hall effect is a consequence of a charged carrier moving in the presence of perpendicular electric and magnetic fields. The charged carrier is deflected, inducing a Hall voltage. The polarity of the Hall voltage is a function of the semiconductor conductivity type. The majority carrier concentration and mobility can be determined from the Hall voltage.

GLOSSARY OF IMPORTANT TERMS

conductivity A material parameter related to carrier drift; quantitatively, the ratio of drift current density to electric field.

diffusion The process whereby particles flow from a region of high concentration to a region of low concentration.

diffusion coefficient The parameter relating particle flux to the particle density gradient.

diffusion current The current that results from the diffusion of charged particles.

drift The process whereby charged particles move while under the influence of an electric field.

drift current The current that results from the drift of charged particles.

drift velocity The average velocity of charged particles in the presence of an electric field.

Einstein relation The relation between the mobility and the diffusion coefficient.

Hall voltage The voltage induced across a semiconductor in a Hall effect measurement.

ionized impurity scattering The interaction between a charged carrier and an ionized impurity center.

lattice scattering The interaction between a charged carrier and a thermally vibrating lattice atom.

mobility The parameter relating carrier drift velocity and electric field.

resistivity The reciprocal of conductivity; a material parameter that is a measure of the resistance to current.

velocity saturation The saturation of carrier drift velocity with increasing electric field.

CHECKPOINT

After studying this chapter, the reader should have the ability to:

- Discuss carrier drift current density.
- Explain why carriers reach an average drift velocity in the presence of an applied electric field.
- Discuss the mechanisms of lattice scattering and impurity scattering.
- Define mobility and discuss the temperature and ionized impurity concentration dependence on mobility.
- Define conductivity and resistivity.
- Discuss velocity saturation.
- Discuss carrier diffusion current density.
- State the Einstein relation.
- Describe the Hall effect.

REVIEW QUESTIONS

1. Write the equation for the total drift current density. Is the linear relationship between drift current density and electric field always valid? Why or why not?
2. Define electron and hole mobility. What is the unit of mobility?
3. Explain the temperature dependence of mobility. Why is the carrier mobility a function of the ionized impurity concentrations?
4. Define conductivity. Define resistivity. What are the units of conductivity and resistivity?
5. Sketch the drift velocity of electrons in silicon versus electric field. Repeat for GaAs.
6. Write the equations for the diffusion current densities of electrons and holes.
7. What is the Einstein relation?
8. What is the direction of the induced electric field in a semiconductor with a graded donor impurity concentration? Repeat for a graded acceptor impurity concentration.
9. Describe the Hall effect.
10. Explain why the polarity of the Hall voltage changes depending on the conductivity type (n type or p type) of the semiconductor.

PROBLEMS

(Note: Use the semiconductor parameters given in Appendix B if the parameters are not specifically given in a problem.)

Section 5.1 Carrier Drift

- 5.1** The concentration of donor impurity atoms in silicon is $N_d = 10^{15} \text{ cm}^{-3}$. Assume an electron mobility of $\mu_n = 1300 \text{ cm}^2/\text{V}\cdot\text{s}$ and a hole mobility of $\mu_p = 450 \text{ cm}^2/\text{V}\cdot\text{s}$.
- (a) Calculate the resistivity of the material. (b) What is the conductivity of the material?
- 5.2** A p-type silicon material is to have a conductivity of $\sigma = 1.80 (\Omega\cdot\text{cm})^{-1}$. If the mobility values are $\mu_n = 1250 \text{ cm}^2/\text{V}\cdot\text{s}$ and $\mu_p = 380 \text{ cm}^2/\text{V}\cdot\text{s}$, what must be the acceptor impurity concentration in the material?

- 5.3** (a) The required conductivity of an n-type silicon sample at $T = 300$ K is to be $\sigma = 10(\Omega\text{-cm})^{-1}$. What donor impurity concentration is required? What is the electron mobility corresponding to this impurity concentration? (b) A p-type silicon material is required to have a resistivity of $\rho = 0.20(\Omega\text{-cm})$. What acceptor impurity concentration is required and what is the corresponding hole mobility?
- 5.4** (a) The resistivity of a p-type GaAs material at $T = 300$ K is required to be $\rho = 0.35(\Omega\text{-cm})$. Determine the acceptor impurity concentration that is required. What is the hole mobility corresponding to this impurity concentration? (b) An n-type GaAs material is required to have a conductivity of $\sigma = 120(\Omega\text{-cm})^{-1}$. What donor impurity concentration is required and what is the corresponding electron mobility?
- 5.5** A silicon sample is 2.5 cm long and has a cross-sectional area of 0.1 cm^2 . The silicon is n type with a donor impurity concentration of $N_d = 2 \times 10^{15}\text{ cm}^{-3}$. The resistance of the sample is measured and found to be 70Ω . What is the electron mobility?
- 5.6** Consider a homogeneous gallium arsenide semiconductor at $T = 300$ K with $N_d = 10^{16}\text{ cm}^{-3}$ and $N_a = 0$. (a) Calculate the thermal-equilibrium values of electron and hole concentrations. (b) For an applied E-field of 10 V/cm, calculate the drift current density. (c) Repeat parts (a) and (b) if $N_d = 0$ and $N_a = 10^{16}\text{ cm}^{-3}$.
- 5.7** A silicon crystal having a cross-sectional area of 0.001 cm^2 and a length of 10^{-3} cm is connected at its ends to a 10-V battery. At $T = 300$ K, we want a current of 100 mA in the silicon. Calculate (a) the required resistance R , (b) the required conductivity, (c) the density of donor atoms to be added to achieve this conductivity, and (d) the concentration of acceptor atoms to be added to form a compensated p-type material with the conductivity given from part (b) if the initial concentration of donor atoms is $N_d = 10^{15}\text{ cm}^{-3}$.
- 5.8** (a) A silicon semiconductor resistor is in the shape of a rectangular bar with a cross-sectional area of $8.5 \times 10^{-4}\text{ cm}^2$, a length of 0.075 cm, and is doped with a concentration of $2 \times 10^{16}\text{ cm}^{-3}$ boron atoms. Let $T = 300$ K. A bias of 2 volts is applied across the length of the silicon device. Calculate the current in the resistor. (b) Repeat part (a) if the length is increased by a factor of three. (c) Determine the average drift velocity of holes in parts (a) and (b).
- 5.9** (a) A GaAs semiconductor resistor is doped with donor impurities at a concentration of $N_d = 2 \times 10^{15}\text{ cm}^{-3}$ and has a cross-sectional area of $5 \times 10^{-5}\text{ cm}^2$. A current of $I = 25\text{ mA}$ is induced in the resistor with an applied bias of 5 V. Determine the length of the resistor. (b) Using the results of part (a), calculate the drift velocity of the electrons. (c) If the bias applied to the resistor in part (a) increases to 20 V, determine the resulting current if the electrons are traveling at their saturation velocity of $5 \times 10^6\text{ cm/s}$.
- 5.10** (a) Three volts is applied across a 1-cm-long semiconductor bar. The average electron drift velocity is 10^4 cm/s . Find the electron mobility. (b) If the electron mobility in part (a) were $800\text{ cm}^2/\text{V}\cdot\text{s}$, what is the average electron drift velocity?
- 5.11** Use the velocity–field relations for silicon and gallium arsenide shown in Figure 5.7 to determine the transit time of electrons through a $1\text{-}\mu\text{m}$ distance in these materials for an electric field of (a) 1 kV/cm and (b) 50 kV/cm.
- 5.12** A perfectly compensated semiconductor is one in which the donor and acceptor impurity concentrations are exactly equal. Assuming complete ionization, determine the resistivity of silicon at $T = 300$ K in which the impurity concentrations are (a) $N_a = N_d = 10^{14}\text{ cm}^{-3}$, (b) $N_a = N_d = 10^{16}\text{ cm}^{-3}$, and (c) $N_a = N_d = 10^{18}\text{ cm}^{-3}$.

- 5.13** (a) In a p-type gallium arsenide semiconductor, the conductivity is $\sigma = 5 \text{ } (\Omega\text{-cm})^{-1}$ at $T = 300 \text{ K}$. Calculate the thermal-equilibrium values of the electron and hole concentrations. (b) Repeat part (a) for n-type silicon if the resistivity is $\rho = 8 \text{ } \Omega\text{-cm}$.
- 5.14** In a particular semiconductor material, $\mu_n = 1000 \text{ cm}^2/\text{V}\cdot\text{s}$, $\mu_p = 600 \text{ cm}^2/\text{V}\cdot\text{s}$, and $N_c = N_v = 10^{19} \text{ cm}^{-3}$. These parameters are independent of temperature. The measured conductivity of the intrinsic material is $\sigma = 10^{-6} \text{ } (\Omega\text{-cm})^{-1}$ at $T = 300 \text{ K}$. Find the conductivity at $T = 500 \text{ K}$.
- 5.15** (a) Calculate the resistivity at $T = 300 \text{ K}$ of intrinsic (i) silicon, (ii) germanium, and (iii) gallium arsenide. (b) If rectangular semiconductor bars are fabricated using the materials in part (a), determine the resistance of each bar if its cross-sectional area is $85 \text{ } \mu\text{m}^2$ and length is $200 \text{ } \mu\text{m}$.
- 5.16** An n-type silicon material at $T = 300 \text{ K}$ has a conductivity of $0.25 \text{ } (\Omega\text{-cm})^{-1}$. (a) What is the donor impurity concentration and the corresponding electron mobility? (b) Determine the expected conductivity of the material at (i) $T = 250 \text{ K}$ and (ii) $T = 400 \text{ K}$.
- 5.17** The conductivity of a semiconductor layer varies with depth as $\sigma(x) = \sigma_o \exp(-x/d)$ where $\sigma_o = 20 \text{ } (\Omega\text{-cm})^{-1}$ and $d = 0.3 \text{ } \mu\text{m}$. If the thickness of the semiconductor layer is $t = 1.5 \text{ } \mu\text{m}$, determine the average conductivity of this layer.
- 5.18** An n-type silicon resistor has a length $L = 150 \text{ } \mu\text{m}$, width $W = 7.5 \text{ } \mu\text{m}$, and thickness $T = 1 \text{ } \mu\text{m}$. A voltage of 2 V is applied across the length of the resistor. The donor impurity concentration varies linearly through the thickness of the resistor with $N_d = 2 \times 10^{16} \text{ cm}^{-3}$ at the top surface and $N_d = 2 \times 10^{15} \text{ cm}^{-3}$ at the bottom surface. Assume an average carrier mobility of $\mu_n = 750 \text{ cm}^2/\text{V}\cdot\text{s}$. (a) What is the electric field in the resistor? (b) Determine the average conductivity of the silicon. (c) Calculate the current in the resistor. (d) Determine the current density near the top surface and the current density near the bottom surface.
- 5.19** Consider silicon doped at impurity concentrations of $N_d = 2 \times 10^{16} \text{ cm}^{-3}$ and $N_a = 0$. An empirical expression relating electron drift velocity to electric field is given by

$$v_d = \frac{\mu_{n0}E}{\sqrt{1 + \left(\frac{\mu_{n0}E}{v_{sat}}\right)^2}}$$

where $\mu_{n0} = 1350 \text{ cm}^2/\text{V}\cdot\text{s}$, $v_{sat} = 1.8 \times 10^7 \text{ cm/s}$, and E is given in V/cm . Plot electron drift current density (magnitude) versus electric field (log–log scale) over the range $0 \leq E \leq 10^6 \text{ V/cm}$.

- 5.20** Consider silicon at $T = 300 \text{ K}$. Assume the electron mobility is $\mu_n = 1350 \text{ cm}^2/\text{V}\cdot\text{s}$. The kinetic energy of an electron in the conduction band is $(1/2)m_n^*v_d^2$, where m_n^* is the effective mass and v_d is the drift velocity. Determine the kinetic energy of an electron in the conduction band if the applied electric field is (a) 10 V/cm and (b) 1 kV/cm .
- 5.21** Consider a semiconductor that is uniformly doped with $N_d = 10^{14} \text{ cm}^{-3}$ and $N_a = 0$, with an applied electric field of $E = 100 \text{ V/cm}$. Assume that $\mu_n = 1000 \text{ cm}^2/\text{V}\cdot\text{s}$ and $\mu_p = 0$. Also assume the following parameters:

$$N_c = 2 \times 10^{19} (T/300)^{3/2} \text{ cm}^{-3}$$

$$N_v = 1 \times 10^{19} (T/300)^{3/2} \text{ cm}^{-3}$$

$$E_g = 1.10 \text{ eV}$$

(a) Calculate the electric-current density at $T = 300 \text{ K}$. (b) At what temperature will this current increase by 5 percent? (Assume the mobilities are independent of temperature.)

- 5.22** A semiconductor material has electron and hole mobilities μ_n and μ_p , respectively. When the conductivity is considered as a function of the hole concentration p_0 , (a) show that the minimum value of conductivity, σ_{\min} , can be written as

$$\sigma_{\min} = \frac{2\sigma_i(\mu_n\mu_p)^{1/2}}{(\mu_n + \mu_p)}$$

where σ_i is the intrinsic conductivity, and (b) show that the corresponding hole concentration is $p_0 = n_i(\mu_n/\mu_p)^{1/2}$.

- 5.23** Consider three samples of silicon at $T = 300$ K. The n-type sample is doped with arsenic atoms to a concentration of $N_d = 5 \times 10^{16} \text{ cm}^{-3}$. The p-type sample is doped with boron atoms to a concentration of $N_a = 2 \times 10^{16} \text{ cm}^{-3}$. The compensated sample is doped with both the donors and acceptors described in the n-type and p-type samples. (a) Find the equilibrium electron and hole concentrations in each sample, (b) determine the majority carrier mobility in each sample, (c) calculate the conductivity of each sample, (d) and determine the electric field required in each sample to induce a drift current density of $J = 120 \text{ A/cm}^2$.
- 5.24** Three scattering mechanisms are present in a particular semiconductor material. If only the first scattering mechanism were present, the mobility would be $\mu_1 = 2000 \text{ cm}^2/\text{V-s}$, if only the second mechanism were present, the mobility would be $\mu_2 = 1500 \text{ cm}^2/\text{V-s}$, and if only the third mechanism were present, the mobility would be $\mu_3 = 500 \text{ cm}^2/\text{V-s}$. What is the net mobility?
- 5.25** Assume that the mobility of electrons in silicon at $T = 300$ K is $\mu_n = 1300 \text{ cm}^2/\text{V-s}$. Also assume that the mobility is limited by lattice scattering and varies as $T^{-3/2}$. Determine the electron mobility at (a) $T = 200$ K and (b) $T = 400$ K.
- 5.26** Two scattering mechanisms exist in a semiconductor. If only the first mechanism were present, the mobility would be $250 \text{ cm}^2/\text{V-s}$. If only the second mechanism were present, the mobility would be $500 \text{ cm}^2/\text{V-s}$. Determine the mobility when both scattering mechanisms exist at the same time.
- 5.27** The effective density of states functions in silicon can be written in the form

$$N_c = 2.8 \times 10^{19} \left(\frac{T}{300} \right)^{3/2} \quad N_v = 1.04 \times 10^{19} \left(\frac{T}{300} \right)^{3/2}$$

Assume the mobilities are given by

$$\mu_n = 1350 \left(\frac{T}{300} \right)^{-3/2} \quad \mu_p = 480 \left(\frac{T}{300} \right)^{-3/2}$$

Assume the bandgap energy is $E_g = 1.12 \text{ eV}$ and independent of temperature. Plot the intrinsic conductivity as a function of T over the range $200 \leq T \leq 600$ K.

- 5.28** (a) Assume that the electron mobility in an n-type semiconductor is given by

$$\mu_n = \frac{1350}{\left(1 + \frac{N_d}{5 \times 10^{16}} \right)^{1/2}} \text{ cm}^2/\text{V-s}$$

where N_d is the donor concentration in cm^{-3} . Assuming complete ionization, plot the conductivity as a function of N_d over the range $10^{15} \leq N_d \leq 10^{18} \text{ cm}^{-3}$. (b) Compare the results of part (a) to that if the mobility were assumed to be a constant equal to $1350 \text{ cm}^2/\text{V-s}$. (c) If an electric field of $E = 10 \text{ V/cm}$ is applied to the semiconductor, plot the electron drift current density of parts (a) and (b).

Section 5.2 Carrier Diffusion

- 5.29** Consider a sample of silicon at $T = 300$ K. Assume that the electron concentration varies linearly with distance, as shown in Figure P5.29. The diffusion current density is found to be $J_n = 0.19$ A/cm². If the electron diffusion coefficient is $D_n = 25$ cm²/s, determine the electron concentration at $x = 0$.
- 5.30** The steady-state electron distribution in silicon can be approximated by a linear function of x . The maximum electron concentration occurs at $x = 0$ and is $n(0) = 2 \times 10^{16}$ cm⁻³. At $x = 0.012$ cm, the electron concentration is 5×10^{15} cm⁻³. If the electron diffusion coefficient is $D_n = 27$ cm²/s, determine the electron diffusion current density.
- 5.31** The electron diffusion current density in a semiconductor is a constant and is given by $J_n = -2$ A/cm². The electron concentration at $x = 0$ is $n(0) = 10^{15}$ cm⁻³. (a) Calculate the electron concentration at $x = 20$ μm if the material is silicon with $D_n = 30$ cm²/s. (b) Repeat part (a) if the material is GaAs with $D_n = 230$ cm²/s.
- 5.32** The hole concentration in p-type GaAs is given by $p(x) = 10^{16}(1 + x/L)^2$ cm⁻³ for $-L \leq x \leq 0$ where $L = 12$ μm . The hole diffusion coefficient is $D_p = 10$ cm²/s. Calculate the hole diffusion current density at (a) $x = 0$, (b) $x = -6$ μm , and (c) $x = -12$ μm .
- 5.33** In silicon, the electron concentration is given by $n(x) = 10^{15}e^{-x/L_n}$ cm⁻³ for $x \geq 0$ and the hole concentration is given by $p(x) = 5 \times 10^{15}e^{+x/L_p}$ cm⁻³ for $x \leq 0$. The parameter values are $L_n = 2 \times 10^{-3}$ cm and $L_p = 5 \times 10^{-4}$ cm. The electron and hole diffusion coefficients are $D_n = 25$ cm²/s and $D_p = 10$ cm²/s, respectively. The total current density is defined as the sum of the electron and hole diffusion current densities at $x = 0$. Calculate the total current density.
- 5.34** The concentration of holes in a semiconductor is given by $p(x) = 5 \times 10^{15}e^{-x/L_p}$ cm⁻³ for $x \geq 0$. Determine the hole diffusion current density at (a) $x = 0$ and (b) $x = L_p$ if the material is (i) silicon with $D_p = 10$ cm²/s and $L_p = 50$ μm , and (ii) germanium with $D_p = 48$ cm²/s and $L_p = 22.5$ μm .
- 5.35** The electron concentration in silicon at $T = 300$ K is given by

$$n(x) = 10^{16} \exp\left(\frac{-x}{18}\right) \text{ cm}^{-3}$$

where x is measured in μm and is limited to $0 \leq x \leq 25$ μm . The electron diffusion coefficient is $D_n = 25$ cm²/s and the electron mobility is $\mu_n = 960$ cm²/V-s. The total electron current density through the semiconductor is constant and equal to

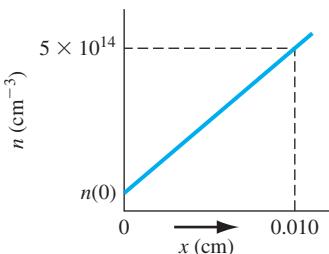


Figure P5.29 | Figure for Problem 5.29.

$J_n = -40 \text{ A/cm}^2$. The electron current has both diffusion and drift current components. Determine the electric field as a function of x which must exist in the semiconductor.

- 5.36** The total current in a semiconductor is constant and equal to $J = -10 \text{ A/cm}^2$. The total current is composed of a hole drift current and electron diffusion current. Assume that the hole concentration is a constant and equal to 10^{16} cm^{-3} and assume that the electron concentration is given by $n(x) = 2 \times 10^{15} e^{-x/L} \text{ cm}^{-3}$ where $L = 15 \mu\text{m}$. The electron diffusion coefficient is $D_n = 27 \text{ cm}^2/\text{s}$ and the hole mobility is $\mu_p = 420 \text{ cm}^2/\text{V}\cdot\text{s}$. Calculate (a) the electron diffusion current density for $x > 0$, (b) the hole drift current density for $x > 0$, and (c) the required electric field for $x > 0$.
- *5.37** A constant electric field, $E = 12 \text{ V/cm}$, exists in the $+x$ direction of an n-type gallium arsenide semiconductor for $0 \leq x \leq 50 \mu\text{m}$. The total current density is a constant and is $J = 100 \text{ A/cm}^2$. At $x = 0$, the drift and diffusion currents are equal. Let $T = 300 \text{ K}$ and $\mu_n = 8000 \text{ cm}^2/\text{V}\cdot\text{s}$. (a) Determine the expression for the electron concentration $n(x)$. (b) Calculate the electron concentration at $x = 0$ and at $x = 50 \mu\text{m}$. (c) Calculate the drift and diffusion current densities at $x = 50 \mu\text{m}$.
- *5.38** In n-type silicon, the Fermi energy level varies linearly with distance over a short range. At $x = 0$, $E_F - E_{Fi} = 0.4 \text{ eV}$ and, at $x = 10^{-3} \text{ cm}$, $E_F - E_{Fi} = 0.15 \text{ eV}$. (a) Write the expression for the electron concentration over the distance. (b) If the electron diffusion coefficient is $D_n = 25 \text{ cm}^2/\text{s}$, calculate the electron diffusion current density at (i) $x = 0$ and (ii) $x = 5 \times 10^{-4} \text{ cm}$.
- *5.39** (a) The electron concentration in a semiconductor is given by $n = 10^{16}(1 - x/L) \text{ cm}^{-3}$ for $0 \leq x \leq L$, where $L = 10 \mu\text{m}$. The electron mobility and diffusion coefficient are $\mu_n = 1000 \text{ cm}^2/\text{V}\cdot\text{s}$ and $D_n = 25.9 \text{ cm}^2/\text{s}$. An electric field is applied such that the total electron current density is a constant over the given range of x and is $J_n = -80 \text{ A/cm}^2$. Determine the required electric field versus distance function. (b) Repeat part (a) if $J_n = -20 \text{ A/cm}^2$.

Section 5.3 Graded Impurity Distribution

- 5.40** Consider an n-type semiconductor at $T = 300 \text{ K}$ in thermal equilibrium (no current). Assume that the donor concentration varies as $N_d(x) = N_{d0}e^{-x/L}$ over the range $0 \leq x \leq L$ where $N_{d0} = 10^{16} \text{ cm}^{-3}$ and $L = 10 \mu\text{m}$. (a) Determine the electric field as a function of x for $0 \leq x \leq L$. (b) Calculate the potential difference between $x = 0$ and $x = L$ (with the potential at $x = 0$ being positive with respect to that at $x = L$).
- 5.41** Using the data in Example 5.6, calculate the potential difference between $x = 0$ and $x = 1 \mu\text{m}$.
- 5.42** Determine the doping profile in a semiconductor at $T = 300 \text{ K}$ that will induce a constant electric field of 500 V/cm over a length of 0.1 cm .
- *5.43** In GaAs, the donor impurity concentration varies as $N_{d0} \exp(-x/L)$ for $0 \leq x \leq L$, where $L = 0.1 \mu\text{m}$ and $N_{d0} = 5 \times 10^{16} \text{ cm}^{-3}$. Assume $\mu_n = 6000 \text{ cm}^2/\text{V}\cdot\text{s}$ and $T = 300 \text{ K}$. (a) Derive the expression for the electron diffusion current density versus distance over the given range of x . (b) Determine the induced electric field that generates a drift current density that compensates the diffusion current density.
- 5.44** (a) Consider the electron mobility in silicon for $N_d = 10^{17} \text{ cm}^{-3}$ from Figure 5.2a. Calculate and plot the electron diffusion coefficient versus temperature over the range $-50 \leq T \leq 200^\circ\text{C}$. (b) Repeat part (a) if the electron diffusion coefficient is given by $D_n = (0.0259)\mu_n$ for all temperatures. What conclusion can be made about the temperature dependence of the diffusion coefficient?

*Asterisks next to problems indicate problems that are more difficult.

- 5.45** Consider a semiconductor at $T = 300$ K. (a) (i) Determine the electron diffusion coefficient if the electron mobility is $\mu_n = 1150 \text{ cm}^2/\text{V}\cdot\text{s}$. (ii) Repeat (i) of part (a) if the electron mobility is $\mu_n = 6200 \text{ cm}^2/\text{V}\cdot\text{s}$. (b) (i) Determine the hole mobility if the hole diffusion coefficient is $D_p = 8 \text{ cm}^2/\text{s}$. (ii) Repeat (i) of part (b) if the hole diffusion coefficient is $D_p = 35 \text{ cm}^2/\text{s}$.

Section 5.4 The Hall Effect

(Note: Refer to Figure 5.13 for the geometry of the Hall effect.)

- 5.46** Silicon, at $T = 300$ K, is uniformly doped with phosphorus atoms at a concentration of $2 \times 10^{16} \text{ cm}^{-3}$. A Hall device has the same geometrical dimensions as given in Example 5.8. The current is $I_x = 1.2 \text{ mA}$ and the magnetic field is $B_z = 500 \text{ gauss} = 5 \times 10^{-2} \text{ tesla}$. Determine (a) the Hall voltage and (b) the Hall field.
- 5.47** Germanium is doped with 5×10^{15} donor atoms per cm^3 at $T = 300$ K. The dimensions of the Hall device are $d = 5 \times 10^{-3} \text{ cm}$, $W = 2 \times 10^{-2} \text{ cm}$, and $L = 10^{-1} \text{ cm}$. The current is $I_x = 250 \mu\text{A}$, the applied voltage is $V_x = 100 \text{ mV}$, and the magnetic flux density is $B_z = 500 \text{ gauss} = 5 \times 10^{-2} \text{ tesla}$. Calculate: (a) the Hall voltage, (b) the Hall field, and (c) the carrier mobility.
- 5.48** A semiconductor Hall device at $T = 300$ K has the following geometry: $d = 10^{-3} \text{ cm}$, $W = 10^{-2} \text{ cm}$, and $L = 10^{-1} \text{ cm}$. The following parameters are measured: $I_x = 0.50 \text{ mA}$, $V_x = 15 \text{ V}$, $V_H = -5.2 \text{ mV}$, and $B_z = 0.10 \text{ tesla}$. Determine the (a) conductivity type, (b) majority carrier concentration, and (c) majority carrier mobility.
- 5.49** Consider silicon at $T = 300$ K. A Hall effect device is fabricated with the following geometry: $d = 5 \times 10^{-3} \text{ cm}$, $W = 5 \times 10^{-2} \text{ cm}$, and $L = 0.50 \text{ cm}$. The electrical parameters measured are: $I_x = 0.50 \text{ mA}$, $V_x = 1.25 \text{ V}$, and $B_z = 650 \text{ gauss} = 6.5 \times 10^{-2} \text{ tesla}$. The Hall field is $E_H = -16.5 \text{ mV/cm}$. Determine (a) the Hall voltage, (b) the conductivity type, (c) the majority carrier concentration, and (d) the majority carrier mobility.
- 5.50** Consider a gallium arsenide sample at $T = 300$ K. A Hall effect device has been fabricated with the following geometry: $d = 0.01 \text{ cm}$, $W = 0.05 \text{ cm}$, and $L = 0.5 \text{ cm}$. The electrical parameters are: $I_x = 2.5 \text{ mA}$, $V_x = 2.2 \text{ V}$, and $B_z = 2.5 \times 10^{-2} \text{ tesla}$. The Hall voltage is $V_H = -4.5 \text{ mV}$. Find: (a) the conductivity type, (b) the majority carrier concentration, (c) the mobility, and (d) the resistivity.

Summary and Review

- 5.51** An n-type silicon semiconductor resistor is to be designed so that it carries a current of 5 mA with an applied voltage of 5 V. (a) If $N_d = 3 \times 10^{14} \text{ cm}^{-3}$ and $N_a = 0$, design a resistor to meet the required specifications. (b) If $N_d = 3 \times 10^{16} \text{ cm}^{-3}$ and $N_a = 2.5 \times 10^{16} \text{ cm}^{-3}$, redesign the resistor. (c) Discuss the relative lengths of the two designs compared to the doping concentration. Is there a linear relationship?
- 5.52** In fabricating a Hall effect device, the two points at which the Hall voltage is measured may not be lined up exactly perpendicular to the current I_x (see Figure 5.13). Discuss the effect this misalignment will have on the Hall voltage. Show that a valid Hall voltage can be obtained from two measurements: first with the magnetic field in the $+z$ direction, and then in the $-z$ direction.
- 5.53** Another technique for determining the conductivity type of a semiconductor is called the hot probe method. It consists of two probes and an ammeter that indicates the

direction of current. One probe is heated and the other is at room temperature. No voltage is applied, but a current will exist when the probes touch the semiconductor. Explain the operation of this hot probe technique and sketch a diagram indicating the direction of current for p- and n-type semiconductor samples.

READING LIST

- *1. Bube, R. H. *Electrons in Solids: An Introductory Survey*, 3rd ed. San Diego, CA: Academic Press, 1992.
- 2. Caughey, D. M., and R. E. Thomas. "Carrier Mobilities in Silicon Empirically, Related to Doping and Field." *Proc. IEEE* 55 (1967), p. 2192.
- 3. Dimitrijev, S. *Principles of Semiconductor Devices*. New York: Oxford University, 2006.
- 4. Kano, K. *Semiconductor Devices*. Upper Saddle River, NJ: Prentice Hall, 1998.
- *5. Lundstrom, M. *Fundamentals of Carrier Transport*. Vol. X of *Modular Series on Solid State Devices*. Reading, MA: Addison-Wesley, 1990.
- 6. Muller, R. S., and T. I. Kamins. *Device Electronics for Integrated Circuits*, 2nd ed. New York: Wiley, 1986.
- 7. Navon, D. H. *Semiconductor Microdevices and Materials*. New York: Holt, Rinehart & Winston, 1986.
- 8. Pierret, R. F. *Semiconductor Device Fundamentals*. Reading, MA: Addison-Wesley, 1996.
- 9. Shur, M. *Introduction to Electronic Devices*. New York: John Wiley and Sons, 1996.
- *10. Shur, M. *Physics of Semiconductor Devices*. Englewood Cliffs, NJ: Prentice Hall, 1990.
- 11. Singh, J. *Semiconductor Devices: An Introduction*. New York: McGraw-Hill, 1994.
- 12. Singh, J. *Semiconductor Devices: Basic Principles*. New York: John Wiley and Sons, 2001.
- 13. Streetman, B. G., and S. K. Banerjee. *Solid State Electronic Devices*, 6th ed. Upper Saddle River, NJ: Pearson Prentice Hall, 2006.
- 14. Sze, S. M. and K. K. Ng. *Physics of Semiconductor Devices*, 3rd ed. Hoboken, NJ: John Wiley and Sons, 2007.
- 15. Sze, S. M. *Semiconductor Devices: Physics and Technology*, 2nd ed. New York: John Wiley and Sons, 2001.
- *16. van der Ziel, A. *Solid State Physical Electronics*, 2nd ed. Englewood Cliffs, NJ: Prentice Hall, 1968.
- 17. Wang, S. *Fundamentals of Semiconductor Theory and Device Physics*. Englewood Cliffs, NJ: Prentice Hall, 1989.
- 18. Yang, E. S. *Microelectronic Devices*. New York: McGraw-Hill, 1988.

*Indicates references that are at an advanced level compared to this text.

The pn Junction

Up to this point in the text, we have been considering the properties of the semiconductor material. We calculated electron and hole concentrations in thermal equilibrium and determined the position of the Fermi level. We then considered the nonequilibrium condition in which excess electrons and holes are present in the semiconductor. We now wish to consider the situation in which a p-type and an n-type semiconductor are brought into contact with one another to form a pn junction.

Most semiconductor devices contain at least one junction between p-type and n-type semiconductor regions. Semiconductor device characteristics and operation are intimately connected to these pn junctions, so considerable attention is devoted initially to this basic device.

The electrostatics of the zero-biased and reverse-biased pn junction is considered in this chapter. The current–voltage characteristics of the pn junction diode are developed in the next chapter. ■

7.0 | PREVIEW

In this chapter, we will:

- Consider a uniformly doped pn junction, in which one region of the semiconductor is uniformly doped with acceptor atoms and the adjacent region is uniformly doped with donor atoms.
- Determine the energy-band diagram of a pn junction in thermal equilibrium.
- Discuss the creation of a space charge region between the p and n regions.
- Apply Poisson's equation to determine the electric field in the space charge region and calculate the built-in potential barrier.
- Analyze the changes that occur in the pn junction when a reverse-biased voltage is applied. Derive expressions for space charge width and depletion capacitance.
- Analyze the voltage breakdown characteristics of a pn junction.
- Consider the properties of a nonuniformly doped pn junction. Specific doping profiles can lead to desirable properties of the pn junction.

7.1 | BASIC STRUCTURE OF THE pn JUNCTION

Figure 7.1a schematically shows the pn junction. It is important to realize that the entire semiconductor is a single-crystal material in which one region is doped with acceptor impurity atoms to form the p region and the adjacent region is doped with donor atoms to form the n region. The interface separating the n and p regions is referred to as the *metallurgical junction*.

The impurity doping concentrations in the p and n regions are shown in Figure 7.1b. For simplicity, we will consider a *step junction* in which the doping concentration is uniform in each region and there is an abrupt change in doping at the junction. Initially, at the metallurgical junction, there is a very large density gradient in both electron and hole concentrations. Majority carrier electrons in the n region will begin diffusing into the p region, and majority carrier holes in the p region will begin diffusing into the n region. If we assume there are no external connections to the semiconductor, then this diffusion process cannot continue indefinitely. As electrons diffuse from the n region, positively charged donor atoms are left behind. Similarly, as holes diffuse from the p region, they uncover negatively charged acceptor atoms. The net positive and negative charges in the n and p regions induce an electric field in the region near the metallurgical junction, in the direction from the positive to the negative charge, or from the n to the p region.

The net positively and negatively charged regions are shown in Figure 7.2. These two regions are referred to as the *space charge region*. Essentially all electrons and holes are swept out of the space charge region by the electric field. Since the space charge region is depleted of any mobile charge, this region is also referred to as the

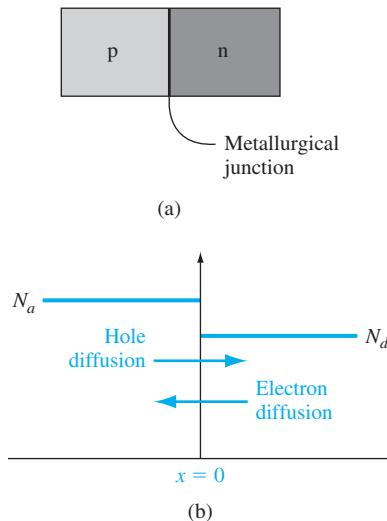


Figure 7.1 | (a) Simplified geometry of a pn junction; (b) doping profile of an ideal uniformly doped pn junction.

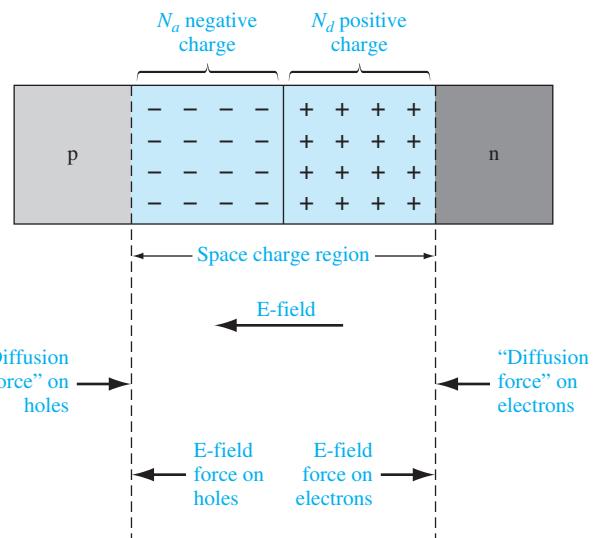


Figure 7.2 | The space charge region, the electric field, and the forces acting on the charged carriers.

depletion region; these two terms will be used interchangeably. Density gradients still exist in the majority carrier concentrations at each edge of the space charge region. We can think of a density gradient as producing a “diffusion force” that acts on the majority carriers. These diffusion forces, acting on the electrons and holes at the edges of the space charge region, are shown in the figure. The electric field in the space charge region produces another force on the electrons and holes, which is in the opposite direction to the diffusion force for each type of particle. In thermal equilibrium, the diffusion force and the E-field force exactly balance each other.

7.2 | ZERO APPLIED BIAS

We have considered the basic pn junction structure and discussed briefly how the space charge region is formed. In this section we will examine the properties of the step junction in thermal equilibrium, where no currents exist and no external excitation is applied. We will determine the space charge region width, electric field, and potential through the depletion region.

The analysis in this chapter is based on two assumptions that we have considered in previous chapters. The first assumption is that the Boltzmann approximation is valid, which means that each semiconductor region is nondegenerately doped. The second assumption is that complete ionization exists, which means that the temperature of the pn junction is not “too low.”

7.2.1 Built-in Potential Barrier

If we assume that no voltage is applied across the pn junction, then the junction is in thermal equilibrium—the Fermi energy level is constant throughout the entire system. Figure 7.3 shows the energy-band diagram for the pn junction in thermal equilibrium. The conduction and valence band energies must bend as we go through the space charge region, since the relative position of the conduction and valence bands with respect to the Fermi energy changes between p and n regions.

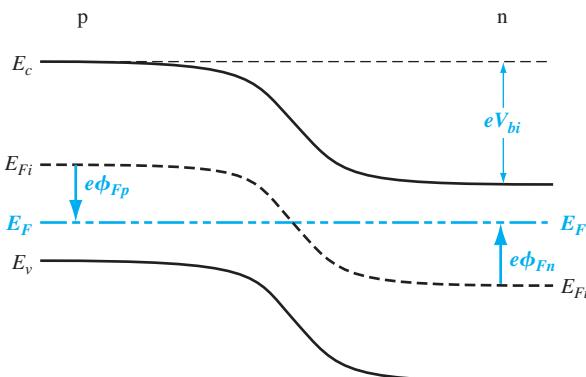


Figure 7.3 | Energy-band diagram of a pn junction in thermal equilibrium.

Electrons in the conduction band of the n region see a potential barrier in trying to move into the conduction band of the p region. This potential barrier is referred to as the *built-in potential barrier* and is denoted by V_{bi} . The built-in potential barrier maintains equilibrium between majority carrier electrons in the n region and minority carrier electrons in the p region, and also between majority carrier holes in the p region and minority carrier holes in the n region. This potential difference across the junction cannot be measured with a voltmeter because new potential barriers will be formed between the probes and the semiconductor that will cancel V_{bi} . The potential V_{bi} maintains equilibrium, so no current is produced by this voltage.

The intrinsic Fermi level is equidistant from the conduction band edge through the junction; thus, the built-in potential barrier can be determined as the difference between the intrinsic Fermi levels in the p and n regions. We can define the potentials ϕ_{Fn} and ϕ_{Fp} as shown in Figure 7.3, so we have

$$V_{bi} = |\phi_{Fn}| + |\phi_{Fp}| \quad (7.1)$$

In the n region, the electron concentration in the conduction band is given by

$$n_0 = N_c \exp \left[\frac{-(E_c - E_F)}{kT} \right] \quad (7.2)$$

which can also be written in the form

$$n_0 = n_i \exp \left[\frac{E_F - E_{Fi}}{kT} \right] \quad (7.3)$$

where n_i and E_{Fi} are the intrinsic carrier concentration and the intrinsic Fermi energy, respectively. We may define the potential ϕ_{Fn} in the n region as

$$e\phi_{Fn} = E_{Fi} - E_F \quad (7.4)$$

Equation (7.3) may then be written as

$$n_0 = n_i \exp \left[\frac{-(e\phi_{Fn})}{kT} \right] \quad (7.5)$$

Taking the natural log of both sides of Equation (7.5), setting $n_0 = N_d$, and solving for the potential, we obtain

$$\phi_{Fn} = \frac{-kT}{e} \ln \left(\frac{N_d}{n_i} \right) \quad (7.6)$$

Similarly, in the p region, the hole concentration is given by

$$p_0 = N_a = n_i \exp \left[\frac{E_{Fi} - E_F}{kT} \right] \quad (7.7)$$

where N_a is the acceptor concentration. We can define the potential ϕ_{Fp} in the p region as

$$e\phi_{Fp} = E_{Fi} - E_F \quad (7.8)$$

Combining Equations (7.7) and (7.8), we find that

$$\phi_{Fp} = +\frac{kT}{e} \ln \left(\frac{N_a}{n_i} \right) \quad (7.9)$$

Finally, the built-in potential barrier for the step junction is found by substituting Equations (7.6) and (7.9) into Equation (7.1), which yields

$$V_{bi} = \frac{kT}{e} \ln \left(\frac{N_a N_d}{n_i^2} \right) = V_t \ln \left(\frac{N_a N_d}{n_i^2} \right) \quad (7.10)$$

where $V_t = kT/e$ and is defined as the thermal voltage.

At this time, we should note a subtle but important point concerning notation. Previously in the discussion of a semiconductor material, N_d and N_a denoted donor and acceptor impurity concentrations in the same region, thereby forming a compensated semiconductor. From this point on in the text, N_d and N_a will denote the net donor and acceptor concentrations in the individual n and p regions, respectively. If the p region, for example, is a compensated material, then N_a will represent the difference between the actual acceptor and donor impurity concentrations. The parameter N_d is defined in a similar manner for the n region.

Objective: Calculate the built-in potential barrier in a pn junction.

Consider a silicon pn junction at $T = 300$ K with doping concentrations of $N_a = 2 \times 10^{17}$ cm $^{-3}$ and $N_d = 10^{15}$ cm $^{-3}$.

EXAMPLE 7.1

■ Solution

The built-in potential barrier is determined from Equation (7.10) as

$$V_{bi} = V_t \ln \left(\frac{N_a N_d}{n_i^2} \right) = (0.0259) \ln \left[\frac{(2 \times 10^{17})(10^{15})}{(1.5 \times 10^{10})^2} \right] = 0.713 \text{ V}$$

If we change the doping concentration in the p region of the pn junction such that the doping concentrations become $N_a = 10^{16}$ cm $^{-3}$ and $N_d = 10^{15}$ cm $^{-3}$, then the built-in potential barrier becomes $V_{bi} = 0.635$ V.

■ Comment

The built-in potential barrier changes only slightly as the doping concentrations change by orders of magnitude because of the logarithmic dependence.

■ EXERCISE PROBLEM

- Ex 7.1** (a) Calculate the built-in potential barrier in a silicon pn junction at $T = 300$ K for
 (i) $N_a = 5 \times 10^{15}$ cm $^{-3}$, $N_d = 10^{17}$ cm $^{-3}$ and (ii) $N_a = 2 \times 10^{16}$ cm $^{-3}$, $N_d = 2 \times 10^{15}$ cm $^{-3}$.
 (b) Repeat part (a) for a GaAs pn junction.

[Ans. (a) (i) 0.736 V, (iii) 0.671 V; (b) (i) 1.20 V, (iii) 1.14 V]

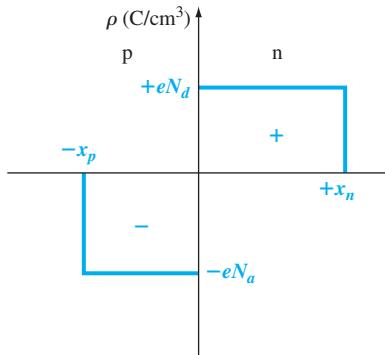


Figure 7.4 | The space charge density in a uniformly doped pn junction assuming the abrupt junction approximation.

7.2.2 Electric Field

An electric field is created in the depletion region by the separation of positive and negative space charge densities. Figure 7.4 shows the volume charge density distribution in the pn junction assuming uniform doping and assuming an abrupt junction approximation. We will assume that the space charge region abruptly ends in the n region at $x = +x_n$ and abruptly ends in the p region at $x = -x_p$ (x_p is a positive quantity).

The electric field is determined from Poisson's equation, which, for a one-dimensional analysis, is

$$\frac{d^2\phi(x)}{dx^2} = \frac{-\rho(x)}{\epsilon_s} = -\frac{dE(x)}{dx} \quad (7.11)$$

where $\phi(x)$ is the electric potential, $E(x)$ is the electric field, $\rho(x)$ is the volume charge density, and ϵ_s is the permittivity of the semiconductor. From Figure 7.4, the charge densities are

$$\rho(x) = -eN_a \quad -x_p < x < 0 \quad (7.12a)$$

and

$$\rho(x) = eN_d \quad 0 < x < x_n \quad (7.12b)$$

The electric field in the p region is found by integrating Equation (7.11). We have

$$E = \int \frac{\rho(x)}{\epsilon_s} dx = - \int \frac{eN_a}{\epsilon_s} dx = \frac{-eN_a}{\epsilon_s} x + C_1 \quad (7.13)$$

where C_1 is a constant of integration. The electric field is assumed to be zero in the neutral p region for $x < -x_p$ since the currents are zero in thermal equilibrium. Since there are no surface charge densities within the pn junction structure, the electric

field is a continuous function. The constant of integration is determined by setting $E = 0$ at $x = -x_p$. The electric field in the p region is then given by

$$E = \frac{-eN_a}{\epsilon_s}(x + x_p) \quad -x_p \leq x \leq 0 \quad (7.14)$$

In the n region, the electric field is determined from

$$E = \int \frac{(eN_d)}{\epsilon_s} dx = \frac{eN_d}{\epsilon_s} x + C_2 \quad (7.15)$$

where C_2 is again a constant of integration and is determined by setting $E = 0$ at $x = x_n$, since the E-field is assumed to be zero in the n region and is a continuous function. Then

$$E = \frac{-eN_d}{\epsilon_s}(x_n - x) \quad 0 \leq x \leq x_n \quad (7.16)$$

The electric field is also continuous at the metallurgical junction, or at $x = 0$. Setting Equations (7.14) and (7.16) equal to each other at $x = 0$ gives

$$N_a x_p = N_d x_n \quad (7.17)$$

Equation (7.17) states that the number of negative charges per unit area in the p region is equal to the number of positive charges per unit area in the n region.

Figure 7.5 is a plot of the electric field in the depletion region. The electric field direction is from the n to the p region, or in the negative x direction for this geometry. For the uniformly doped pn junction, the E-field is a linear function of distance through the junction, and the maximum (magnitude) electric field occurs at the metallurgical junction. An electric field exists in the depletion region even when no voltage is applied between the p and n regions.

The potential in the junction is found by integrating the electric field. In the p region then, we have

$$\phi(x) = - \int E(x) dx = \int \frac{eN_a}{\epsilon_s} (x + x_p) dx \quad (7.18)$$

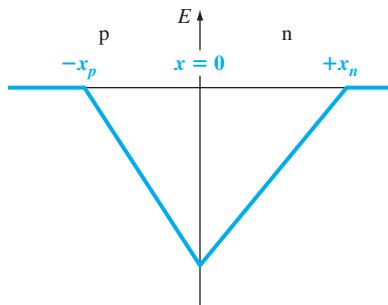


Figure 7.5 | Electric field in the space charge region of a uniformly doped pn junction.

or

$$\phi(x) = \frac{eN_a}{\epsilon_s} \left(\frac{x^2}{2} + x_p \cdot x \right) + C'_1 \quad (7.19)$$

where C'_1 is again a constant of integration. The potential difference through the pn junction is the important parameter, rather than the absolute potential, so we may arbitrarily set the potential equal to zero at $x = -x_p$. The constant of integration is then found as

$$C'_1 = \frac{eN_a}{2\epsilon_s} x_p^2 \quad (7.20)$$

so that the potential in the p region can now be written as

$$\phi(x) = \frac{eN_a}{2\epsilon_s} (x + x_p)^2 \quad (-x_p \leq x \leq 0) \quad (7.21)$$

The potential in the n region is determined by integrating the electric field in the n region, or

$$\phi(x) = \int \frac{eN_d}{\epsilon_s} (x_n - x) dx \quad (7.22)$$

Then

$$\phi(x) = \frac{eN_d}{\epsilon_s} \left(x_n \cdot x - \frac{x^2}{2} \right) + C'_2 \quad (7.23)$$

where C'_2 is another constant of integration. The potential is a continuous function, so setting Equation (7.21) equal to Equation (7.23) at the metallurgical junction, or at $x = 0$, gives

$$C'_2 = \frac{eN_a}{2\epsilon_s} x_p^2 \quad (7.24)$$

The potential in the n region can thus be written as

$$\phi(x) = \frac{eN_d}{\epsilon_s} \left(x_n \cdot x - \frac{x^2}{2} \right) + \frac{eN_a}{2\epsilon_s} x_p^2 \quad (0 \leq x \leq x_n) \quad (7.25)$$

Figure 7.6 is a plot of the potential through the junction and shows the quadratic dependence on distance. The magnitude of the potential at $x = x_n$ is equal to the built-in potential barrier. Then from Equation (7.25), we have

$$V_{bi} = |\phi(x = x_n)| = \frac{e}{2\epsilon_s} (N_d x_n^2 + N_a x_p^2) \quad (7.26)$$

The potential energy of an electron is given by $E = -e\phi$, which means that the electron potential energy also varies as a quadratic function of distance through the space charge region. The quadratic dependence on distance was shown in the energy-band diagram of Figure 7.3, although we did not explicitly know the shape of the curve at that time.

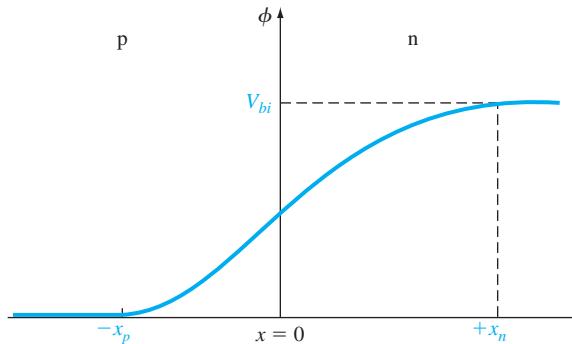


Figure 7.6 | Electric potential through the space charge region of a uniformly doped pn junction.

7.2.3 Space Charge Width

We can determine the distance that the space charge region extends into the p and n regions from the metallurgical junction. This distance is known as the space charge width. From Equation (7.17), we may write, for example,

$$x_p = \frac{N_d x_n}{N_a} \quad (7.27)$$

Then, substituting Equation (7.27) into Equation (7.26) and solving for x_n , we obtain

$$x_n = \left\{ \frac{2\epsilon_s V_{bi}}{e} \left[\frac{N_a}{N_d} \right] \left[\frac{1}{N_a + N_d} \right] \right\}^{1/2} \quad (7.28)$$

Equation (7.28) gives the space charge width, or the width of the depletion region, x_n extending into the n-type region for the case of zero applied voltage.

Similarly, if we solve for x_p from Equation (7.17) and substitute into Equation (7.26), we find

$$x_p = \left\{ \frac{2\epsilon_s V_{bi}}{e} \left[\frac{N_d}{N_a} \right] \left[\frac{1}{N_a + N_d} \right] \right\}^{1/2} \quad (7.29)$$

where x_p is the width of the depletion region extending into the p region for the case of zero applied voltage.

The total depletion or space charge width W is the sum of the two components, or

$$W = x_n + x_p \quad (7.30)$$

Using Equations (7.28) and (7.29), we obtain

$$W = \left\{ \frac{2\epsilon_s V_{bi}}{e} \left[\frac{N_a + N_d}{N_a N_d} \right] \right\}^{1/2} \quad (7.31)$$

The built-in potential barrier can be determined from Equation (7.10), and then the total space charge region width is obtained using Equation (7.31).

EXAMPLE 7.2

Objective: Calculate the space charge width and electric field in a pn junction for zero bias.

Consider a silicon pn junction at $T = 300$ K with doping concentrations of $N_a = 10^{16} \text{ cm}^{-3}$ and $N_d = 10^{15} \text{ cm}^{-3}$.

■ Solution

In Example 7.1, we determined the built-in potential barrier as $V_{bi} = 0.635$ V. From Equation (7.31), the space charge width is

$$\begin{aligned} W &= \left\{ \frac{2\epsilon_s V_{bi}}{e} \left[\frac{N_a + N_d}{N_a N_d} \right] \right\}^{1/2} \\ &= \left\{ \frac{2(11.7)(8.85 \times 10^{-14})(0.635)}{1.6 \times 10^{-19}} \left[\frac{10^{16} + 10^{15}}{(10^{16})(10^{15})} \right] \right\}^{1/2} \\ &= 0.951 \times 10^{-4} \text{ cm} = 0.951 \mu\text{m} \end{aligned}$$

Using Equations (7.28) and (7.29), we can find $x_n = 0.8644 \mu\text{m}$, and $x_p = 0.0864 \mu\text{m}$.

The peak electric field at the metallurgical junction, using Equation (7.16) for example, is

$$E_{max} = -\frac{e N_d x_n}{\epsilon_s} = -\frac{(1.6 \times 10^{-19})(10^{15})(0.8644 \times 10^{-4})}{(11.7)(8.85 \times 10^{-14})} = -1.34 \times 10^4 \text{ V/cm}$$

■ Comment

The peak electric field in the space charge region of a pn junction is quite large. We must keep in mind, however, that there is no mobile charge in this region; hence there will be no drift current. We may also note, from this example, that the width of each space charge region is a reciprocal function of the doping concentration: The depletion region will extend further into the lower-doped region.

■ EXERCISE PROBLEM

Ex 7.2 A silicon pn junction at $T = 300$ K with zero applied bias has doping concentrations of $N_d = 5 \times 10^{16} \text{ cm}^{-3}$ and $N_a = 5 \times 10^{15} \text{ cm}^{-3}$. Determine x_n , x_p , W , and $|E_{max}|$.

$$x_n = 4.11 \times 10^{-5} \text{ cm}, W = 4.52 \times 10^{-5} \text{ cm}, |E_{max}| = 3.18 \times 10^4 \text{ V/cm}$$

$$(\text{Ans. } x_p = 4.11 \times 10^{-6} \text{ cm})$$

TEST YOUR UNDERSTANDING

TYU 7.1 Calculate V_{bi} , x_n , x_p , W , and $|E_{max}|$ for a silicon pn junction at zero bias and $T = 300$ K for doping concentrations of (a) $N_a = 2 \times 10^{17} \text{ cm}^{-3}$, $N_d = 10^{16} \text{ cm}^{-3}$ and

$$(b) N_a = 4 \times 10^{15} \text{ cm}^{-3}$$

$$N_d = 3 \times 10^{16} \text{ cm}^{-3}, W = 0.5064 \mu\text{m}, |E_{max}| = 2.76 \times 10^4 \text{ V/cm}$$

$$E_{max} = 4.77 \times 10^4 \text{ V/cm}; (b) V_{bi} = 0.669 \text{ V}, x_n = 0.0596 \mu\text{m}, x_p = 0.4469 \mu\text{m},$$

$$W = 0.3240 \mu\text{m}, (\text{Ans. } (a) V_{bi} = 0.772 \text{ V}, x_n = 0.3085 \mu\text{m}, x_p = 0.0154 \mu\text{m}, W = 0.3240 \mu\text{m},$$

TYU 7.2 Repeat Exercise Problem Ex 7.2 for a GaAs pn junction. ($E_{max} = 3.86 \times 10^4 \text{ V/cm}$)

$$(\text{Ans. } V_{bi} = 1.186 \text{ V}, x_n = 0.05590 \mu\text{m}, x_p = 0.5590 \mu\text{m}, W = 0.6149 \mu\text{m})$$

7.3 | REVERSE APPLIED BIAS

If we apply a potential between the p and n regions, we will no longer be in an equilibrium condition—the Fermi energy level will no longer be constant through the system. Figure 7.7 shows the energy-band diagram of the pn junction for the case when a positive voltage is applied to the n region with respect to the p region. As the positive potential is downward, the Fermi level on the n side is below the Fermi level on the p side. The difference between the two is equal to the applied voltage in units of energy.

The total potential barrier, indicated by V_{total} , has increased. The applied potential is the reverse-biased condition. The total potential barrier is now given by

$$V_{\text{total}} = |\phi_{F_n}| + |\phi_{F_p}| + V_R \quad (7.32)$$

where V_R is the magnitude of the applied reverse-biased voltage. Equation (7.32) can be rewritten as

$$V_{\text{total}} = V_{bi} + V_R \quad (7.33)$$

where V_{bi} is the same built-in potential barrier we had defined in thermal equilibrium.

7.3.1 Space Charge Width and Electric Field

Figure 7.8 shows a pn junction with an applied reverse-biased voltage V_R . Also indicated in the figure are the electric field in the space charge region and the electric field E_{app} , induced by the applied voltage. The electric fields in the neutral p and n regions are essentially zero, or at least very small, which means that the magnitude of the electric field in the space charge region must increase above the thermal-equilibrium value due to the applied voltage. The electric field originates on positive charge and terminates on negative charge; this means that the number of positive and negative

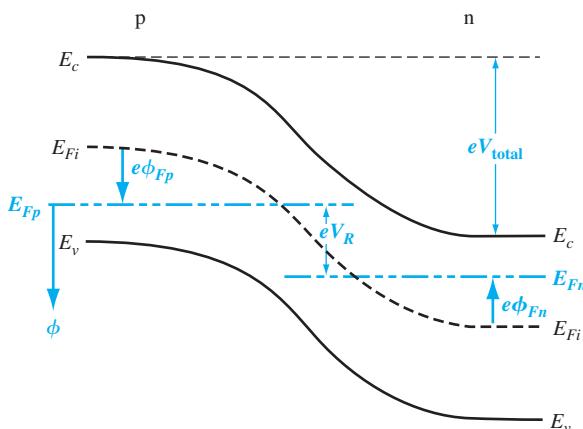


Figure 7.7 | Energy-band diagram of a pn junction under reverse bias.

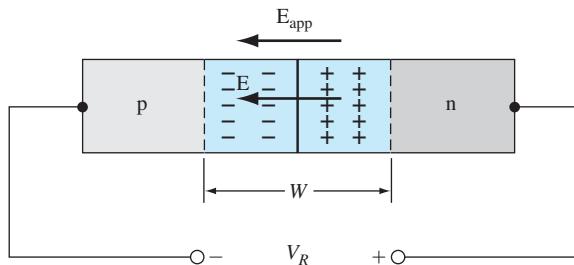


Figure 7.8 | A pn junction, with an applied reverse-biased voltage, showing the directions of the electric field induced by V_R and the space charge electric field.

charges must increase if the electric field increases. For given impurity doping concentrations, the number of positive and negative charges in the depletion region can be increased only if the space charge width W increases. The space charge width W increases, therefore, with an increasing reverse-biased voltage V_R . We are assuming that the electric field in the bulk n and p regions is zero. This assumption will become clearer in the next chapter when we discuss the current–voltage characteristics.

In all of the previous equations, the built-in potential barrier can be replaced by the total potential barrier. The total space charge width can be written from Equation (7.31) as

$$W = \left\{ \frac{2\epsilon_s(V_{bi} + V_R)}{e} \left[\frac{N_a + N_d}{N_a N_d} \right] \right\}^{1/2} \quad (7.34)$$

showing that the total space charge width increases as we apply a reverse-biased voltage. By substituting the total potential barrier V_{total} into Equations (7.28) and (7.29), the space charge widths in the n and p regions, respectively, can be found as a function of applied reverse-biased voltage.

EXAMPLE 7.3

Objective: Calculate the width of the space charge region in a pn junction when a reverse-biased voltage is applied.

Again consider a silicon pn junction at $T = 300$ K with doping concentrations of $N_a = 10^{16} \text{ cm}^{-3}$ and $N_d = 10^{15} \text{ cm}^{-3}$. Assume that $n_i = 1.5 \times 10^{10} \text{ cm}^{-3}$ and $V_R = 5 \text{ V}$.

■ Solution

The built-in potential barrier was calculated in Example 7.1 for this case and is $V_{bi} = 0.635 \text{ V}$. The space charge width is determined from Equation (7.34). We have

$$W = \left\{ \frac{2(11.7)(8.85 \times 10^{-14})(0.635 + 5)}{1.6 \times 10^{-19}} \left[\frac{10^{16} + 10^{15}}{(10^{16})(10^{15})} \right] \right\}^{1/2}$$

so that

$$W = 2.83 \times 10^{-4} \text{ cm} = 2.83 \mu\text{m}$$

Comment

The space charge width has increased from $0.951 \mu\text{m}$ at zero bias to $2.83 \mu\text{m}$ at a reverse bias of 5 V .

EXERCISE PROBLEM

- Ex 7.3** (a) A silicon pn junction at $T = 300 \text{ K}$ has doping concentrations of $N_a = 5 \times 10^{15} \text{ cm}^{-3}$ and $N_d = 5 \times 10^{16} \text{ cm}^{-3}$. A reverse-biased voltage of $V_R = 4 \text{ V}$ is applied. Determine V_{bi} , x_n , x_p , and W . (b) Repeat part (a) for $V_R = 8 \text{ V}$.
 (q) $V_b = 0.718 \text{ V}$, $x_n = 0.1432 \mu\text{m}$, $x_p = 1.432 \mu\text{m}$, $W = 1.576 \mu\text{m}$
 [Ans. (a) $V_b = 0.718 \text{ V}$, $x_n = 0.1054 \mu\text{m}$, $x_p = 1.054 \mu\text{m}$, $W = 1.159 \mu\text{m}$]

The magnitude of the electric field in the depletion region increases with an applied reverse-biased voltage. The electric field is still given by Equations (7.14) and (7.16) and is still a linear function of distance through the space charge region. Since x_n and x_p increase with reverse-biased voltage, the magnitude of the electric field also increases. The maximum electric field still occurs at the metallurgical junction.

The maximum electric field at the metallurgical junction, from Equations (7.14) and (7.16), is

$$E_{\max} = \frac{-eN_d x_n}{\epsilon_s} = \frac{-eN_d x_p}{\epsilon_s} \quad (7.35)$$

If we use either Equation (7.28) or (7.29) in conjunction with the total potential barrier, $V_{bi} + V_R$, then

$$E_{\max} = -\left\{\frac{2e(V_{bi} + V_R)}{\epsilon_s} \left(\frac{N_a N_d}{N_a + N_d}\right)\right\}^{1/2} \quad (7.36)$$

We can show that the maximum electric field in the pn junction can also be written as

$$E_{\max} = \frac{-2(V_{bi} + V_R)}{W} \quad (7.37)$$

where W is the total space charge width.

Objective: Design a pn junction to meet maximum electric field and voltage specifications.

Consider a silicon pn junction at $T = 300 \text{ K}$ with a p-type doping concentration of $N_a = 2 \times 10^{17} \text{ cm}^{-3}$. Determine the n-type doping concentration such that the maximum electric field is $|E_{\max}| = 2.5 \times 10^5 \text{ V/cm}$ at a reverse-biased voltage of $V_R = 25 \text{ V}$.

Solution

The maximum electric field is given by Equation (7.36). Neglecting V_{bi} compared to V_R , we can write

$$|E_{\max}| \cong \left\{\frac{2eV_R}{\epsilon_s} \left(\frac{N_a N_d}{N_a + N_d}\right)\right\}^{1/2}$$

DESIGN EXAMPLE 7.4

or

$$2.5 \times 10^5 = \left\{ \frac{2(1.6 \times 10^{-19})(25)}{(11.7)(8.85 \times 10^{-14})} \left[\frac{(2 \times 10^{17})N_d}{2 \times 10^{17} + N_d} \right] \right\}^{1/2}$$

which yields

$$N_d = 8.43 \times 10^{15} \text{ cm}^{-3}$$

Conclusion

A smaller value of N_d results in a smaller value of $|E_{\max}|$ for a given reverse-biased voltage. The value of N_d determined in this example, then, is the maximum value that will meet the specifications.

EXERCISE PROBLEM

- Ex 7.4** The maximum electric field in a reverse-biased GaAs pn junction at $T = 300$ K is to be limited to $|E_{\max}| = 7.2 \times 10^4$ V/cm. The doping concentrations are $N_d = 5 \times 10^{15} \text{ cm}^{-3}$ and $N_a = 3 \times 10^{16} \text{ cm}^{-3}$. Determine the maximum reverse-biased voltage that can be applied. ($\Lambda I_2 \Sigma = 3 \text{ A} \cdot \text{s} \cdot \text{V}$)

7.3.2 Junction Capacitance

Since we have a separation of positive and negative charges in the depletion region, a capacitance is associated with the pn junction. Figure 7.9 shows the charge densities in the depletion region for applied reverse-biased voltages of V_R and $V_R + dV_R$. An increase

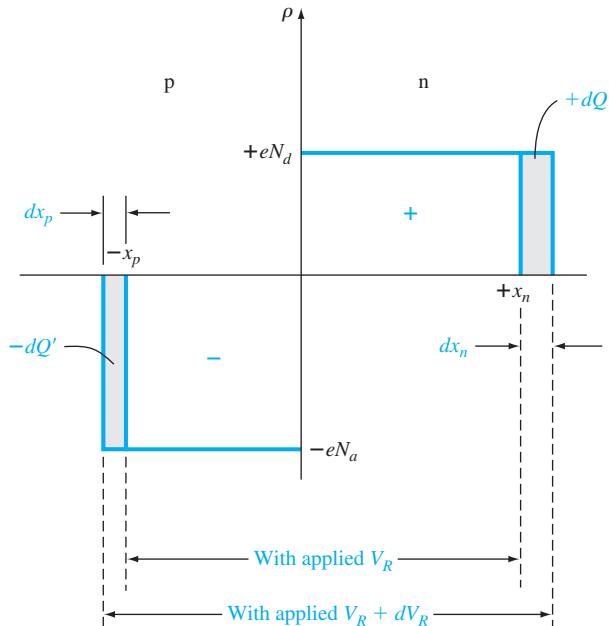


Figure 7.9 | Differential change in the space charge width with a differential change in reverse-biased voltage for a uniformly doped pn junction.

in the reverse-biased voltage dV_R will uncover additional positive charges in the n region and additional negative charges in the p region. The junction capacitance is defined as

$$C' = \frac{dQ'}{dV_R} \quad (7.38)$$

where

$$dQ' = eN_d dx_n = eN_a dx_p \quad (7.39)$$

The differential charge dQ' is in units of C/cm² so that the capacitance C' is in units of farads per square centimeter F/cm², or capacitance per unit area.

For the total potential barrier, Equation (7.28) may be written as

$$x_n = \left\{ \frac{2\epsilon_s(V_{bi} + V_R)}{e} \left[\frac{N_a}{N_d} \right] \left[\frac{1}{N_a + N_d} \right] \right\}^{1/2} \quad (7.40)$$

The junction capacitance can be written as

$$C' = \frac{dQ'}{dV_R} = eN_d \frac{dx_n}{dV_R} \quad (7.41)$$

so that

$$C' = \left\{ \frac{e\epsilon_s N_a N_d}{2(V_{bi} + V_R)(N_a + N_d)} \right\}^{1/2} \quad (7.42)$$

Exactly the same capacitance expression is obtained by considering the space charge region extending into the p region x_p . The junction capacitance is also referred to as the *depletion layer capacitance*.

Objective: Calculate the junction capacitance of a pn junction.

EXAMPLE 7.5

Consider the same pn junction as that in Example 7.3. Again assume that $V_R = 5$ V.

Solution

The junction capacitance is found from Equation (7.42) as

$$C' = \left\{ \frac{(1.6 \times 10^{-19})(11.7)(8.85 \times 10^{-14})(10^{16})(10^{15})}{2(0.635 + 5)(10^{16} + 10^{15})} \right\}^{1/2}$$

or

$$C' = 3.66 \times 10^{-9} \text{ F/cm}^2$$

If the cross-sectional area of the pn junction is, for example, $A = 10^{-4} \text{ cm}^2$, then the total junction capacitance is

$$C = C' \cdot A = 0.366 \times 10^{-12} \text{ F} = 0.366 \text{ pF}$$

Comment

The value of junction capacitance is usually in the pF, or smaller, range.

■ EXERCISE PROBLEM

- Ex 7.5** Consider a GaAs pn junction at $T = 300$ K doped to $N_a = 5 \times 10^{15} \text{ cm}^{-3}$ and $N_d = 2 \times 10^{16} \text{ cm}^{-3}$. (a) Calculate V_{bi} . (b) Determine the junction capacitance C' for $V_R = 4$ V. (c) Repeat part (b) for $V_R = 8$ V.

[Ans. (a) $V_{bi} = 1.16$ V; (b) $C' = 8.48 \times 10^{-9} \text{ F/cm}^2$; (c) $C' = 6.36 \times 10^{-9} \text{ F/cm}^2$]

If we compare Equation (7.34) for the total depletion width W of the space charge region under reverse bias and Equation (7.42) for the junction capacitance C' , we find that we can write

$$C' = \frac{\epsilon_s}{W} \quad (7.43)$$

Equation (7.43) is the same as the capacitance per unit area of a parallel plate capacitor. Considering Figure 7.9, we may have come to this same conclusion earlier. Keep in mind that the space charge width is a function of the reverse-biased voltage so that the junction capacitance is also a function of the reverse-biased voltage applied to the pn junction.

7.3.3 One-Sided Junctions

Consider a special pn junction called the one-sided junction. If, for example, $N_a \gg N_d$, this junction is referred to as a p⁺n junction. The total space charge width, from Equation (7.34), reduces to

$$W \approx \left\{ \frac{2\epsilon_s(V_{bi} + V_R)}{eN_d} \right\}^{1/2} \quad (7.44)$$

Considering the expressions for x_n and x_p , we have for the p⁺n junction

$$x_p \ll x_n \quad (7.45)$$

and

$$W \approx x_n \quad (7.46)$$

Almost the entire space charge layer extends into the low-doped region of the junction. This effect can be seen in Figure 7.10.

The junction capacitance of the p⁺n junction reduces to

$$C' \approx \left\{ \frac{e\epsilon_s N_d}{2(V_{bi} + V_R)} \right\}^{1/2} \quad (7.47)$$

The depletion layer capacitance of a one-sided junction is a function of the doping concentration in the low-doped region. Equation (7.47) may be manipulated to give

$$\left(\frac{1}{C'} \right)^2 = \frac{2(V_{bi} + V_R)}{e\epsilon_s N_d} \quad (7.48)$$

which shows that the inverse capacitance squared is a linear function of applied reverse-biased voltage.

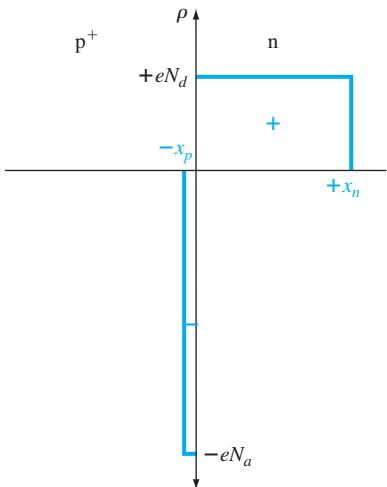


Figure 7.10 | Space charge density of a one-sided p^-n junction.

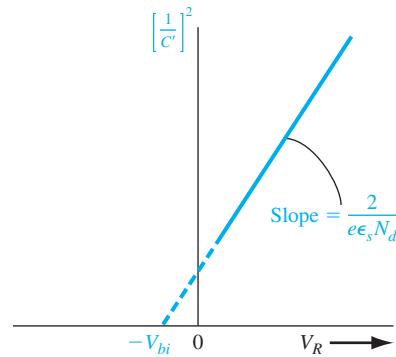


Figure 7.11 | $(1/C')^2$ versus V_R of a uniformly doped pn junction.

Figure 7.11 shows a plot of Equation (7.48). The built-in potential of the junction can be determined by extrapolating the curve to the point where $(1/C')^2 = 0$. The slope of the curve is inversely proportional to the doping concentration of the low-doped region in the junction; thus, this doping concentration can be experimentally determined. The assumptions used in the derivation of this capacitance include uniform doping in both semiconductor regions, the abrupt junction approximation, and a planar junction.

Objective: Determine the impurity doping concentrations in a p^+n junction given the parameters from Figure 7.11.

EXAMPLE 7.6

Assume that the intercept and the slope of the curve in Figure 7.11 are $V_{bi} = 0.725$ V and $6.15 \times 10^{15} (\text{F/cm}^2)^{-2} (\text{V})^{-1}$, respectively, for a silicon p^+n junction at $T = 300$ K.

Solution

The slope of the curve in Figure 7.11 is given by $2/e \epsilon_s N_d$, so we may write

$$N_d = \frac{2}{e \epsilon_s} \cdot \frac{1}{\text{slope}} = \frac{2}{(1.6 \times 10^{-19})(11.7)(8.85 \times 10^{-14})(6.15 \times 10^{15})}$$

or

$$N_d = 1.96 \times 10^{15} \text{ cm}^{-3}$$

From the expression for V_{bi} , which is

$$V_{bi} = V_t \ln \left(\frac{N_a N_d}{n_i^2} \right)$$

we can solve for N_a as

$$N_a = \frac{n_i^2}{N_d} \exp \left(\frac{V_{bi}}{V_t} \right) = \frac{(1.5 \times 10^{10})^2}{1.963 \times 10^{15}} \exp \left(\frac{0.725}{0.0259} \right)$$

which yields

$$N_a = 1.64 \times 10^{17} \text{ cm}^{-3}$$

Comment

The results of this example show that $N_a \gg N_d$; therefore the assumption of a one-sided junction was valid.

EXERCISE PROBLEM

- Ex 7.6** The experimentally measured junction capacitance of a one-sided silicon n⁺p junction biased at $V_R = 3$ V and at $T = 300$ K is $C = 0.105$ pF. The built-in potential barrier is found to be $V_{bi} = 0.765$ V. The cross-sectional area is $A = 10^{-5}$ cm². Find the doping concentrations. (Ans. $N_n = 5.01 \times 10^{15} \text{ cm}^{-3}$, $N_p = 3.02 \times 10^{17} \text{ cm}^{-3}$)

A one-sided pn junction is useful for experimentally determining the doping concentrations and built-in potential.

TEST YOUR UNDERSTANDING

- TYU 7.3** (a) A silicon pn junction at $T = 300$ K is reverse biased at $V_R = 8$ V. The doping concentrations are $N_a = 5 \times 10^{16} \text{ cm}^{-3}$ and $N_d = 5 \times 10^{15} \text{ cm}^{-3}$. Determine x_n , x_p , W , and $|E_{max}|$. (b) Repeat part (a) for a reverse-biased voltage of $V_R = 12$ V.

$$\begin{aligned} x_n &= 1.73 \times 10^{-5} \text{ cm}, W = 1.90 \times 10^{-4} \text{ cm}, |E_{max}| = 1.34 \times 10^5 \text{ V/cm} \\ x_p &= 1.57 \times 10^{-4} \text{ cm}, |E_{max}| = 1.11 \times 10^5 \text{ V/cm}; (b) x_n = 1.73 \times 10^{-4} \text{ cm}, \\ W &= 1.43 \times 10^{-4} \text{ cm}, x_p = 1.43 \times 10^{-5} \text{ cm}, \end{aligned}$$

(Ans. (a) $x_n = 1.43 \times 10^{-4} \text{ cm}$, $x_p = 1.43 \times 10^{-5} \text{ cm}$,

- TYU 7.4** A silicon pn junction at $T = 300$ K has doping concentrations of $N_d = 3 \times 10^{16} \text{ cm}^{-3}$ and $N_a = 8 \times 10^{15} \text{ cm}^{-3}$, and has a cross-sectional area of $A = 5 \times 10^{-5} \text{ cm}^2$. Determine the junction capacitance at (a) $V_R = 2$ V and (b) $V_R = 5$ V. (Ans. (a) 0.694 pF; (b) 0.478 pF)

7.4 | JUNCTION BREAKDOWN

In the last section, we determined the effects of applying a reverse-biased voltage across the pn junction. However, the reverse-biased voltage may not increase without limit; at some particular voltage, the reverse-biased current will increase rapidly. The applied voltage at this point is called the *breakdown voltage*.

Two physical mechanisms give rise to the reverse-biased breakdown in a pn junction: the *Zener effect* and the *avalanche effect*. Zener breakdown occurs in highly doped pn junctions through a tunneling mechanism. In a highly doped junction, the conduction and valence bands on opposite sides of the junction are sufficiently close during reverse bias that electrons may tunnel directly from the valence band on the p side into the conduction band on the n side. This tunneling process is schematically shown in Figure 7.12a.

The avalanche breakdown process occurs when electrons and/or holes, moving across the space charge region, acquire sufficient energy from the electric field to create electron–hole pairs by colliding with atomic electrons within the depletion region. The avalanche process is schematically shown in Figure 7.12b. The newly

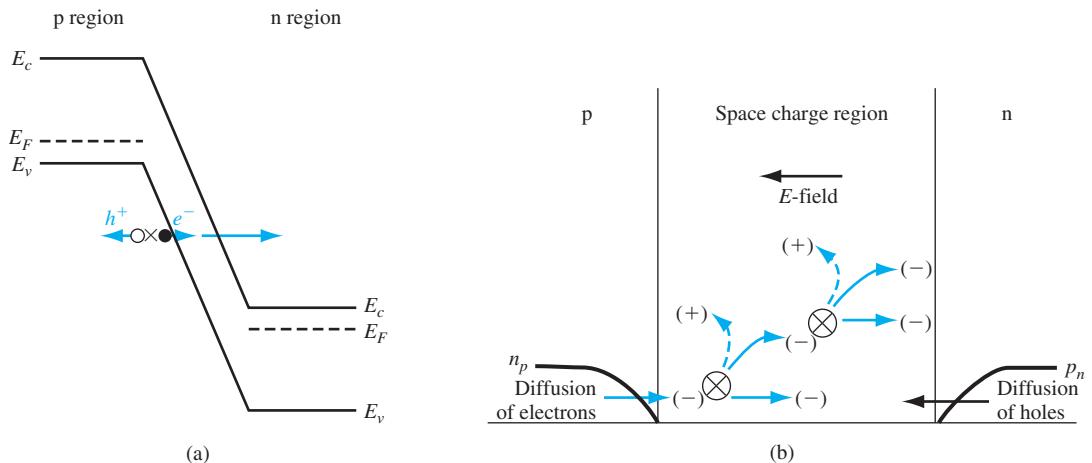


Figure 7.12 | (a) Zener breakdown mechanism in a reverse-biased pn junction; (b) avalanche breakdown process in a reverse-biased pn junction.

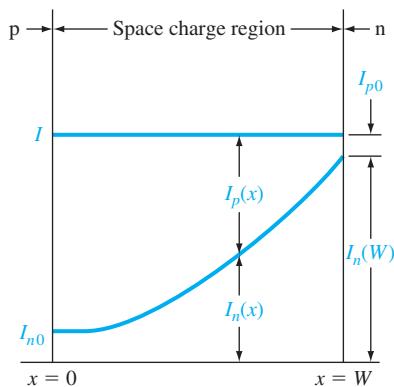


Figure 7.13 | Electron and hole current components through the space charge region during avalanche multiplication.

created electrons and holes move in opposite directions due to the electric field and thereby create a reverse-biased current. In addition, the newly generated electrons and/or holes may acquire sufficient energy to ionize other atoms, leading to the avalanche process. For most pn junctions, the predominant breakdown mechanism will be the avalanche effect.

If we assume that a reverse-biased electron current I_{n0} enters the depletion region at $x = 0$ as shown in Figure 7.13, the electron current I_n will increase with distance through the depletion region due to the avalanche process. At $x = W$, the electron current may be written as

$$I_n(W) = M_n I_{n0} \quad (7.49)$$

where M_n is a multiplication factor. The hole current is increasing through the depletion region from the n to p region and reaches a maximum value at $x = 0$. The total current is constant through the pn junction in steady state.

We can write an expression for the incremental electron current at some point x as

$$dI_n(x) = I_n(x)\alpha_n dx + I_p(x)\alpha_p dx \quad (7.50)$$

where α_n and α_p are the electron and hole ionization rates, respectively. The ionization rates are the number of electron–hole pairs generated per unit length by an electron (α_n) or by a hole (α_p). Equation (7.50) may be written as

$$\frac{dI_n(x)}{dx} = I_n(x)\alpha_n + I_p(x)\alpha_p \quad (7.51)$$

The total current I is given by

$$I = I_n(x) + I_p(x) \quad (7.52)$$

which is a constant. Solving for $I_p(x)$ from Equation (7.52) and substituting into Equation (7.51), we obtain

$$\frac{dI_n(x)}{dx} + (\alpha_p - \alpha_n)I_n(x) = \alpha_p I \quad (7.53)$$

If we make the assumption that the electron and hole ionization rates are equal so that

$$\alpha_n = \alpha_p \equiv \alpha \quad (7.54)$$

then Equation (7.53) may be simplified and integrated through the space charge region. We will obtain

$$I_n(W) - I_n(0) = I \int_0^W \alpha dx \quad (7.55)$$

Using Equation (7.49), Equation (7.55) may be written as

$$\frac{M_n I_{n0} - I_n(0)}{I} = \int_0^W \alpha dx \quad (7.56)$$

Since $M_n I_{n0} \approx I$ and since $I_n(0) = I_{n0}$, Equation (7.56) becomes

$$1 - \frac{1}{M_n} = \int_0^W \alpha dx \quad (7.57)$$

The avalanche breakdown voltage is defined to be the voltage at which M_n approaches infinity. The avalanche breakdown condition is then given by

$$\int_0^W \alpha dx = 1 \quad (7.58)$$

The ionization rates are strong functions of electric field and, since the electric field is not constant through the space charge region, Equation (7.58) is not easy to evaluate.

If we consider, for example, a one-sided p⁺n junction, the maximum electric field is given by

$$E_{\max} = \frac{eN_d x_n}{\epsilon_s} \quad (7.59)$$

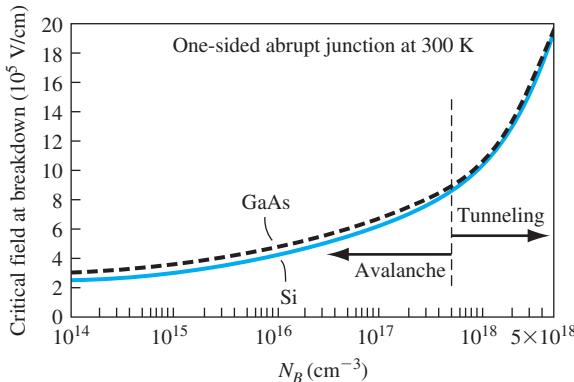


Figure 7.14 | Critical electric field at breakdown in a one-sided junction as a function of impurity doping concentrations.
(From Sze and Ng[14].)

The depletion width x_n is given approximately as

$$x_n \approx \left\{ \frac{2\epsilon_s V_R}{e} \cdot \frac{1}{N_d} \right\}^{1/2} \quad (7.60)$$

where V_R is the magnitude of the applied reverse-biased voltage. We have neglected the built-in potential V_{bi} .

If we now define V_R to be the breakdown voltage V_B , the maximum electric field, E_{max} , will be defined as a critical electric field, E_{crit} , at breakdown. Combining Equations (7.59) and (7.60), we may write

$$V_B = \frac{\epsilon_s E_{crit}^2}{2eN_B} \quad (7.61)$$

where N_B is the semiconductor doping in the low-doped region of the one-sided junction. The critical electric field, plotted in Figure 7.14, is a slight function of doping.

We have been considering a uniformly doped planar junction. The breakdown voltage will decrease for a linearly graded junction. (See Section 7.5.) Figure 7.15 shows a plot of the breakdown voltage for a one-sided abrupt junction and a linearly graded junction. If we take into account the curvature of a diffused junction as well, the breakdown voltage will be further degraded.

Objective: Design an ideal one-sided n⁺p junction diode to meet a breakdown voltage specification.

Consider a silicon pn junction diode at $T = 300 \text{ K}$. Assume that $N_d = 3 \times 10^{18} \text{ cm}^{-3}$. Design the diode such that the breakdown voltage is $V_B = 100 \text{ V}$.

■ Solution

From Figure 7.15, we find that the doping concentration in the low-doped side of a one-sided abrupt junction should be approximately $4 \times 10^{15} \text{ cm}^{-3}$ for a breakdown voltage of 100 V.

DESIGN EXAMPLE 7.7

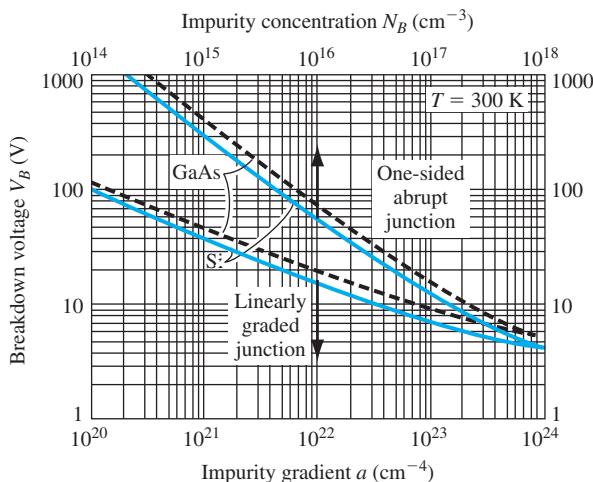


Figure 7.15 | Breakdown voltage versus impurity concentration in uniformly doped and linearly graded junctions.
(From Sze [14].)

For a doping concentration of $4 \times 10^{15} \text{ cm}^{-3}$, the critical electric field, from Figure 7.14, is approximately $3.7 \times 10^5 \text{ V/cm}$. Then, using Equation (7.61), we find the breakdown voltage as

$$V_B = \frac{\epsilon_s E_{crit}^2}{2eN_B} = \frac{(11.7)(8.85 \times 10^{-14})(3.7 \times 10^5)^2}{2(1.6 \times 10^{-19})(4 \times 10^{15})} = 110 \text{ V}$$

which correlates very well with the results from Figure 7.15.

Conclusion

As Figure 7.15 shows, the breakdown voltage increases as the doping concentration decreases in the low-doped region.

EXERCISE PROBLEM

- Ex 7.7** A one-sided, planar, uniformly doped silicon pn junction diode is required to have a reverse-biased breakdown voltage of $V_B = 60 \text{ V}$. What is the maximum doping concentration in the low-doped region such that this specification is met?
($\epsilon_s = 11.7 \text{ a.u.} \times 8 \approx 90 \text{ a.u.}$)

*7.5 | NONUNIFORMLY DOPED JUNCTIONS

In the pn junctions considered so far, we have assumed that each semiconductor region has been uniformly doped. In actual pn junction structures, this is rarely true. In some electronic applications, specific nonuniform doping profiles are used to obtain special pn junction capacitance characteristics.

7.5.1 Linearly Graded Junctions

If we start with a uniformly doped n-type semiconductor, for example, and diffuse acceptor atoms through the surface, the impurity concentrations will tend to be like those shown in Figure 7.16. The point $x = x'$ on the figure corresponds to the metallurgical junction. The depletion region extends into the p and n regions from the metallurgical junction as we have discussed previously. The net p-type doping concentration near the metallurgical junction may be approximated as a linear function of distance from the metallurgical junction. Likewise, as a first approximation, the net n-type doping concentration is also a linear function of distance extending into the n region from the metallurgical junction. This effective doping profile is referred to as a linearly graded junction.

Figure 7.17 shows the space charge density in the depletion region of the linearly graded junction. For convenience, the metallurgical junction is placed at $x = 0$. The space charge density can be written as

$$\rho(x) = eax \quad (7.62)$$

where a is the gradient of the net impurity concentration.

The electric field and potential in the space charge region can be determined from Poisson's equation. We can write

$$\frac{dE}{dx} = \frac{\rho(x)}{\epsilon_s} = \frac{eax}{\epsilon_s} \quad (7.63)$$

so that the electric field can be found by integration as

$$E = \int \frac{eax}{\epsilon_s} dx = \frac{ea}{2\epsilon_s} (x^2 - x_0^2) \quad (7.64)$$

The electric field in the linearly graded junction is a quadratic function of distance rather than the linear function found in the uniformly doped junction. The maximum electric field again occurs at the metallurgical junction. We may note that the electric field is zero at both $x = +x_0$ and at $x = -x_0$. The electric field in a nonuniformly

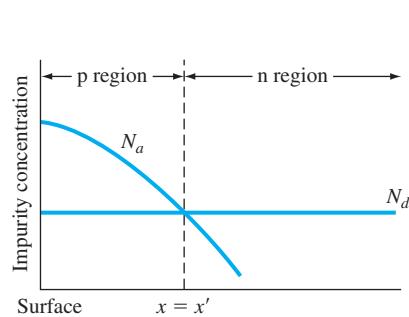


Figure 7.16 | Impurity concentrations of a pn junction with a nonuniformly doped p region.

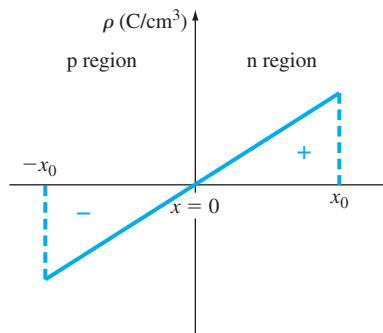


Figure 7.17 | Space charge density in a linearly graded pn junction.

doped semiconductor is not exactly zero, but the magnitude of this field is small, so setting $E = 0$ in the bulk regions is still a good approximation.

The potential is again found by integrating the electric field as

$$\phi(x) = - \int E dx \quad (7.65)$$

If we arbitrarily set $\phi = 0$ at $x = -x_0$, then the potential through the junction is

$$\phi(x) = \frac{-ea}{2\epsilon_s} \left(\frac{x^3}{3} - x_0^2 x \right) + \frac{ea}{3\epsilon_s} x_0^3 \quad (7.66)$$

The magnitude of the potential at $x = +x_0$ will equal the built-in potential barrier for this function. We then have that

$$\phi(x_0) = \frac{2}{3} \cdot \frac{eax_0^3}{\epsilon_s} = V_{bi} \quad (7.67)$$

Another expression for the built-in potential barrier for a linearly graded junction can be approximated from the expression used for a uniformly doped junction. We can write

$$V_{bi} = V_t \ln \left[\frac{N_d(x_0)N_a(-x_0)}{n_i^2} \right] \quad (7.68)$$

where $N_d(x_0)$ and $N_a(-x_0)$ are the doping concentrations at the edges of the space charge region. We can relate these doping concentrations to the gradient, so that

$$N_d(x_0) = ax_0 \quad (7.69a)$$

and

$$N_a(-x_0) = ax_0 \quad (7.69b)$$

Then the built-in potential barrier for the linearly graded junction becomes

$$V_{bi} = V_t \ln \left(\frac{ax_0}{n_i} \right)^2 \quad (7.70)$$

There may be situations in which the doping gradient is not the same on either side of the junction, but we will not consider that condition here.

If a reverse-biased voltage is applied to the junction, the potential barrier increases. The built-in potential barrier V_{bi} in the above equations is then replaced by the total potential barrier $V_{bi} + V_R$. Solving for x_0 from Equation (7.67) and using the total potential barrier, we obtain

$$x_0 = \left\{ \frac{3}{2} \cdot \frac{\epsilon_s}{ea} (V_{bi} + V_R) \right\}^{1/3} \quad (7.71)$$

The junction capacitance per unit area can be determined by the same method that we used for the uniformly doped junction. Figure 7.18 shows the differential charge

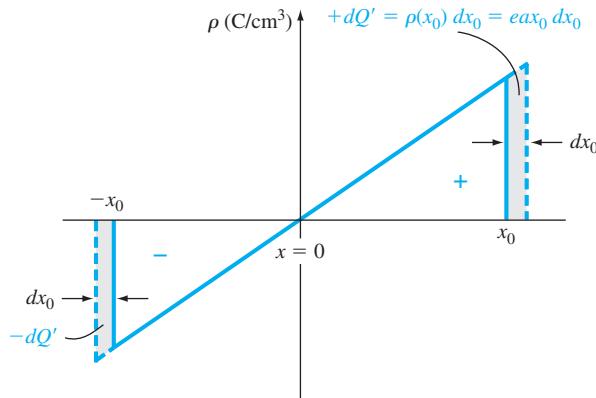


Figure 7.18 | Differential change in space charge width with a differential change in reverse-biased voltage for a linearly graded pn junction.

dQ' , which is uncovered as a differential voltage dV_R is applied. The junction capacitance is then

$$C' = \frac{dQ'}{dV_R} = (eax_0) \frac{dx_0}{dV_R} \quad (7.72)$$

Using Equation (7.71), we obtain¹

$$C' = \left\{ \frac{ea\epsilon_s^2}{12(V_{bi} + V_R)} \right\}^{1/3} \quad (7.73)$$

We may note that C' is proportional to $(V_{bi} + V_R)^{-1/3}$ for the linearly graded junction as compared to $C'\alpha(V_{bi} + V_R)^{-1/2}$ for the uniformly doped junction. In the linearly graded junction, the capacitance is less dependent on reverse-biased voltage than in the uniformly doped junction.

7.5.2 Hyperabrupt Junctions

The uniformly doped junction and linearly graded junction are not the only possible doping profiles. Figure 7.19 shows a generalized one-sided p⁺n junction where the generalized n-type doping concentration for $x > 0$ is given by

$$N = Bx^m \quad (7.74)$$

The case of $m = 0$ corresponds to the uniformly doped junction, and $m = +1$ corresponds to the linearly graded junction just discussed. The cases of $m = +2$ and $m = +3$ shown would approximate a fairly low-doped epitaxial n-type layer grown

¹In a more exact analysis, V_{bi} in Equation (7.73) is replaced by a gradient voltage. However, this analysis is beyond the scope of this text.

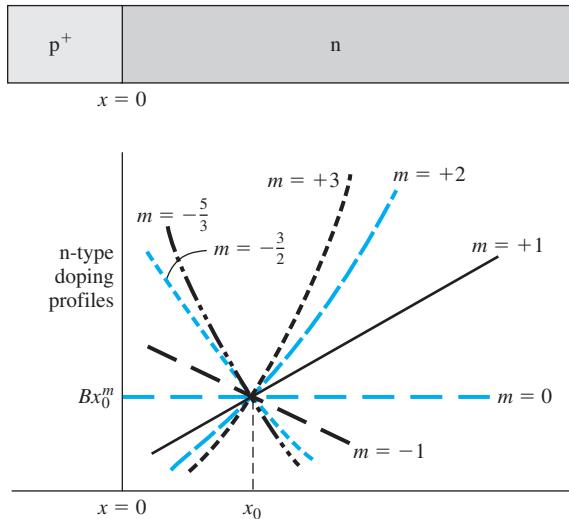


Figure 7.19 | Generalized doping profiles of a one-sided p⁺n junction.
(From Sze [14].)

on a much more heavily doped n⁺ substrate layer. When the value of m is negative, we have what is referred to as a *hyperabrupt junction*. In this case, the n-type doping is larger near the metallurgical junction than in the bulk semiconductor. Equation (7.74) is used to approximate the n-type doping over a small region near $x = x_0$ and does not hold at $x = 0$ when m is negative.

The junction capacitance can be derived using the same analysis method as before and is given by

$$C' = \left\{ \frac{eB\epsilon_s^{(m+1)}}{(m+2)(V_{bi} + V_R)} \right\}^{1/(m+2)} \quad (7.75)$$

When m is negative, the capacitance becomes a very strong function of reverse-biased voltage, a desired characteristic in *varactor diodes*. The term *varactor* comes from the words *variable reactor* and means a device whose reactance can be varied in a controlled manner with bias voltage.

If a varactor diode and an inductance are in parallel, the resonant frequency of the LC circuit is

$$f_r = \frac{1}{2\pi\sqrt{LC}} \quad (7.76)$$

The capacitance of the diode, from Equation (7.75), can be written in the form

$$C = C_0(V_{bi} + V_R)^{-1/(m+2)} \quad (7.77)$$

In a circuit application, we would, in general, like to have the resonant frequency be a linear function of reverse-biased voltage V_R , so we need

$$C \propto V^{-2} \quad (7.78)$$

From Equation (7.77), the parameter m required is found from

$$\frac{1}{m+2} = 2 \quad (7.79a)$$

or

$$m = -\frac{3}{2} \quad (7.79b)$$

A specific doping profile will yield the desired capacitance characteristic.

7.6 | SUMMARY

- A uniformly doped pn junction is initially considered, in which one region of a semiconductor is uniformly doped with acceptor impurities and the adjacent region is uniformly doped with donor impurities.
- A space charge region, or depletion region, is formed on either side of the metallurgical junction separating the n and p regions. This region is essentially depleted of any mobile electrons or holes. A net positive charge density, due to the positively charged donor impurity ions, exists in the n region and a net negative charge density, due to the negatively charged acceptor impurity ions, exists in the p region.
- An electric field exists in the depletion region due to the net space charge density. The direction of the electric field is from the n region to the p region.
- A potential difference exists across the space charge region. Under zero applied bias, this potential difference, known as the built-in potential barrier, maintains thermal equilibrium and holds back majority carrier electrons in the n region and majority carrier holes in the p region.
- An applied reverse-biased voltage (n region positive with respect to the p region) increases the potential barrier, the space charge width, and the magnitude of the electric field.
- As the reverse-biased voltage changes, the amount of charge in the depletion region changes. This change in charge with voltage defines the junction capacitance.
- Avalanche breakdown occurs when a sufficiently large reverse-biased voltage is applied to the pn junction. A large reverse-biased current may then be induced in the pn junction. The breakdown voltage, as a function of the doping concentrations in the pn junction, is derived. In a one-sided pn junction, the breakdown voltage is a function of the doping concentration in the low-doped region.
- The linearly graded junction represents a nonuniformly doped pn junction. Expressions for the electric field, built-in potential barrier, and junction capacitance are derived. The functional relationships differ from those of the uniformly doped junction.
- Specific doping profiles can be used to obtain specific capacitance characteristics. A hyperabrupt junction is one in which the doping decreases away from the metallurgical junction. This type of junction is advantageous in varactor diodes that are used in resonant circuits.

GLOSSARY OF IMPORTANT TERMS

abrupt junction approximation The assumption that there is an abrupt discontinuity in space charge density between the space charge region and the neutral semiconductor region.

avalanche breakdown The process whereby a large reverse-biased pn junction current is created due to the generation of electron–hole pairs by the collision of electrons and/or holes with atomic electrons within the space charge region.

built-in potential barrier The electrostatic potential difference between the p and n regions of a pn junction in thermal equilibrium.

critical electric field The peak electric field in the space charge region at breakdown.

depletion layer capacitance Another term for junction capacitance.

depletion region Another term for space charge region.

hyperabrupt junction A pn junction in which the doping concentration on one side decreases away from the metallurgical junction to achieve a specific capacitance–voltage characteristic.

junction capacitance The capacitance of the pn junction under reverse bias.

linearly graded junction A pn junction in which the doping concentrations on either side of the metallurgical junction are approximated by a linear distribution.

metallurgical junction The interface between the p- and n-doped regions of a pn junction.

one-sided junction A pn junction in which one side of the junction is much more heavily doped than the adjacent side.

reverse bias The condition in which a positive voltage is applied to the n region with respect to the p region of a pn junction so that the potential barrier between the two regions increases above the thermal-equilibrium built-in potential barrier.

space charge region The region on either side of the metallurgical junction in which there is a net charge density due to ionized donors in the n region and ionized acceptors in the p region.

space charge width The width of the space charge region, a function of doping concentrations and applied voltage.

varactor diode A diode whose reactance can be varied in a controlled manner with bias voltage.

CHECKPOINT

After studying this chapter, the reader should have the ability to:

- Describe why and how the space charge region is formed.
- Draw the energy-band diagram of a zero-biased and reverse-biased pn junction.
- Define and derive the expression of the built-in potential barrier voltage.
- Derive the expression for the electric field in space charge region of the pn junction.
- Describe what happens to the parameters of the space charge region when a reverse-biased voltage is applied.
- Define and explain the junction capacitance.
- Describe the characteristics and properties of a one-sided pn junction.
- Describe the avalanche breakdown mechanism in a reverse-biased pn junction.
- Describe how a linearly graded junction is formed.
- Define a hyperabrupt junction.

REVIEW QUESTIONS

1. Define the built-in potential voltage and describe how it maintains thermal equilibrium.
2. Why is an electric field formed in the space charge region? Why is the electric field a linear function of distance in a uniformly doped pn junction?
3. Where does the maximum electric field occur in the space charge region?
4. Why is the space charge width larger in the lower doped side of a pn junction?
5. What is the functional dependence of the space charge width on reverse-biased voltage?
6. Why does the space charge width increase with reverse-biased voltage?
7. Why does a capacitance exist in a reverse-biased pn junction? Why does the capacitance decrease with increasing reverse-biased voltage?
8. What is a one-sided pn junction? What parameters can be determined in a one-sided pn junction?
9. Why does the breakdown voltage of a pn junction decrease as the doping concentration increases?
10. What is a linearly graded junction?
11. What is a hyperabrupt junction and what is one advantage or characteristic of such a junction?

PROBLEMS

Section 7.2 Zero Applied Bias

- 7.1** (a) Calculate V_{bi} in a silicon pn junction at $T = 300$ K for (a) $N_a = 2 \times 10^{15} \text{ cm}^{-3}$ and $N_d = (i) 2 \times 10^{15}$, (ii) 2×10^{16} , and (iii) $2 \times 10^{17} \text{ cm}^{-3}$. (b) Repeat part (a) for $N_a = 2 \times 10^{17} \text{ cm}^{-3}$.
- 7.2** Calculate the built-in potential barrier, V_{bi} , for Si, Ge, and GaAs pn junctions if they each have the following dopant concentrations at $T = 300$ K:
- (a) $N_d = 10^{14} \text{ cm}^{-3}$ $N_a = 10^{17} \text{ cm}^{-3}$
 (b) $N_d = 5 \times 10^{16}$ $N_a = 5 \times 10^{16}$
 (c) $N_d = 10^{17}$ $N_a = 10^{17}$
- 7.3** (a) Plot the built-in potential barrier for a symmetrical ($N_a = N_d$) silicon pn junction at $T = 300$ K over the range $10^{14} \leq N_a = N_d \leq 10^{17} \text{ cm}^{-3}$. (b) Repeat part (a) for a GaAs pn junction. (c) Repeat parts (a) and (b) for $T = 400$ K.
- 7.4** An abrupt silicon pn junction at zero bias has dopant concentrations of $N_a = 10^{17} \text{ cm}^{-3}$ and $N_d = 5 \times 10^{15} \text{ cm}^{-3}$. $T = 300$ K. (a) Calculate the Fermi level on each side of the junction with respect to the intrinsic Fermi level. (b) Sketch the equilibrium energy-band diagram for the junction and determine V_{bi} from the diagram and the results of part (a). (c) Calculate V_{bi} using Equation (7.10), and compare the results to part (b). (d) Determine x_n , x_p , and the peak electric field for this junction.
- 7.5** Repeat problem 7.4 for the case when the doping concentrations are $N_a = N_d = 2 \times 10^{16} \text{ cm}^{-3}$.
- 7.6** A silicon pn junction in thermal equilibrium at $T = 300$ K is doped such that $E_F - E_{Fi} = 0.365 \text{ eV}$ in the n region and $E_{Fi} - E_F = 0.330 \text{ eV}$ in the p region.

(a) Sketch the energy-band diagram for the pn junction. (b) Find the impurity doping concentration in each region. (c) Determine V_{bi} .

- 7.7** Consider a uniformly doped GaAs pn junction with doping concentrations of $N_a = 2 \times 10^{15} \text{ cm}^{-3}$ and $N_d = 4 \times 10^{16} \text{ cm}^{-3}$. Plot the built-in potential barrier V_{bi} versus temperature over the range $200 \leq T \leq 400 \text{ K}$.
- 7.8** (a) Consider a uniformly doped silicon pn junction at $T = 300 \text{ K}$. At zero bias, 25 percent of the total space charge region is in the n-region. The built-in potential barrier is $V_{bi} = 0.710 \text{ V}$. Determine (i) N_a , (ii) N_d , (iii) x_n , (iv) x_p , and (v) $|E_{\max}|$. (b) Repeat part (a) for a GaAs pn junction with $V_{bi} = 1.180 \text{ V}$.
- 7.9** Consider the impurity doping profile shown in Figure P7.9 in a silicon pn junction. For zero applied voltage, (a) determine V_{bi} , (b) calculate x_n and x_p , (c) sketch the thermal equilibrium energy-band diagram, and (d) plot the electric field versus distance through the junction.
- 7.10** Consider a uniformly doped silicon pn junction with doping concentrations $N_a = 2 \times 10^{17} \text{ cm}^{-3}$ and $N_d = 4 \times 10^{16} \text{ cm}^{-3}$. (a) Determine V_{bi} at $T = 300 \text{ K}$. (b) Determine the temperature at which V_{bi} increases by 2 percent. (Trial and error may have to be used.)
- 7.11** The doping concentrations in a uniformly doped silicon pn junction are $N_a = 4 \times 10^{16} \text{ cm}^{-3}$ and $N_d = 2 \times 10^{15} \text{ cm}^{-3}$. The measured built-in potential barrier is $V_{bi} = 0.550 \text{ V}$. Determine the temperature at which this result occurs.
- 7.12** An “isotype” step junction is one in which the same impurity type doping changes from one concentration value to another value. An n-n isotype doping profile is shown in Figure P7.12. (a) Sketch the thermal equilibrium energy-band diagram of the isotype junction. (b) Using the energy-band diagram, determine the built-in potential barrier. (c) Discuss the charge distribution through the junction.
- 7.13** A particular type of junction is an n region adjacent to an intrinsic region. This junction can be modeled as an n-type region to a lightly doped p-type region. Assume the doping concentrations in silicon at $T = 300 \text{ K}$ are $N_d = 10^{16} \text{ cm}^{-3}$ and $N_a = 10^{12} \text{ cm}^{-3}$. For zero applied bias, determine (a) V_{bi} , (b) x_n , (c) x_p , and (d) $|E_{\max}|$. Sketch the electric field versus distance through the junction.
- 7.14** We are assuming an abrupt depletion approximation for the space charge region. That is, no free carriers exist within the depletion region and the semiconductor abruptly changes to a neutral region outside the space charge region. This approximation is

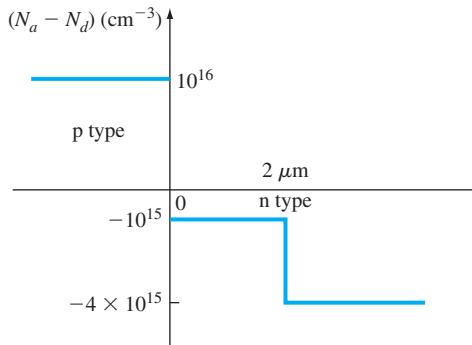


Figure P7.9 | Figure for Problem 7.9.

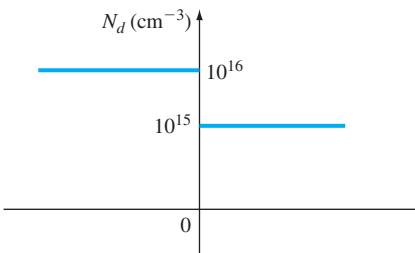


Figure P7.12 | Figure for Problem 7.12.

adequate for most applications, but the abrupt transition does not exist. The space charge region changes over a distance of a few Debye lengths, where the Debye length in the n region is given by

$$L_D = \left[\frac{\epsilon_s k T}{e^2 N_d} \right]^{1/2}$$

Calculate L_D and find the ratio of L_D/x_n for the following conditions. The p-type doping concentration is $N_a = 8 \times 10^{17} \text{ cm}^{-3}$ and the n-type doping concentration is (a) $N_d = 8 \times 10^{14} \text{ cm}^{-3}$, (b) $N_d = 2.2 \times 10^{16} \text{ cm}^{-3}$, and (c) $N_d = 8 \times 10^{17} \text{ cm}^{-3}$.

- 7.15** Examine the electric field versus distance through a uniformly doped silicon pn junction at $T = 300 \text{ K}$ as a function of doping concentrations. Assume zero applied bias. Sketch the electric field versus distance through the space charge region and calculate $|E_{\max}|$ for: (a) $N_a = 10^{17} \text{ cm}^{-3}$ and $10^{14} \leq N_d \leq 10^{17} \text{ cm}^{-3}$ and (b) $N_a = 10^{14} \text{ cm}^{-3}$ and $10^{14} \leq N_d \leq 10^{17} \text{ cm}^{-3}$. (c) What can be said about the results for $N_d \geq 100 N_a$ or $N_a \geq 100 N_d$?

Section 7.3 Reverse Applied Bias

- 7.16** An abrupt silicon pn junction at $T = 300 \text{ K}$ has impurity doping concentrations of $N_a = 5 \times 10^{16} \text{ cm}^{-3}$ and $N_d = 10^{15} \text{ cm}^{-3}$. Calculate (a) V_{bi} , (b) W at (i) $V_R = 0$ and (ii) $V_R = 5 \text{ V}$, and (c) $|E_{\max}|$ at (i) $V_R = 0$ and (ii) $V_R = 5 \text{ V}$.
- 7.17** Consider the pn junction described in Problem 7.10 for $T = 300 \text{ K}$. The cross-sectional area of the junction is $2 \times 10^{-4} \text{ cm}^2$ and the applied reverse-biased voltage is $V_R = 2.5 \text{ V}$. Calculate (a) V_{bi} , (b) x_n, x_p, W , (c) $|E_{\max}|$, and (d) the junction capacitance.
- 7.18** An ideal one-sided silicon p⁺n junction at $T = 300 \text{ K}$ is uniformly doped on both sides of the metallurgical junction. It is found that the doping relation is $N_a = 80 N_d$ and the built-in potential barrier is $V_{bi} = 0.740 \text{ V}$. A reverse-biased voltage of $V_R = 10 \text{ V}$ is applied. Determine (a) N_a, N_d ; (b) x_p, x_n ; (c) $|E_{\max}|$; and (d) C'_j .
- 7.19** A silicon n⁺p junction is biased at $V_R = 5 \text{ V}$. (a) Determine the change in built-in potential barrier if the doping concentration in the p region increases by a factor of 3. (b) Determine the ratio of junction capacitance when the acceptor doping is $3N_a$ compared to that when the acceptor doping is N_a . (c) Why does the junction capacitance increase when the doping concentration increases?
- 7.20** (a) The peak electric field in a reverse-biased silicon pn junction is $|E_{\max}| = 3 \times 10^5 \text{ V/cm}$. The doping concentrations are $N_d = 4 \times 10^{15} \text{ cm}^{-3}$ and $N_a = 4 \times 10^{17} \text{ cm}^{-3}$. Find the magnitude of the reverse-biased voltage. (b) Repeat part (a) for $N_d = 4 \times 10^{16} \text{ cm}^{-3}$ and $N_a = 4 \times 10^{17} \text{ cm}^{-3}$. (c) Repeat part (a) for $N_d = N_a = 4 \times 10^{17} \text{ cm}^{-3}$.
- 7.21** Consider two p⁺n silicon junctions at $T = 300 \text{ K}$ reverse biased at $V_R = 5 \text{ V}$. The impurity doping concentrations in junction A are $N_a = 10^{18} \text{ cm}^{-3}$ and $N_d = 10^{15} \text{ cm}^{-3}$, and those in junction B are $N_a = 10^{18} \text{ cm}^{-3}$ and $N_d = 10^{16} \text{ cm}^{-3}$. Calculate the ratio of the following parameters for junction A to junction B: (a) W, (b) $|E_{\max}|$, and (c) C'_j .
- 7.22** Consider a uniformly doped GaAs pn junction at $T = 300 \text{ K}$. The junction capacitance at zero bias is $C_j(0)$ and the junction capacitance with a 10-V reverse-biased voltage is $C_j(10)$. The ratio of the capacitances is

$$\frac{C_j(0)}{C_j(10)} = 3.13$$

Also under reverse bias, the space charge width into the p region is 0.2 of the total space charge width. Determine (a) V_{bi} and (b) N_a, N_d .

- 7.23** A GaAs pn junction at $T = 300$ K has impurity doping concentrations of $N_a = 2 \times 10^{16}$ cm $^{-3}$ and $N_d = 5 \times 10^{15}$ cm $^{-3}$. It is determined that the ratio of junction capacitance at two reverse-biased voltages is $C'_j(V_{R1})/C'_j(V_{R2}) = 1.5$, where $V_{R1} = 0.5$ V. Determine V_{R2} .
- 7.24** (a) The impurity doping concentrations in a silicon pn junction at $T = 300$ K are $N_a = 2 \times 10^{15}$ cm $^{-3}$ and $N_d = 4 \times 10^{16}$ cm $^{-3}$. The cross-sectional area of the pn junction is 5×10^{-4} cm 2 . Determine the junction capacitance at (i) $V_R = 0$ and (ii) $V_R = 5$ V. (b) Repeat part (a) for a GaAs pn junction.
- 7.25** An abrupt silicon pn junction at $T = 300$ K is uniformly doped with $N_a = 2 \times 10^{17}$ cm $^{-3}$ and $N_d = 5 \times 10^{15}$ cm $^{-3}$. The cross-sectional area of the pn junction is 8×10^{-4} cm 2 . An inductance is placed in parallel with the pn junction. (a) With a reverse-biased voltage of $V_R = 10$ V applied to the pn junction, the resonant frequency of the circuit is $f = 1.25$ MHz. What is the value of the inductance? (b) Using the results of part (a), what is the resonant frequency if the reverse-biased voltage is (i) $V_R = 1$ V and (ii) $V_R = 5$ V?
- 7.26** (a) A uniformly doped silicon p⁺n junction at $T = 300$ K is to be designed such that, at a reverse-biased of $V_R = 10$ V, the maximum electric field is limited to $|E_{\max}| = 2.5 \times 10^5$ V/cm. Determine the maximum doping concentration in the n region. (Use an approximate value for V_{bi}). (b) Repeat part (a) if the maximum electric field is limited to $|E_{\max}| = 10^5$ V/cm.
- 7.27** (a) A GaAs pn junction at $T = 300$ K, with a cross-sectional area of 10^{-4} cm 2 , is to be designed that meets the following specifications. At a reverse-biased voltage of $V_R = 2$ V, 20 percent of the total space charge width is to be in the p region and the total junction capacitance is to be 0.6 pF. Determine N_a , N_d , and W . (b) Repeat part (a) if $V_R = 5$ V.
- 7.28** A silicon pn junction at $T = 300$ K has the doping profile shown in Figure P7.28. Calculate (a) V_{bi} , (b) x_n and x_p at zero bias, and (c) the applied bias required so that $x_n = 30$ μ m.
- 7.29** Consider a silicon pn junction with the doping profile shown in Figure P7.29. $T = 300$ K. (a) Calculate the applied reverse-biased voltage required so that the space charge region extends entirely through the p region. (b) Determine the space charge width into the n⁺ region with the reverse-biased voltage calculated in part (a). (c) Calculate the peak electric field for this applied voltage.

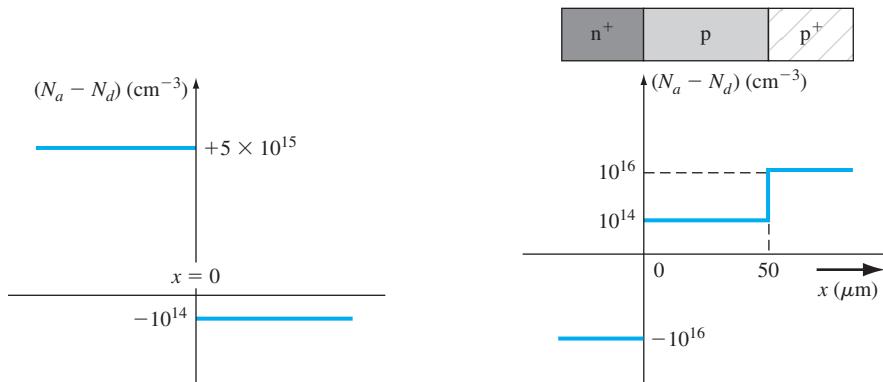


Figure P7.28 | Figure for Problem 7.28.

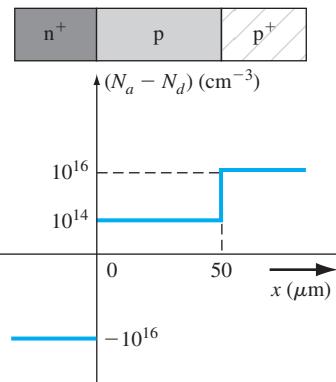
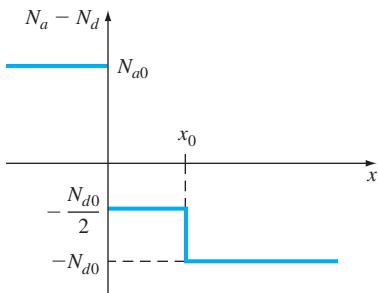
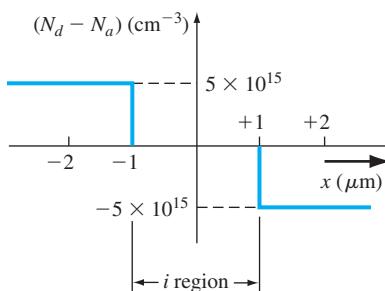


Figure P7.29 | Figure for Problem 7.29.

**Figure P7.33** | Figure for Problem 7.33.**Figure P7.34** | Figure for Problem 7.34.

- 7.30** A silicon p⁺n junction has doping concentrations of $N_a = 2 \times 10^{17} \text{ cm}^{-3}$ and $N_d = 2 \times 10^{15} \text{ cm}^{-3}$. The cross-sectional area is 10^{-5} cm^2 . Calculate (a) V_{bi} and (b) the junction capacitance at (i) $V_R = 1 \text{ V}$, (ii) $V_R = 3 \text{ V}$, and (iii) $V_R = 5 \text{ V}$. (c) Plot $1/C^2$ versus V_R and show that the slope can be used to find N_d and the intercept at the voltage axis yields V_{bi} .
- 7.31** The total junction capacitance of a GaAs pn junction at $T = 300 \text{ K}$ is found to be 1.10 pF at $V_R = 1 \text{ V}$. The doping concentration in one region is measured and found to be $8 \times 10^{16} \text{ cm}^{-3}$, and the built-in potential is found to be $V_{bi} = 1.20 \text{ V}$. Determine (a) the doping in the other region of the pn junction and (b) the cross-sectional area. (c) The reverse-biased voltage is changed and the capacitance is found to be 0.80 pF . What is the value of V_R ?
- 7.32** Examine how the capacitance C' and the function $(1/C')^2$ vary with reverse-biased voltage V_R as the doping concentrations change. In particular, consider these plots versus N_a for $N_a \geq 100 N_d$ and versus N_d for $N_d \geq 100 N_a$.
- *7.33** A pn junction has the doping profile shown in Figure P7.33. Assume that $x_n > x_0$ for all reverse-biased voltages. (a) What is the built-in potential across the junction? (b) For the abrupt junction approximation, sketch the charge density through the junction. (c) Derive the expression for the electric field through the space charge region.
- *7.34** A silicon PIN junction has the doping profile shown in Figure P7.34. The ‘I’ corresponds to an ideal intrinsic region in which there is no impurity doping concentration. A reverse-biased voltage is applied to the PIN junction so that the total depletion width extends from $-2 \mu\text{m}$ to $+2 \mu\text{m}$. (a) Using Poisson’s equation, calculate the magnitude of the electric field at $x = 0$. (b) Sketch the electric field through the PIN junction. (c) Calculate the reverse-biased voltage that must be applied.

Section 7.4 Junction Breakdown

- 7.35** Consider a silicon n⁺p junction diode. The critical electric field for breakdown in silicon is approximately $E_{crit} = 4 \times 10^5 \text{ V/cm}$. Determine the maximum p-type doping concentration such that the breakdown voltage is (a) 40 V and (b) 20 V .
- 7.36** Design an abrupt silicon n⁺p junction diode that has a reverse breakdown voltage of 80 V .
- 7.37** (a) The n-type doping concentration in an abrupt p⁺n GaAs junction diode is $N_d = 10^{16} \text{ cm}^{-3}$. Determine the breakdown voltage. (b) Repeat part (a) for $N_d = 10^{15} \text{ cm}^{-3}$.

*Asterisks next to problems indicate problems that are more difficult.

- 7.38** (a) A symmetrically doped silicon pn junction diode has doping concentrations of $N_a = N_d = 2 \times 10^{16} \text{ cm}^{-3}$. Assuming the critical electric field is $E_{crit} = 4 \times 10^5 \text{ V/cm}$, determine the breakdown voltage. (b) Repeat part (a) if the doping concentrations are $N_a = N_d = 5 \times 10^{15} \text{ cm}^{-3}$.
- 7.39** An abrupt silicon p⁺n junction has an n-region doping concentration of $N_d = 5 \times 10^{15} \text{ cm}^{-3}$. What must be the minimum n-region width such that avalanche breakdown occurs before the depletion region reaches an ohmic contact (punchthrough)?
- 7.40** A silicon pn junction diode is doped with $N_a = N_d = 10^{18} \text{ cm}^{-3}$. Zener breakdown occurs when the peak electric field reaches 10^6 V/cm . Determine the reverse-biased breakdown voltage.
- 7.41** A diode will very often have the doping profile shown in Figure P7.29, which is known as an n⁺pp⁺ diode. Under reverse bias, the depletion region must remain within the p region to avoid premature breakdown. Assume the p-region doping is 10^{15} cm^{-3} . Determine the reverse-biased voltage such that the depletion region remains within the p region and does not reach breakdown if the p-region width is (a) 75 μm and (b) 150 μm . For each case, state whether the maximum depletion width or the breakdown voltage is reached first.
- 7.42** Consider a silicon pn junction at $T = 300 \text{ K}$ whose doping profile varies linearly from $N_a = 10^{18} \text{ cm}^{-3}$ to $N_d = 10^{18} \text{ cm}^{-3}$ over a distance of 2 μm . Estimate the breakdown voltage.

Section 7.5 Nonuniformly Doped Junctions

- 7.43** Consider a linearly graded junction. (a) Starting with Equation (7.62), derive the expression for the electric field given in Equation (7.64). (b) Derive the expression for the potential through the space charge region given by Equation (7.66).
- 7.44** The built-in potential barrier of a linearly graded silicon pn junction at $T = 300 \text{ K}$ is $V_{bi} = 0.70 \text{ V}$. The junction capacitance measured at $V_R = 3.5 \text{ V}$ is $C' = 7.2 \times 10^{-9} \text{ F/cm}^2$. Find the gradient, a , of the net impurity concentration.

Summary and Review

- 7.45** (a) A one-sided silicon n⁺p junction at $T = 300 \text{ K}$ is doped at $N_d = 3 \times 10^{17} \text{ cm}^{-3}$. Design the junction such that $C_j = 0.45 \text{ pF}$ at $V_R = 5 \text{ V}$. (b) Calculate the junction capacitance at (i) $V_R = 2.5 \text{ V}$ and (ii) $V_R = 0 \text{ V}$.
- 7.46** A one-sided p⁺n junction with a cross-sectional area of 10^{-5} cm^2 has a measured built-in potential of $V_{bi} = 0.8 \text{ V}$ at $T = 300 \text{ K}$. A plot of $(1/C_j)^2$ versus V_R is linear for $V_R < 1 \text{ V}$ and is essentially constant for $V_R > 1 \text{ V}$. The capacitance is $C_j = 0.082 \text{ pF}$ at $V_R = 1 \text{ V}$. Determine the doping concentrations on either side of the metallurgical junction that will produce this capacitance characteristic.
- *7.47** Silicon, at $T = 300 \text{ K}$, is doped at $N_{d1} = 10^{15} \text{ cm}^{-3}$ for $x < 0$ and $N_{d2} = 5 \times 10^{16} \text{ cm}^{-3}$ for $x > 0$ to form an n – n step junction. (a) Sketch the energy-band diagram. (b) Derive an expression for V_{bi} . (c) Sketch the charge density, electric field, and potential through the junction. (d) Explain where the charge density came from and is located.
- *7.48** A diffused silicon pn junction has a linearly graded junction on the p side with $a = 2 \times 10^{19} \text{ cm}^{-4}$, and a uniform doping of 10^{15} cm^{-3} on the n side. (a) If the depletion width on the p side is 0.7 μm at zero bias, find the total depletion width, built-in potential, and maximum electric field at zero bias. (b) Plot the potential function through the junction.

READING LIST

1. Dimitrijev, S. *Principles of Semiconductor Devices*. New York: Oxford University Press, 2006.
2. Kano, K. *Semiconductor Devices*. Upper Saddle River, NJ: Prentice Hall, 1998.
- *3. Li, S. S. *Semiconductor Physical Electronics*. New York: Plenum Press, 1993.
4. Muller, R. S., and T. I. Kamins. *Device Electronics for Integrated Circuits*, 2nd ed. New York: John Wiley and Sons, 1986.
5. Navon, D. H. *Semiconductor Microdevices and Materials*. New York: Holt, Rinehart & Winston, 1986.
6. Neudeck, G. W. *The PN Junction Diode*. Vol. 2 of the *Modular Series on Solid State Devices*, 2nd ed. Reading, MA: Addison-Wesley, 1989.
- *7. Ng, K. K. *Complete Guide to Semiconductor Devices*. New York: McGraw-Hill, 1995.
8. Pierret, R. F. *Semiconductor Device Fundamentals*. Reading, MA: Addison-Wesley, 1996.
- *9. Roulston, D. J. *An Introduction to the Physics of Semiconductor Devices*. New York: Oxford University Press, 1999.
10. Shur, M. *Introduction to Electronic Devices*. New York: John Wiley and Sons, 1996.
- *11. Shur, M. *Physics of Semiconductor Devices*. Englewood Cliffs, NJ: Prentice Hall, 1990.
12. Singh, J. *Semiconductor Devices: Basic Principles*. New York: John Wiley and Sons, 2001.
13. Streetman, B. G., and S. K. Banerjee. *Solid State Electronic Devices*, 6th ed. Upper Saddle River, NJ: Pearson Prentice Hall, 2006.
14. Sze, S. M., and K. K. Ng. *Physics of Semiconductor Devices*, 3rd ed. Hoboken, NJ: John Wiley and Sons, 2007.
15. Sze, S. M. *Semiconductor Devices: Physics and Technology*, 2nd ed. New York: John Wiley and Sons, 2001.
- *16. Wang, S. *Fundamentals of Semiconductor Theory and Device Physics*. Englewood Cliffs, NJ: Prentice Hall, 1989.
17. Yang, E. S. *Microelectronic Devices*. New York: McGraw-Hill, 1988.

*Indicates references that are at an advanced level compared to this text.

8

The pn Junction Diode

In the last chapter, we discussed the electrostatics of the pn junction in thermal equilibrium and under reverse bias. We determined the built-in potential barrier at thermal equilibrium and calculated the electric field in the space charge region. We also considered the junction capacitance.

In this chapter, we consider the pn junction with a forward-bias voltage applied and determine the current–voltage characteristics. The potential barrier of the pn junction is lowered when a forward-bias voltage is applied, allowing electrons and holes to flow across the space charge region. When holes flow from the p region across the space charge region into the n region, they become excess minority carrier holes and are subject to the excess minority carrier diffusion, drift, and recombination processes discussed in Chapter 6. Likewise, when electrons from the n region flow across the space charge region into the p region, they become excess minority carrier electrons and are subject to these same processes. ■

8.0 | PREVIEW

In this chapter, we will:

- Consider the process by which the potential barrier of a pn junction is lowered when a forward-bias voltage is applied, so holes and electrons can flow across the junction generating a diode current.
- Derive the boundary conditions for excess holes in the n region and excess electrons in the p region, and analyze the behavior of these excess carriers under a forward bias.
- Derive the ideal current–voltage relation of the forward-biased pn junction diode.
- Describe and analyze nonideal effects in the pn junction diode such as high-level injection, and generation and recombination currents.
- Develop a small-signal equivalent circuit of the pn junction diode. This equivalent circuit is used to relate small time-varying currents and voltages in the pn junction.

- Discuss large signal diode switching characteristics.
- Describe a specialized pn junction called a tunnel diode.

8.1 | pn JUNCTION CURRENT

When a forward-bias voltage is applied to a pn junction, a current will be induced in the device. We initially consider a qualitative discussion of how charges flow in the pn junction and then consider the mathematical derivation of the current–voltage relationship.

8.1.1 Qualitative Description of Charge Flow in a pn Junction

We can qualitatively understand the mechanism of the current in a pn junction by again considering the energy-band diagrams. Figure 8.1a shows the energy-band diagram of a pn junction in thermal equilibrium that was developed in the last chapter. We argued that the potential barrier seen by the electrons, for example, holds back the large concentration of electrons in the n region and keeps them from flowing into the p region. Similarly, the potential barrier seen by the holes holds back the large concentration of holes in the p region and keeps them from flowing into the n region. The potential barrier, then, maintains thermal equilibrium.

Figure 8.1b shows the energy-band diagram of a reverse-biased pn junction. The potential of the n region is positive with respect to the p region so the Fermi energy

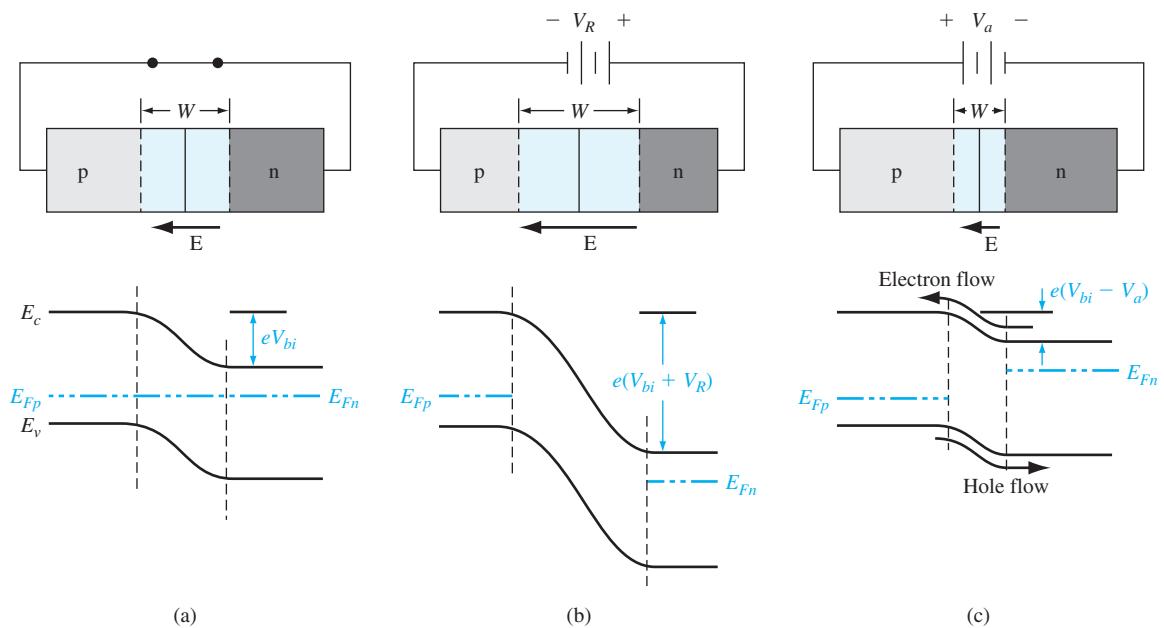


Figure 8.1 | A pn junction and its associated energy-band diagram for (a) zero bias, (b) reverse bias, and (c) forward bias.

in the n region is lower than that in the p region. The total potential barrier is now larger than that for the zero-bias case. We argued in the last chapter that the increased potential barrier continues to hold back the electrons and holes so that there is still essentially no charge flow and hence essentially no current.

Figure 8.1c shows the energy-band diagram for the case when a positive voltage is applied to the p region with respect to the n region. The Fermi level in the p region is now lower than that in the n region. The total potential barrier is now reduced. The smaller potential barrier means that the electric field in the depletion region is also reduced. The smaller electric field means that the electrons and holes are no longer held back in the n and p regions, respectively. There will be a diffusion of holes from the p region across the space charge region where they will flow into the n region. Similarly, there will be a diffusion of electrons from the n region across the space charge region where they will flow into the p region. The flow of charge generates a current through the pn junction.

The injection of holes into the n region means that these holes are minority carriers. Likewise, the injection of electrons into the p region means that these electrons are minority carriers. The behavior of these minority carriers is described by the ambipolar transport equations that were discussed in Chapter 6. There will be diffusion as well as recombination of excess carriers in these regions. The diffusion of carriers implies that there will be diffusion currents. The mathematical derivation of the pn junction current–voltage relationship is considered in the next section.

8.1.2 Ideal Current–Voltage Relationship

The ideal current–voltage relationship of a pn junction is derived on the basis of four assumptions. (The last assumption has three parts, but each part deals with current.) They are:

1. The abrupt depletion layer approximation applies. The space charge regions have abrupt boundaries, and the semiconductor is neutral outside of the depletion region.
2. The Maxwell–Boltzmann approximation applies to carrier statistics.
3. The concepts of low injection and complete ionization apply.
- 4a. The total current is a constant throughout the entire pn structure.
- 4b. The individual electron and hole currents are continuous functions through the pn structure.
- 4c. The individual electron and hole currents are constant throughout the depletion region.

Notation can sometimes appear to be overwhelming in the equations in this chapter. Table 8.1 lists some of the various electron and hole concentration terms that appear. Many terms have already been used in previous chapters but are repeated here for convenience.

Table 8.1 | Commonly used terms and notation for this chapter

Term	Meaning
N_a	Acceptor concentration in the p region of the pn junction
N_d	Donor concentration in the n region of the pn junction
$n_{n0} = N_d$	Thermal-equilibrium majority carrier electron concentration in the n region
$p_{p0} = N_a$	Thermal-equilibrium majority carrier hole concentration in the p region
$n_{p0} = n_i^2/N_a$	Thermal-equilibrium minority carrier electron concentration in the p region
$p_{n0} = n_i^2/N_d$	Thermal-equilibrium minority carrier hole concentration in the n region
n_p	Total minority carrier electron concentration in the p region
p_n	Total minority carrier hole concentration in the n region
$n_p(-x_p)$	Minority carrier electron concentration in the p region at the space charge edge
$p_n(x_n)$	Minority carrier hole concentration in the n region at the space charge edge
$\delta n_p = n_p - n_{p0}$	Excess minority carrier electron concentration in the p region
$\delta p_n = p_n - p_{n0}$	Excess minority carrier hole concentration in the n region

8.1.3 Boundary Conditions

Figure 8.2 shows the conduction-band energy through the pn junction in thermal equilibrium. The n region contains many more electrons in the conduction band than the p region; the built-in potential barrier prevents this large density of electrons from flowing into the p region. The built-in potential barrier maintains equilibrium between the carrier distributions on either side of the junction.

An expression for the built-in potential barrier was derived in the last chapter and was given by Equation (7.10) as

$$V_{bi} = V_t \ln\left(\frac{N_a N_d}{n_i^2}\right)$$

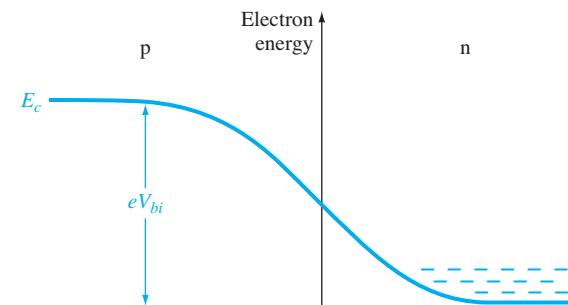


Figure 8.2 | Conduction-band energy through a pn junction.

If we divide the equation by $V_t = kT/e$, take the exponential of both sides, and then take the reciprocal, we obtain

$$\frac{n_i^2}{N_a N_d} = \exp\left(\frac{-eV_{bi}}{kT}\right) \quad (8.1)$$

If we assume complete ionization, we can write

$$n_{n0} \approx N_d \quad (8.2)$$

where n_{n0} is the thermal-equilibrium concentration of majority carrier electrons in the n region. In the p region, we can write

$$n_{p0} \approx \frac{n_i^2}{N_a} \quad (8.3)$$

where n_{p0} is the thermal-equilibrium concentration of minority carrier electrons. Substituting Equations (8.2) and (8.3) into Equation (8.1), we obtain

$$n_{p0} = n_{n0} \exp\left(\frac{-eV_{bi}}{kT}\right) \quad (8.4)$$

This equation relates the minority carrier electron concentration on the p side of the junction to the majority carrier electron concentration on the n side of the junction in thermal equilibrium.

If a positive voltage is applied to the p region with respect to the n region, the potential barrier is reduced. Figure 8.3a shows a pn junction with an applied voltage V_a . The electric field in the bulk p and n regions is normally very small. Essentially all of the applied voltage is across the junction region. The electric field E_{app} induced by the applied voltage is in the opposite direction to the thermal-equilibrium space charge electric field, so the net electric field in the space charge region is reduced below the equilibrium value. The delicate balance between diffusion and the E-field

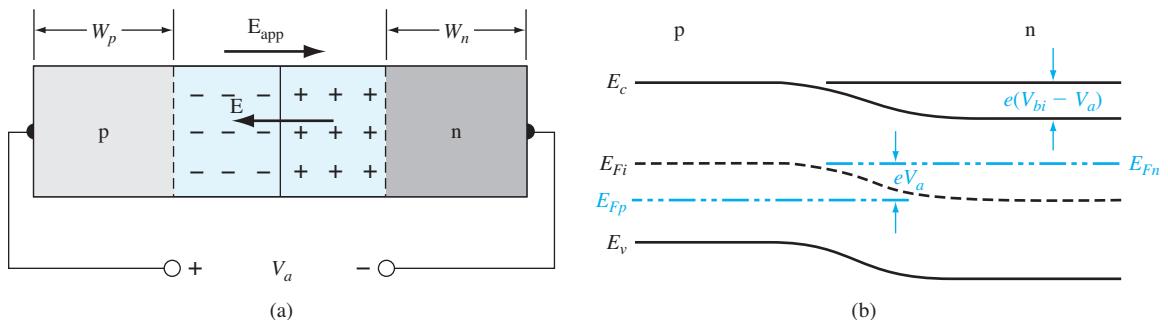


Figure 8.3 | (a) A pn junction with an applied forward-bias voltage showing the directions of the electric field induced by V_a and the space charge electric field. (b) Energy-band diagram of the forward-biased pn junction.

force achieved at thermal equilibrium is upset. The electric field force that prevented majority carriers from crossing the space charge region is reduced; majority carrier electrons from the n side are now injected across the depletion region into the p material, and majority carrier holes from the p side are injected across the depletion region into the n material. As long as the bias V_a is applied, the injection of carriers across the space charge region continues and a current is created in the pn junction. This bias condition is known as forward bias; the energy-band diagram of the forward-biased pn junction is shown in Figure 8.3b.

The potential barrier V_{bi} in Equation (8.4) can be replaced by $(V_{bi} - V_a)$ when the junction is forward biased. Equation (8.4) becomes

$$n_p = n_{p0} \exp\left(\frac{-e(V_{bi} - V_a)}{kT}\right) = n_{p0} \exp\left(\frac{-eV_{bi}}{kT}\right) \exp\left(\frac{+eV_a}{kT}\right) \quad (8.5)$$

If we assume low injection, the majority carrier electron concentration n_{n0} , for example, does not change significantly. However, the minority carrier concentration, n_p , can deviate from its thermal-equilibrium value n_{p0} by orders of magnitude. Using Equation (8.4), we can write Equation (8.5) as

$$n_p = n_{p0} \exp\left(\frac{eV_a}{kT}\right) \quad (8.6)$$

When a forward-bias voltage is applied to the pn junction, the junction is no longer in thermal equilibrium. The left side of Equation (8.6) is the total minority carrier electron concentration in the p region, which is now greater than the thermal equilibrium value. The forward-bias voltage lowers the potential barrier so that majority carrier electrons from the n region are injected across the junction into the p region, thereby increasing the minority carrier electron concentration. We have produced excess minority carrier electrons in the p region.

When the electrons are injected into the p region, these excess carriers are subject to the diffusion and recombination processes we discussed in Chapter 6. Equation (8.6), then, is the expression for the minority carrier electron concentration at the edge of the space charge region in the p region.

Exactly the same process occurs for majority carrier holes in the p region, which are injected across the space charge region into the n region under a forward-bias voltage. We can write that

$$p_n = p_{n0} \exp\left(\frac{eV_a}{kT}\right) \quad (8.7)$$

where p_n is the concentration of minority carrier holes at the edge of the space charge region in the n region. Figure 8.4 shows these results. By applying a forward-bias voltage, we create excess minority carriers in each region of the pn junction.

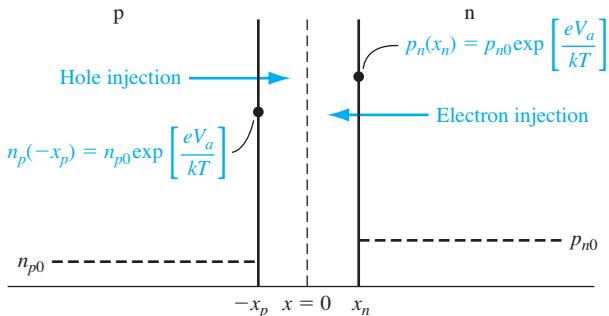


Figure 8.4 | Excess minority carrier concentrations at the space charge edges generated by the forward-bias voltage.

EXAMPLE 8.1

Objective: Calculate the **minority carrier concentrations at the edge** of the space charge regions in a forward-biased pn junction.

Consider a silicon pn junction at $T = 300$ K. Assume the doping concentration in the n region is $N_d = 10^{16} \text{ cm}^{-3}$ and the doping concentration in the p region is $N_a = 6 \times 10^{15} \text{ cm}^{-3}$, and assume that a forward bias of 0.60 V is applied to the pn junction.

■ Solution

From Equations (8.6) and (8.7) and from Figure 8.4, we have

$$n_p(-x_p) = n_{p0} \exp\left(\frac{eV_a}{kT}\right) \quad \text{and} \quad p_n(x_n) = p_{n0} \exp\left(\frac{eV_a}{kT}\right)$$

The thermal-equilibrium minority carrier concentrations are

$$n_{p0} = \frac{n_i^2}{N_a} = \frac{(1.5 \times 10^{10})^2}{6 \times 10^{15}} = 3.75 \times 10^4 \text{ cm}^{-3}$$

and

$$p_{n0} = \frac{n_i^2}{N_d} = \frac{(1.5 \times 10^{10})^2}{10^{16}} = 2.25 \times 10^4 \text{ cm}^{-3}$$

We then have

$$n_p(-x_p) = 3.75 \times 10^4 \exp\left(\frac{0.60}{0.0259}\right) = 4.31 \times 10^{14} \text{ cm}^{-3}$$

and

$$p_n(x_n) = 2.25 \times 10^4 \exp\left(\frac{0.60}{0.0259}\right) = 2.59 \times 10^{14} \text{ cm}^{-3}$$

■ Comment

The minority carrier concentrations can increase by many orders of magnitude when a relatively small forward-bias voltage is applied. Low injection still applies, however, since the excess minority carrier concentrations at the space-charge edges are much less than the thermal-equilibrium majority carrier concentrations.

■ EXERCISE PROBLEM

Ex 8.1 A silicon pn junction at $T = 300$ K is doped with impurity concentrations of $N_d = 2 \times 10^{16} \text{ cm}^{-3}$ and $N_a = 5 \times 10^{16} \text{ cm}^{-3}$. The junction is forward biased at $V_a = 0.650$ V. Determine the minority carrier concentrations at the space charge edges. Does low injection still apply?

$$[\text{Ans. } n_p(-x) = 3.57 \times 10^{14} \text{ cm}^{-3}, n_p(x) = 8.92 \times 10^{14} \text{ cm}^{-3}, \text{ yes}]$$

The minority carrier concentrations at the space charge edges, given by Equations (8.6) and (8.7), were derived assuming that a forward-bias voltage ($V_a > 0$) was applied across the pn junction. However, nothing in the derivation prevents V_a from being negative (reverse bias). If a reverse-biased voltage greater than a few tenths of a volt is applied to the pn junction, then we see from Equations (8.6) and (8.7) that the minority carrier concentrations at the space charge edge are essentially zero. The minority carrier concentrations for the reverse-biased condition drop below the thermal-equilibrium values.

8.1.4 Minority Carrier Distribution

We developed, in Chapter 6, the ambipolar transport equation for excess minority carrier holes in an n region. This equation, in one dimension, is

$$D_p \frac{\partial^2 (\delta p_n)}{\partial x^2} - \mu_p E \frac{\partial (\delta p_n)}{\partial x} + g' - \frac{\delta p_n}{\tau_{p0}} = \frac{\partial (\delta p_n)}{\partial t} \quad (8.8)$$

where $\delta p_n = p_n - p_{n0}$ is the excess minority carrier hole concentration and is the difference between the total and thermal equilibrium minority carrier concentrations. The ambipolar transport equation describes the behavior of excess carriers as a function of time and spatial coordinates.

In Chapter 5, we calculated drift current densities in a semiconductor. We determined that relatively large currents could be created with fairly small electric fields. As a first approximation, we assume that the electric field is zero in both the neutral p and n regions. In the n region for $x > x_n$, we have that $E = 0$ and $g' = 0$. If we also assume steady state so $\partial(\delta p_n)/\partial t = 0$, then Equation (8.8) reduces to

$$\frac{d^2 (\delta p_n)}{dx^2} - \frac{\delta p_n}{L_p^2} = 0 \quad (x > x_n) \quad (8.9)$$

where $L_p^2 = D_p \tau_{p0}$. For the same set of conditions, the excess minority carrier electron concentration in the p region is determined from

$$\frac{d^2 (\delta n_p)}{dx^2} - \frac{\delta n_p}{L_n^2} = 0 \quad (x < x_p) \quad (8.10)$$

where $L_n^2 = D_n \tau_{n0}$.

The boundary conditions for the total minority carrier concentrations are

$$p_n(x_n) = p_{n0} \exp\left(\frac{eV_a}{kT}\right) \quad (8.11a)$$

$$n_p(-x_p) = n_{p0} \exp\left(\frac{eV_a}{kT}\right) \quad (8.11b)$$

$$p_n(x \rightarrow +\infty) = p_{n0} \quad (8.11c)$$

$$n_p(x \rightarrow -\infty) = n_{p0} \quad (8.11d)$$

As minority carriers diffuse from the space charge edge into the neutral semiconductor regions, they recombine with majority carriers. We assume that the lengths W_n and W_p shown in Figure 8.3a are very long, meaning in particular that $W_n \gg L_p$ and $W_p \gg L_n$. The excess minority carrier concentrations must approach zero at distances far from the space charge region. The structure is referred to as a long pn junction.

The general solution to Equation (8.9) is

$$\delta p_n(x) = p_n(x) - p_{n0} = Ae^{x/L_p} + Be^{-x/L_p} \quad (x \geq x_n) \quad (8.12)$$

and the general solution to Equation (8.10) is

$$\delta n_p(x) = n_p(x) - n_{p0} = Ce^{x/L_n} + De^{-x/L_n} \quad (x \leq -x_p) \quad (8.13)$$

Applying the boundary conditions from Equations (8.11c) and (8.11d), the coefficients A and D must be zero. The coefficients B and C may be determined from the boundary conditions given by Equations (8.11a) and (8.11b). The excess carrier concentrations are then found to be, for ($x \geq x_n$),

$$\boxed{\delta p_n(x) = p_n(x) - p_{n0} = p_{n0} \left[\exp\left(\frac{eV_a}{kT}\right) - 1 \right] \exp\left(\frac{x_n - x}{L_p}\right)} \quad (8.14)$$

and, for ($x \leq -x_p$),

$$\boxed{\delta n_p(x) = n_p(x) - n_{p0} = n_{p0} \left[\exp\left(\frac{eV_a}{kT}\right) - 1 \right] \exp\left(\frac{x_p + x}{L_n}\right)} \quad (8.15)$$

The minority carrier concentrations decay exponentially with distance away from the junction to their thermal-equilibrium values. Figure 8.5 shows these results. Again, we have assumed that both the n-region and the p-region lengths are long compared to the minority carrier diffusion lengths.

In Chapter 6, we discussed the concept of quasi-Fermi levels, which apply to excess carriers in a nonequilibrium condition. Since excess electrons exist in the neutral p region and excess holes exist in the neutral n region, we can apply quasi-Fermi levels to these regions. We had defined quasi-Fermi levels in terms of carrier concentrations as

$$p = p_o + \delta p = n_i \exp\left(\frac{E_{Fi} - E_{Fp}}{kT}\right) \quad (8.16)$$

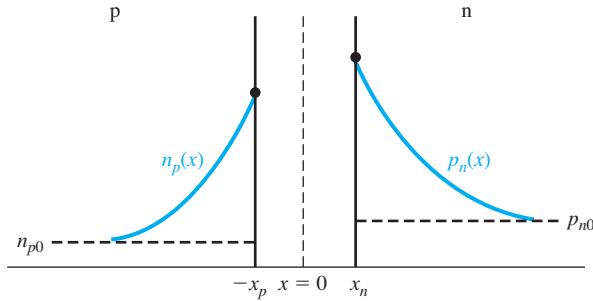


Figure 8.5 | Steady-state minority carrier concentrations in a pn junction under forward bias.

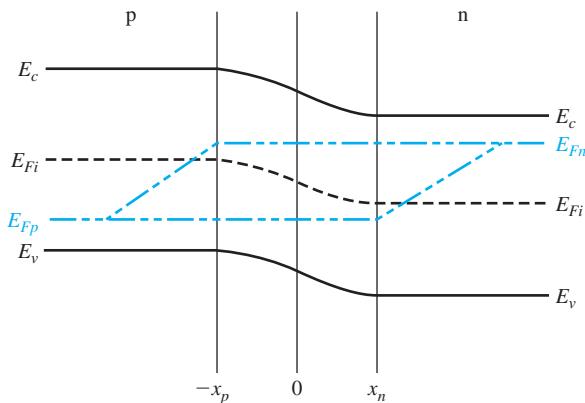


Figure 8.6 | Quasi-Fermi levels through a forward-biased pn junction.

and

$$n = n_o + \delta n = n_i \exp\left(\frac{E_{Fn} - E_{Fi}}{kT}\right) \quad (8.17)$$

Figure 8.6 shows the quasi-Fermi levels through the pn junction. From Equations (8.14) and (8.15), the carrier concentrations are exponential functions of distance, and from Equations (8.16) and (8.17), the carrier concentrations are exponential functions of the quasi-Fermi levels. The quasi-Fermi levels are then linear functions of distance in the neutral p and n regions as shown in Figure 8.6.

We may note that close to the space charge edge in the p region, $E_{Fn} - E_{Fi} > 0$ which means that $\delta n > n_i$. Further from the space charge edge, $E_{Fn} - E_{Fi} < 0$ which means that $\delta n < n_i$ and the excess electron concentration is approaching zero. The same discussion applies to the excess hole concentration in the n region.

At the space charge edge at $x = x_n$, we can write, for low injection

$$n_o p_n(x_n) = n_o p_{no} \exp\left(\frac{V_a}{V_t}\right) = n_i^2 \exp\left(\frac{V_a}{V_t}\right) \quad (8.18)$$

Combining Equations (8.16) and (8.17), we can write

$$np = n_i^2 \exp\left(\frac{E_{Fn} - E_{Fp}}{kT}\right) \quad (8.19)$$

Comparing Equations (8.18) and (8.19), the difference in quasi-Fermi levels is related to the applied bias V_a and represents the deviation from thermal equilibrium. The difference between E_{Fn} and E_{Fp} is nearly constant through the depletion region.

To review, a forward-bias voltage lowers the built-in potential barrier of a pn junction so that electrons from the n region are injected across the space charge region, creating excess minority carriers in the p region. These excess electrons begin diffusing into the bulk p region where they can recombine with majority carrier holes. The excess minority carrier electron concentration then decreases with distance from the junction. The same discussion applies to holes injected across the space charge region into the n region.

8.1.5 Ideal pn Junction Current

The approach we use to determine the current in a pn junction is based on the three parts of the fourth assumption stated earlier in this section. The total current in the junction is the sum of the individual electron and hole currents that are constant through the depletion region. Since the electron and hole currents are continuous functions through the pn junction, the total pn junction current will be the minority carrier hole diffusion current at $x = x_n$ plus the minority carrier electron diffusion current at $x = -x_p$. The gradients in the minority carrier concentrations, as shown in Figure 8.5, produce diffusion currents, and since we are assuming the electric field to be zero at the space charge edges, we can neglect any minority carrier drift current component. This approach in determining the pn junction current is shown in Figure 8.7.

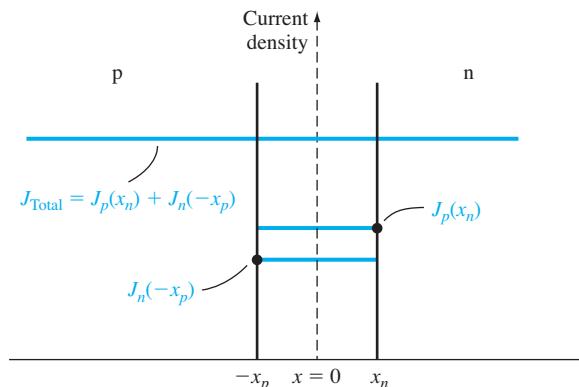


Figure 8.7 | Electron and hole current densities through the space charge region of a pn junction.

We can calculate the minority carrier hole diffusion current density at $x = x_n$ from the relation

$$J_p(x_n) = -eD_p \frac{dp_n(x)}{dx} \Big|_{x=x_n} \quad (8.20)$$

Since we are assuming uniformly doped regions, the thermal-equilibrium carrier concentration is constant, so the hole diffusion current density may be written as

$$J_p(x_n) = -eD_p \frac{d(\delta p_n(x))}{dx} \Big|_{x=x_n} \quad (8.21)$$

Taking the derivative of Equation (8.14) and substituting into Equation (8.21), we obtain

$$J_p(x_n) = \frac{eD_p p_{n0}}{L_p} \left[\exp\left(\frac{eV_a}{kT}\right) - 1 \right] \quad (8.22)$$

The hole current density for this forward-bias condition is in the $+x$ direction, which is from the p to the n region.

Similarly, we may calculate the electron diffusion current density at $x = -x_p$. This may be written as

$$J_n(-x_p) = eD_n \frac{d(\delta n_p(x))}{dx} \Big|_{x=-x_p} \quad (8.23)$$

Using Equation (8.15), we obtain

$$J_n(-x_p) = \frac{eD_n n_{p0}}{L_n} \left[\exp\left(\frac{eV_a}{kT}\right) - 1 \right] \quad (8.24)$$

The electron current density is also in the $+x$ direction.

An assumption we made at the beginning was that the individual electron and hole currents were continuous functions and constant through the space charge region. The total current is the sum of the electron and hole currents and is constant through the entire junction. Figure 8.7 again shows a plot of the magnitudes of these currents.

The total current density in the pn junction is then

$$J = J_p(x_n) + J_n(-x_p) = \left[\frac{eD_p p_{n0}}{L_p} + \frac{eD_n n_{p0}}{L_n} \right] \left[\exp\left(\frac{eV_a}{kT}\right) - 1 \right] \quad (8.25)$$

Equation (8.25) is the ideal current–voltage relationship of a pn junction.

We may define a parameter J_s as

$$J_s = \left[\frac{eD_p p_{n0}}{L_p} + \frac{eD_n n_{p0}}{L_n} \right] \quad (8.26)$$

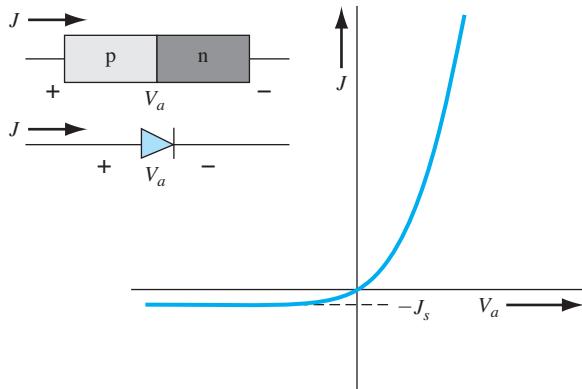


Figure 8.8 | Ideal I - V characteristic of a pn junction diode.

so that Equation (8.25) may be written as

$$J = J_s \left[\exp \left(\frac{eV_a}{kT} \right) - 1 \right] \quad (8.27)$$

Equation (8.27), known as the ideal-diode equation, gives a good description of the current–voltage characteristics of the pn junction over a wide range of currents and voltages. Although Equation (8.27) was derived assuming a forward-bias voltage ($V_a > 0$), there is nothing to prevent V_a from being negative (reverse bias). Equation (8.27) is plotted in Figure 8.8 as a function of forward-bias voltage V_a . If the voltage V_a becomes negative (reverse bias) by a few kT/eV , then the reverse-biased current density becomes independent of the reverse-biased voltage. The parameter J_s is then referred to as the reverse-saturation current density. The current–voltage characteristics of the pn junction diode are obviously not bilateral.

EXAMPLE 8.2

Objective: Determine the ideal reverse-saturation current density in a silicon pn junction at $T = 300$ K.

Consider the following parameters in a silicon pn junction:

$$\begin{aligned} N_a = N_d &= 10^{16} \text{ cm}^{-3} & n_i &= 1.5 \times 10^{10} \text{ cm}^{-3} \\ D_n &= 25 \text{ cm}^2/\text{s} & \tau_{p0} = \tau_{n0} &= 5 \times 10^{-7} \text{ s} \\ D_p &= 10 \text{ cm}^2/\text{s} & \epsilon_r &= 11.7 \end{aligned}$$

■ Solution

The ideal reverse-saturation current density is given by

$$J_s = \frac{eD_n n_{p0}}{L_n} + \frac{eD_p p_{n0}}{L_p}$$

which may be rewritten as

$$J_s = en_i^2 \left(\frac{1}{N_a} \sqrt{\frac{D_n}{\tau_{n0}}} + \frac{1}{N_d} \sqrt{\frac{D_p}{\tau_{p0}}} \right)$$

Then

$$J_s = (1.6 \times 10^{-19})(1.5 \times 10^{10})^2 \left(\frac{1}{10^{16}} \sqrt{\frac{25}{5 \times 10^{-7}}} + \frac{1}{10^{16}} \sqrt{\frac{10}{5 \times 10^{-7}}} \right)$$

or $J_s = 4.16 \times 10^{-11} \text{ A/cm}^2$

Comment

The ideal reverse-biased saturation current density is very small. If the pn junction cross-sectional area were $A = 10^{-4} \text{ cm}^2$, for example, then the ideal reverse-biased diode current would be $I_s = 4.15 \times 10^{-15} \text{ A}$.

EXERCISE PROBLEM

- Ex 8.2** Consider a GaAs pn junction diode at $T = 300 \text{ K}$. The parameters of the device are $N_d = 2 \times 10^{16} \text{ cm}^{-3}$, $N_a = 8 \times 10^{15} \text{ cm}^{-3}$, $D_n = 210 \text{ cm}^2/\text{s}$, $D_p = 8 \text{ cm}^2/\text{s}$, $\tau_{n0} = 10^{-7} \text{ s}$, and $\tau_{p0} = 5 \times 10^{-8} \text{ s}$. Determine the ideal reverse-saturation current density.
 Ans. $J_s = 3.30 \times 10^{-11} \text{ A/cm}^2$

If the forward-bias voltage in Equation (8.27) is positive by more than a few kT/eV , then the (-1) term in Equation (8.27) becomes negligible. Figure 8.9 shows the forward-bias current–voltage characteristic when the current is plotted on a log scale. Ideally, this plot yields a straight line when V_a is greater than a few kT/eV . The forward-bias current is an exponential function of the forward-bias voltage.

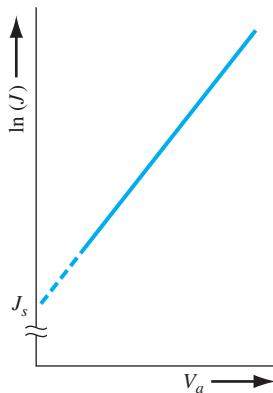


Figure 8.9 | Ideal I – V characteristic of a pn junction diode with the current plotted on a log scale.

EXAMPLE 8.3

Objective: Design a pn junction diode to produce particular electron and hole current densities at a given forward-bias voltage.

Consider a silicon pn junction diode at $T = 300\text{ K}$. Design the diode such that $J_n = 20\text{ A/cm}^2$ and $J_p = 5\text{ A/cm}^2$ at $V_a = 0.65\text{ V}$. Assume the remaining semiconductor parameters are as given in Example 8.2.

■ **Solution**

The electron diffusion current density is given by Equation (8.24) as

$$J_n = \frac{eD_n n_{p0}}{L_n} \left[\exp\left(\frac{eV_a}{kT}\right) - 1 \right] = e \sqrt{\frac{D_n}{\tau_{n0}}} \cdot \frac{n_i^2}{N_a} \left[\exp\left(\frac{eV_a}{kT}\right) - 1 \right]$$

Substituting the numbers, we have

$$20 = (1.6 \times 10^{-19}) \sqrt{\frac{25}{5 \times 10^{-7}}} \cdot \frac{(1.5 \times 10^{10})^2}{N_a} \left[\exp\left(\frac{0.65}{0.0259}\right) - 1 \right]$$

which yields

$$N_a = 1.01 \times 10^{15} \text{ cm}^{-3}$$

The hole diffusion current density is given by Equation (8.22) as

$$J_p = \frac{eD_p p_{n0}}{L_p} \left[\exp\left(\frac{eV_a}{kT}\right) - 1 \right] = e \sqrt{\frac{D_p}{\tau_{p0}}} \cdot \frac{n_i^2}{N_d} \left[\exp\left(\frac{eV_a}{kT}\right) - 1 \right]$$

Substituting the numbers, we have

$$5 = (1.6 \times 10^{-19}) \sqrt{\frac{10}{5 \times 10^{-7}}} \cdot \frac{(1.5 \times 10^{10})^2}{N_d} \left[\exp\left(\frac{0.65}{0.0259}\right) - 1 \right]$$

which yields

$$N_d = 2.55 \times 10^{15} \text{ cm}^{-3}$$

■ **Comment**

The relative magnitude of the electron and hole current densities through a diode can be varied by changing the doping concentrations in the device.

■ **EXERCISE PROBLEM**

Ex 8.3 Using the parameters given in Ex 8.2 for the GaAs diode, determine the electron and hole current densities at the space charge edges, and determine the total current density in the diode for a forward-bias voltage of $V_a = 1.05\text{ V}$.

[Ans. $J_n(-x_e) = 1.20\text{ A/cm}^2$, $J_n(x_e) = 0.1325\text{ A/cm}^2$, $J_p = 1.33\text{ A/cm}^2$]

8.1.6 Summary of Physics

We have been considering the case of a forward-bias voltage being applied to a pn junction. The forward-bias voltage lowers the potential barrier so that electrons and holes are injected across the space charge region. The injected carriers become minority carriers which then diffuse from the junction and recombine with majority carriers.

We calculated the minority carrier diffusion current densities at the edge of the space charge region. We can reconsider Equations (8.14) and (8.15) and determine

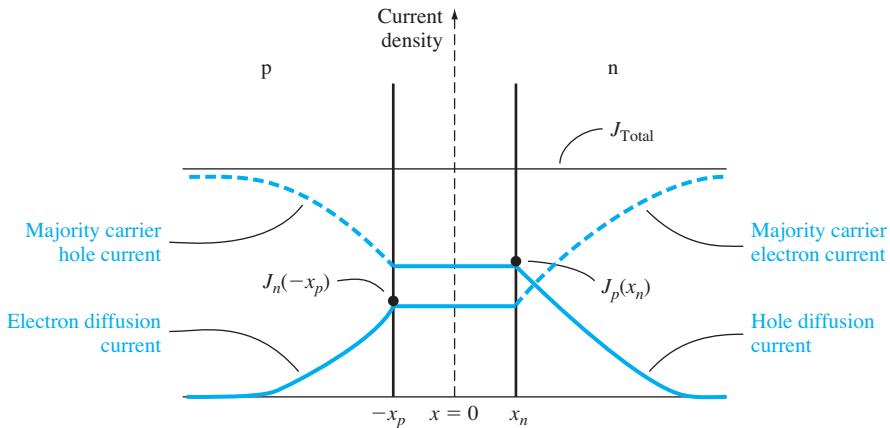


Figure 8.10 | Ideal electron and hole current components through a pn junction under forward bias.

the minority carrier diffusion current densities as a function of distance through the p and n regions. These results are

$$J_p(x) = \frac{eD_p p_{n0}}{L_p} \left[\exp\left(\frac{eV_a}{kT}\right) - 1 \right] \exp\left(\frac{x_n - x}{L_p}\right) \quad (x \geq x_n) \quad (8.28)$$

and

$$J_n(x) = \frac{eD_n n_{p0}}{L_n} \left[\exp\left(\frac{eV_a}{kT}\right) - 1 \right] \exp\left(\frac{x_p + x}{L_n}\right) \quad (x \leq -x_p) \quad (8.29)$$

The minority carrier diffusion current densities decay exponentially in each region. However, the total current through the pn junction is constant. The difference between total current and minority carrier diffusion current is a majority carrier current. Figure 8.10 shows the various current components through the pn structure. The drift of majority carrier holes in the p region far from the junction, for example, is to supply holes that are being injected across the space charge region into the n region and also to supply holes that are lost by recombination with excess minority carrier electrons. The same discussion applies to the drift of electrons in the n region.

We have seen that excess carriers are created in a forward-biased pn junction. From the results of the ambipolar transport theory derived in Chapter 6, the behavior of the excess carriers is determined by the minority carrier parameters for low injection. In determining the current–voltage relationship of the pn junction, we consider the flow of minority carriers since we know the behavior and characteristics of these particles. It may seem strange, at times, that we concern ourselves so much with minority carriers rather than with the vast number of majority carriers, but the reason for this can be found in the results derived from the ambipolar transport theory.

The fact that we now have drift current densities in the p and n regions implies that the electric field in these regions is not zero as we had originally assumed. We can calculate the electric field in the neutral regions and determine the validity of our zero-field approximation.

EXAMPLE 8.4

Objective: Calculate the electric field in a neutral region of a silicon diode to produce a given majority carrier drift current density.

Consider a silicon pn junction at $T = 300$ K with the parameters given in Example 8.2 and with an applied forward-bias voltage $V_a = 0.65$ V.

■ Solution

The total forward-bias current density is given by

$$J = J_s \left[\exp\left(\frac{eV}{kT}\right) - 1 \right]$$

We determined the reverse-saturation current density in Example 8.2, so we can write

$$J = (4.155 \times 10^{-11}) \left[\exp\left(\frac{0.65}{0.0259}\right) - 1 \right] = 3.295 \text{ A/cm}^2$$

The total current far from the junction in the n region will be majority carrier electron drift current, so we can write

$$J = J_n \approx e\mu_n N_d E$$

The doping concentration is $N_d = 10^{16}$ cm⁻³, and, if we assume $\mu_n = 1350$ cm²/V-s, then the electric field must be

$$E = \frac{J_n}{e\mu_n N_d} = \frac{3.295}{(1.6 \times 10^{-19})(1350)(10^{16})} = 1.525 \text{ V/cm}$$

■ Comment

We assumed, in the derivation of the current–voltage equation, that the electric field in the neutral p and n regions was zero. Although the electric field is not zero, this example shows that the magnitude is very small—thus the approximation of zero electric field is very good.

■ EXERCISE PROBLEM

Ex 8.4 Determine the electric field in the neutral n region and neutral p region for the GaAs pn junction diode described in Ex 8.3.

(Ans. $E_n = 0.0694$ V/cm, $E_p = 3.25$ V/cm)

8.1.7 Temperature Effects

The ideal reverse-saturation current density J_s , given by Equation (8.26), is a function of the thermal-equilibrium minority carrier concentrations n_{p0} and p_{n0} . These minority carrier concentrations are proportional to n^2 , which is a very strong function of temperature. For a silicon pn junction, the ideal reverse-saturation current density will increase by approximately a factor of 4 for every 10°C increase in temperature.

The forward-bias current–voltage relation is given by Equation (8.27). This relation includes J_s as well as the $\exp(-eV_a/kT)$ factor, making the forward-bias current–voltage relation a function of temperature also. As temperature increases, less forward-bias voltage is required to obtain the same diode current. If the voltage is held constant, the diode current will increase as temperature increases. The change in forward-bias current with temperature is less sensitive than the reverse-saturation current.

Objective: Determine the change in the forward-bias voltage on a pn junction with a change in temperature to maintain a constant diode current.

EXAMPLE 8.5

Consider a silicon pn junction initially biased at 0.60 V at $T = 300$ K. Assume the temperature increases to $T = 310$ K. Calculate the change in the forward-bias voltage required to maintain a constant current through the junction.

■ Solution

The forward-bias current can be written as follows:

$$J \propto \exp\left(\frac{-E_g}{kT}\right) \exp\left(\frac{eV_a}{kT}\right)$$

If the temperature changes, we may take the ratio of the diode currents at the two temperatures. This ratio is

$$\frac{J_2}{J_1} = \frac{\exp(-E_g/kT_2) \exp(eV_{a2}/kT_2)}{\exp(-E_g/kT_1) \exp(eV_{a1}/kT_1)}$$

If current is to be held constant, then $J_1 = J_2$, and we must have

$$\frac{E_g - eV_{a2}}{kT_2} = \frac{E_g - eV_{a1}}{kT_1}$$

For $T_1 = 300$ K, $T_2 = 310$ K, $E_g = 1.12$ eV, and $V_{a1} = 0.60$ V. We then find

$$\frac{1.12 - V_{a2}}{310} = \frac{1.12 - 0.60}{300}$$

which yields

$$V_{a2} = 0.5827 \text{ V}$$

■ Comment

The change in the forward-bias voltage is -17.3 mV for a 10°C temperature change.

■ EXERCISE PROBLEM

Ex 8.5 Repeat Example 8.5 for a GaAs pn junction diode biased at $V_a = 1.050$ V for $T = 300$ K.

(Ans. -12.3 mV)

8.1.8 The “Short” Diode

We assumed in the previous analysis that both p and n regions were long compared with the minority carrier diffusion lengths. In many pn junction structures, one region may, in fact, be short compared with the minority carrier diffusion length. Figure 8.11 shows one such example: the length W_n is assumed to be much smaller than the minority carrier hole diffusion length, L_p .

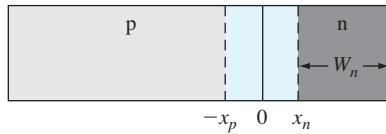


Figure 8.11 | Geometry of a “short” diode.

The steady-state excess minority carrier hole concentration in the n region is determined from Equation (8.9), which is given as

$$\frac{d^2(\delta p_n)}{dx^2} - \frac{\delta p_n}{L_p^2} = 0$$

The original boundary condition at $x = x_n$ still applies, given by Equation (8.11a) as

$$p_n(x_n) = p_{n0} \exp\left(\frac{eV_a}{kT}\right)$$

A second boundary condition needs to be determined. In many cases we assume that an ohmic contact exists at $x = (x_n + W_n)$, implying an infinite surface-recombination velocity and therefore an excess minority carrier concentration of zero. The second boundary condition is then written as

$$p_n(x = x_n + W_n) = p_{n0} \quad (8.30)$$

The general solution to Equation (8.9) is again given by Equation (8.12), which was

$$\delta p_n(x) = p_n(x) - p_{n0} = Ae^{x/L_p} + Be^{-x/L_p} \quad (x \geq x_n)$$

In this case, because of the finite length of the n region, both terms of the general solution must be retained. Applying the boundary conditions of Equations (8.11b) and (8.30), the excess minority carrier concentration is given by

$$\delta p_n(x) = p_{n0} \left[\exp\left(\frac{eV_a}{kT}\right) - 1 \right] \frac{\sinh[(x_n + W_n - x)/L_p]}{\sinh[W_n/L_p]} \quad (8.31)$$

Equation (8.31) is the general solution for the excess minority carrier hole concentration in the n region of a forward-biased pn junction. If $W_n \gg L_p$, the assumption for the long diode, Equation (8.31) reduces to the previous result given by Equation (8.14). If $W_n \ll L_p$, we can approximate the hyperbolic sine terms by

$$\sinh\left(\frac{x_n + W_n - x}{L_p}\right) \approx \left(\frac{x_n + W_n - x}{L_p}\right) \quad (8.32a)$$

and

$$\sinh\left(\frac{W_n}{L_p}\right) \approx \left(\frac{W_n}{L_p}\right) \quad (8.32b)$$

Then Equation (8.31) becomes

$$\delta p_n(x) = p_{n0} \left[\exp\left(\frac{eV_a}{kT}\right) - 1 \right] \left(\frac{x_n + W_n - x}{W_n} \right) \quad (8.33)$$

The minority carrier concentration becomes a linear function of distance.

The minority carrier hole diffusion current density is given by

$$J_p = -eD_p \frac{d[\delta p_n(x)]}{dx}$$

so that in the short n region, we have

$$J_p(x) = \frac{eD_p p_{n0}}{W_n} \left[\exp\left(\frac{eV_a}{kT}\right) - 1 \right] \quad (8.34)$$

The minority carrier hole diffusion current density now contains the length W_n in the denominator, rather than the diffusion length L_p . The diffusion current density is larger for a short diode than for a long diode since $W_n \ll L_p$. In addition, since the minority carrier concentration is approximately a linear function of distance through the n region, the minority carrier diffusion current density is a constant. This constant current implies that there is no recombination of minority carriers in the short region.

TEST YOUR UNDERSTANDING

- TYU 8.1** The doping concentrations in a GaAs pn junction diode at $T = 300$ K are $N_d = 5 \times 10^{15} \text{ cm}^{-3}$ and $N_a = 5 \times 10^{16} \text{ cm}^{-3}$. The minority carrier concentration at either space charge edge is to be no larger than 10 percent of the respective majority carrier concentration. Calculate the maximum forward-bias voltage that can be applied to this junction and still meet the required specifications.

[Ans. $V_a(\text{max}) = 1.067 \text{ V}$]

- TYU 8.2** A silicon pn junction at $T = 300$ K has the following parameters: $N_a = 5 \times 10^{16} \text{ cm}^{-3}$, $N_d = 1 \times 10^{16} \text{ cm}^{-3}$, $D_n = 25 \text{ cm}^2/\text{s}$, $D_p = 10 \text{ cm}^2/\text{s}$, $\tau_{n0} = 5 \times 10^{-7} \text{ s}$, and $\tau_{p0} = 1 \times 10^{-7} \text{ s}$. The cross-sectional area is $A = 10^{-3} \text{ cm}^2$ and the forward-bias voltage is $V_a = 0.625 \text{ V}$. Calculate the (a) minority electron diffusion current at the space charge edge, (b) minority hole diffusion current at the space charge edge, and (c) total current in the pn junction diode.

[Ans. (a) 0.154 mA; (b) 1.09 mA; (c) 1.24 mA]

- TYU 8.3** Consider the silicon pn junction diode described in TYU 8.2. The p region is long and the n region is short with $W_n = 2 \mu\text{m}$. (a) Calculate the electron and hole currents in the depletion region. (b) Why has the hole current increased compared to that found in TYU 8.2?

[Ans. (a) $I_e = 0.154 \text{ mA}$, $I_h = 5.44 \text{ mA}$; (b) The hole density gradient has increased.]

8.2 | GENERATION–RECOMBINATION CURRENTS AND HIGH-INJECTION LEVELS

In the derivation of the ideal current–voltage relationship, we assumed low injection and neglected any effects occurring within the space charge region. High-level injection and other current components generated within the space charge region cause

the I - V relationship to deviate from the ideal expression. The additional currents are generated from the recombination processes discussed in Chapter 6.

8.2.1 Generation–Recombination Currents

The recombination rate of excess electrons and holes, given by the Shockley–Read–Hall recombination theory, was written as

$$R = \frac{C_n C_p N_i (np - n_i^2)}{C_n (n + n') + C_p (p + p')} \quad (8.35)$$

The parameters n and p are, as usual, the concentrations of electrons and holes, respectively.

Reverse-Biased Generation Current For a pn junction under reverse bias, we have argued that the mobile electrons and holes have essentially been swept out of the space charge region. Accordingly, within the space charge region, $n \approx p \approx 0$. The recombination rate from Equation (8.35) becomes

$$R = \frac{-C_n C_p N_i n_i^2}{C_n n' + C_p p'} \quad (8.36)$$

The negative sign implies a negative recombination rate; hence, we are really generating electron–hole pairs within the reverse-biased space charge region. The recombination of excess electrons and holes is the process whereby we are trying to reestablish thermal equilibrium. Since the concentration of electrons and holes is essentially zero within the reverse-biased space charge region, electrons and holes are being generated via the trap level to also try to reestablish thermal equilibrium. This generation process is schematically shown in Figure 8.12. As the electrons and holes

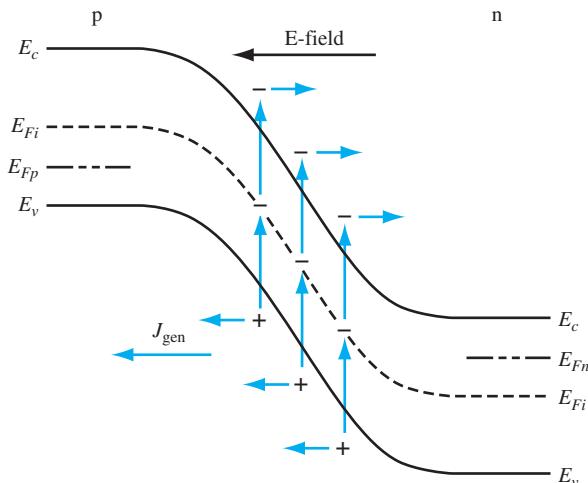


Figure 8.12 | Generation process in a reverse-biased pn junction.

are generated, they are swept out of the space charge region by the electric field. The flow of charge is in the direction of a reverse-biased current. This *reverse-biased generation current*, caused by the generation of electrons and holes in the space charge region, is in addition to the ideal reverse-biased saturation current.

We may calculate the density of the reverse-biased generation current by considering Equation (8.36). If we make a simplifying assumption and let the trap level be at the intrinsic Fermi level, then from Equations (6.92) and (6.97), we have that $n' = n_i$ and $p' = n_i$. Equation (8.36) now becomes

$$R = \frac{-n_i}{\frac{1}{N_t C_p} + \frac{1}{N_t C_n}} \quad (8.37)$$

Using the definitions of lifetimes from Equations (6.103) and (6.104), we may write Equation (8.37) as

$$R = \frac{-n_i}{\tau_{p0} + \tau_{n0}} \quad (8.38)$$

If we define a new lifetime as the average of τ_{p0} and τ_{n0} , or

$$\tau_0 = \frac{\tau_{p0} + \tau_{n0}}{2} \quad (8.39)$$

then the recombination rate can be written as

$$R = \frac{-n_i}{2\tau_0} \equiv -G \quad (8.40)$$

The negative recombination rate implies a generation rate, so G is the generation rate of electrons and holes in the space charge region.

The generation current density may be determined from

$$J_{\text{gen}} = \int_0^W e G dx \quad (8.41)$$

where the integral is over the space charge region. If we assume that the generation rate is constant throughout the space charge region, then we obtain

$$J_{\text{gen}} = \frac{en_i W}{2\tau_0} \quad (8.42)$$

The total reverse-biased current density is the sum of the ideal reverse saturation current density and the generation current density, or

$$J_R = J_s + J_{\text{gen}} \quad (8.43)$$

The ideal reverse-saturation current density J_s is independent of the reverse-biased voltage. However, J_{gen} is a function of the depletion width W , which in turn is a function of the reverse-biased voltage. The actual reverse-biased current density, then, is no longer independent of the reverse-biased voltage.

EXAMPLE 8.6

Objective: Determine the relative magnitudes of the ideal reverse-saturation current density and the generation current density in a reverse-biased pn junction.

Consider a silicon pn junction at $T = 300$ K with parameters $D_n = 25 \text{ cm}^2/\text{s}$, $D_p = 10 \text{ cm}^2/\text{s}$, $N_a = N_d = 10^{16} \text{ cm}^{-3}$, and $\tau_0 = \tau_{n0} = \tau_{p0} = 5 \times 10^{-7} \text{ s}$. Assume the diode is reverse biased at $V_R = 5 \text{ V}$.

Solution

The ideal reverse-saturation current density was calculated in Example 8.2 and was found to be $J_s = 4.155 \times 10^{-11} \text{ A/cm}^2$.

The built-in potential is found as

$$V_{bi} = V_t \ln \left(\frac{N_a N_d}{n_i^2} \right) = (0.0259) \ln \left[\frac{(10^{16})(10^{16})}{(1.5 \times 10^{10})^2} \right] = 0.695 \text{ V}$$

The depletion width is found to be

$$\begin{aligned} W &= \left\{ \frac{2 e_s (V_{bi} + V_R)}{e} \left(\frac{N_a + N_d}{N_a N_d} \right) \right\}^{1/2} \\ &= \left\{ \frac{2(11.7)(8.85 \times 10^{-14})(0.695 + 5)}{1.6 \times 10^{-19}} \left[\frac{10^{16} + 10^{16}}{(10^{16})(10^{16})} \right] \right\}^{1/2} \\ &= 1.214 \times 10^{-4} \text{ cm} \end{aligned}$$

The generation current density is then found to be

$$J_{gen} = \frac{en_i W}{2\tau_0} = \frac{(1.6 \times 10^{-19})(1.5 \times 10^{10})(1.214 \times 10^{-4})}{2(5 \times 10^{-7})}$$

or

$$J_{gen} = 2.914 \times 10^{-7} \text{ A/cm}^2$$

The ratio of the two currents is

$$\frac{J_{gen}}{J_s} = \frac{2.914 \times 10^{-7}}{4.155 \times 10^{-11}} \cong 7 \times 10^3$$

Comment

Comparing the solutions for the two current densities, it is obvious that, for the silicon pn junction diode at room temperature, the generation current density is approximately four orders of magnitude larger than the ideal saturation current density. The generation current is the dominant reverse-biased current in a silicon pn junction diode.

EXERCISE PROBLEM

Ex 8.6 Consider a GaAs pn junction diode at $T = 300$ K with parameters $N_d = 8 \times 10^{16} \text{ cm}^{-3}$, $N_a = 2 \times 10^{15} \text{ cm}^{-3}$, $D_n = 207 \text{ cm}^2/\text{s}$, $D_p = 9.80 \text{ cm}^2/\text{s}$, and $\tau_0 = \tau_{p0} = \tau_{n0} = 5 \times 10^{-8} \text{ s}$. (a) Calculate the ideal reverse-biased saturation current density. (b) Find the reverse-biased generation current density if the diode is reverse biased at $V_R = 5 \text{ V}$. (c) Determine the ratio of J_{gen} to J_s .

[Ans. (a) $1.677 \times 10^{-17} \text{ A/cm}^2$; (b) $6.166 \times 10^{-10} \text{ A/cm}^2$; (c) 3.68×10^7]

Forward-Bias Recombination Current For the reverse-biased pn junction, electrons and holes are essentially completely swept out of the space charge region so that $n \approx p \approx 0$. Under forward bias, however, electrons and holes are injected across the space charge region, so we do, in fact, have some excess carriers in the space charge region. The possibility exists that some of these electrons and holes will recombine within the space charge region and not become part of the minority carrier distribution.

The recombination rate of electrons and holes is again given from Equation (8.35) as

$$R = \frac{C_n C_p N_i (np - n_i^2)}{C_n(n + n') + C_p(p + p')}$$

Dividing both numerator and denominator by $C_n C_p N_i$ and using the definitions of τ_{n0} and τ_{p0} , we may write the recombination rate as

$$R = \frac{np - n_i^2}{\tau_{p0}(n + n') + \tau_{n0}(p + p')} \quad (8.44)$$

Figure 8.13 shows the energy-band diagram of the forward-biased pn junction. Shown in the figure are the intrinsic Fermi level and the quasi-Fermi levels for electrons and holes. From the results of Chapter 6, we may write the electron concentration as

$$n = n_i \exp\left[\frac{E_{Fn} - E_{Fi}}{kT}\right] \quad (8.45)$$

and the hole concentration as

$$p = n_i \exp\left[\frac{E_{Fi} - E_{Fp}}{kT}\right] \quad (8.46)$$

where E_{Fn} and E_{Fp} are the quasi-Fermi levels for electrons and holes, respectively.

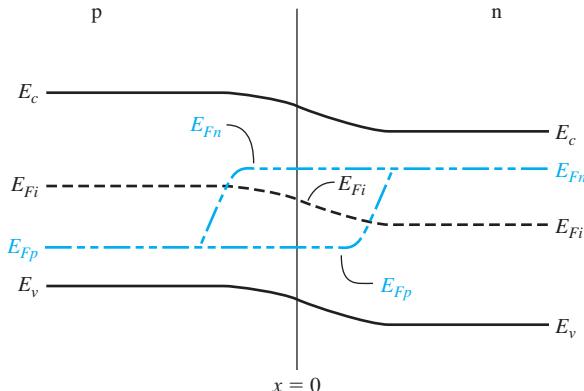


Figure 8.13 | Energy-band diagram of a forward-biased pn junction including quasi-Fermi levels.

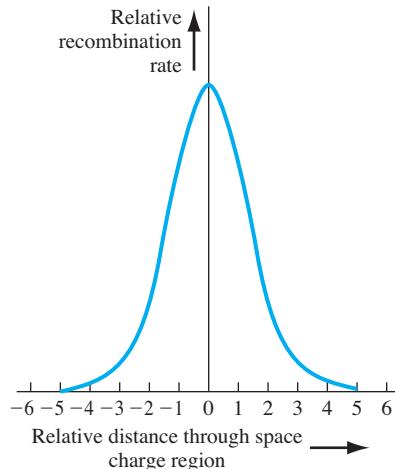


Figure 8.14 | Relative magnitude of the recombination rate through the space charge region of a forward-biased pn junction.

From Figure 8.13, we may note that

$$(E_{Fn} - E_{Fi}) + (E_{Fi} - E_{Fp}) = eV_a \quad (8.47)$$

where V_a is the applied forward-bias voltage. Again, if we assume that the trap level is at the intrinsic Fermi level, then $n' = p' = n_i$. Figure 8.14 shows a plot of the relative magnitude of the recombination rate as a function of distance through the space charge region. This plot was generated using Equations (8.44), (8.45), (8.46), and (8.47). A very sharp peak occurs at the metallurgical junction ($x = 0$).

At the center of the space charge region, we have

$$E_{Fn} - E_{Fi} = E_{Fi} - E_{Fp} = \frac{eV_a}{2} \quad (8.48)$$

Equations (8.45) and (8.46) then become

$$n = n_i \exp\left(\frac{eV_a}{2kT}\right) \quad (8.49)$$

and

$$p = n_i \exp\left(\frac{eV_a}{2kT}\right) \quad (8.50)$$

If we assume that $n' = p' = n_i$ and that $\tau_{n0} = \tau_{p0} = \tau_0$, then Equation (8.44) becomes

$$R_{\max} = \frac{n_i}{2\tau_0} \frac{[\exp(eV_a/kT) - 1]}{[\exp(eV_a/2kT) + 1]} \quad (8.51)$$

which is the maximum recombination rate for electrons and holes that occurs at the center of the forward-biased pn junction. If we assume that $V_a \gg kT/e$, we may

neglect the (-1) term in the numerator and the $(+1)$ term in the denominator. Equation (8.51) then becomes

$$R_{\max} = \frac{n_i}{2\tau_0} \exp\left(\frac{eV_a}{2kT}\right) \quad (8.52)$$

The recombination current density may be calculated from

$$J_{\text{rec}} = \int_0^W eR \, dx \quad (8.53)$$

where again the integral is over the entire space charge region. In this case, however, the recombination rate is not a constant through the space charge region. We have calculated the maximum recombination rate at the center of the space charge region, so we may write

$$J_{\text{rec}} = ex' \frac{n_i}{2\tau_0} \exp\left(\frac{eV_a}{2kT}\right) \quad (8.54)$$

where x' is a length over which the maximum recombination rate is effective. However, since τ_0 may not be a well-defined or known parameter, it is customary to write

$$J_{\text{rec}} = \frac{eWn_i}{2\tau_0} \exp\left(\frac{eV_a}{2kT}\right) = J_{r0} \exp\left(\frac{eV_a}{2kT}\right) \quad (8.55)$$

where W is the space charge width.

Total Forward-Bias Current The total forward-bias current density in the pn junction is the sum of the recombination and the ideal diffusion current densities. Figure 8.15 shows a plot of the minority carrier hole concentration in the neutral

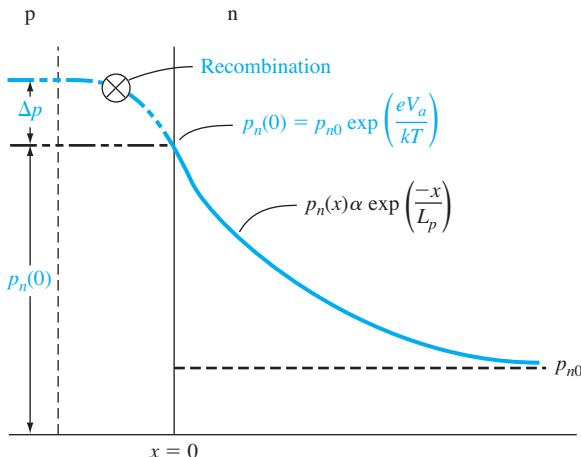


Figure 8.15 | Because of recombination, additional holes from the p region must be injected into the space charge region to establish the minority carrier hole concentration in the n region.

n region. This distribution yields the ideal hole diffusion current density and is a function of the minority carrier hole diffusion length and the applied junction voltage. The distribution is established as a result of holes being injected across the space charge region. If, now, some of the injected holes in the space charge region are lost due to recombination, then additional holes must be injected from the p region to make up for this loss. The flow of these additional injected carriers, per unit time, results in the recombination current. This added component is schematically shown in the figure.

The total forward-bias current density is the sum of the recombination and the ideal diffusion current densities, so we can write

$$J = J_{\text{rec}} + J_D \quad (8.56)$$

where J_{rec} is given by Equation (8.55) and J_D is given by

$$J_D = J_s \exp\left(\frac{eV_a}{kT}\right) \quad (8.57)$$

The (-1) term in Equation (8.27) has been neglected. The parameter J_s is the ideal reverse-saturation current density, and from previous discussion, the value of J_{r0} from the recombination current is larger than the value of J_s .

If we take the natural log of Equations (8.55) and (8.57), we obtain

$$\ln J_{\text{rec}} = \ln J_{r0} + \frac{eV_a}{2kT} = \ln J_{r0} + \frac{V_a}{2V_t} \quad (8.58a)$$

and

$$\ln J_D = \ln J_s + \frac{eV_a}{kT} = \ln J_s + \frac{V_a}{V_t} \quad (8.58b)$$

Figure 8.16 shows the recombination and diffusion current components plotted on a log current scale as a function of V_a/V_t . The slopes of the two curves are not the same. Also shown in the figure is the total current density—the sum of the two current components. We may notice that, at a low current density, the recombination current dominates, and at a higher current density, the ideal diffusion current dominates.

In general, the diode current–voltage relationship may be written as

$I = I_s \left[\exp\left(\frac{eV_a}{nkT}\right) - 1 \right]$

(8.59)

where the parameter n is called the *ideality factor*. For a large forward-bias voltage, $n \approx 1$ when diffusion dominates, and for low forward-bias voltage, $n \approx 2$ when recombination dominates. There is a transition region where $1 < n < 2$.

8.2.2 High-Level Injection

In the derivation of the ideal diode I – V relationship, we assumed that low injection was valid. Low injection implies that the excess minority carrier concentrations are always much less than the majority carrier concentration.

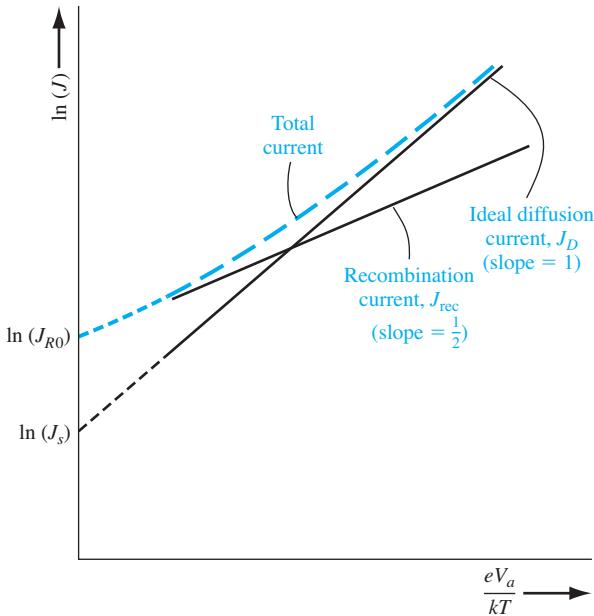


Figure 8.16 | Ideal diffusion, recombination, and total current in a forward-biased pn junction.

However, as the forward-bias voltage increases, the excess carrier concentrations increase and may become comparable or even greater than the majority carrier concentration. From Equation (8.18), we can write

$$np = n_i^2 \exp\left(\frac{V_a}{V_t}\right)$$

We have that $n = n_o + \delta n$ and $p = p_o + \delta p$, so that

$$(n_o + \delta n)(p_o + \delta p) = n_i^2 \exp\left(\frac{V_a}{V_t}\right) \quad (8.60)$$

Under high-level injection, we may have $\delta n > n_o$ and $\delta p > p_o$ so that Equation (8.60) becomes approximately

$$(\delta n)(\delta p) \cong n_i^2 \exp\left(\frac{V_a}{V_t}\right) \quad (8.61)$$

Since $\delta n = \delta p$, then

$$\delta n = \delta p \cong n_i \exp\left(\frac{V_a}{2V_t}\right) \quad (8.62)$$

The diode current is proportional to the excess carrier concentration so that, under high-level injection, we have

$$I \propto \exp\left(\frac{V_a}{2V_t}\right) \quad (8.63)$$

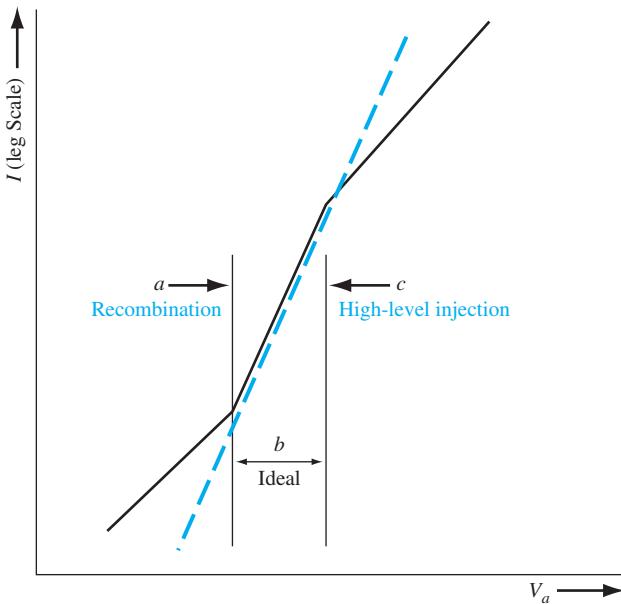


Figure 8.17 | Forward-bias current versus voltage from low forward bias to high forward bias.

In the high-level injection region, it takes a larger increase in diode voltage to produce a given increase in diode current.

The diode forward-bias current, from low-bias levels to high-bias levels, is plotted in Figure 8.17. This plot shows the effect of recombination at low-bias voltages and high-level injection at high-bias voltages.

TEST YOUR UNDERSTANDING

- TYU 8.4** Consider a silicon pn junction diode at $T = 300$ K with parameters $N_a = 2 \times 10^{15} \text{ cm}^{-3}$, $N_d = 8 \times 10^{16} \text{ cm}^{-3}$, $D_p = 10 \text{ cm}^2/\text{s}$, $D_n = 25 \text{ cm}^2/\text{s}$, and $\tau_0 = \tau_{p0} = \tau_{n0} = 10^{-7} \text{ s}$. The diode is forward biased at $V_a = 0.35$ V. (a) Calculate the ideal diode current density. (b) Find the forward-biased recombination current density. (c) Determine the ratio of recombination current to the ideal diffusion current.

[Ans. (a) $2.137 \times 10^{-4} \text{ A/cm}^2$; (b) $5.020 \times 10^{-4} \text{ A/cm}^2$; (c) 2.35]

8.3 | SMALL-SIGNAL MODEL OF THE pn JUNCTION

We have been considering the dc characteristics of the pn junction diode. When semiconductor devices with pn junctions are used in linear amplifier circuits, for example, sinusoidal signals are superimposed on the dc currents and voltages, so that the small-signal characteristics of the pn junction become important.

8.3.1 Diffusion Resistance

The ideal current–voltage relationship of the pn junction diode was given by Equation (8.27), where J and J_s are current densities. If we multiply both sides of the equation by the junction cross-sectional area, we have

$$I_D = I_s \left[\exp\left(\frac{eV_a}{kT}\right) - 1 \right] \quad (8.64)$$

where I_D is the diode current and I_s is the diode reverse-saturation current.

Assume that the diode is forward-biased with a dc voltage V_0 producing a dc diode current I_{DQ} . If we now superimpose a small, low-frequency sinusoidal voltage as shown in Figure 8.18, then a small sinusoidal current will be produced, superimposed on the dc current. The ratio of sinusoidal current to sinusoidal voltage is called the incremental conductance. In the limit of a very small sinusoidal current and voltage, the small-signal incremental conductance is just the slope of the dc current–voltage curve, or

$$g_d = \left. \frac{dI_D}{dV_a} \right|_{V_a=V_0} \quad (8.65)$$

The reciprocal of the incremental conductance is the incremental resistance, defined as

$$r_d = \left. \frac{dV_a}{dI_D} \right|_{I_D=I_{DQ}} \quad (8.66)$$

where I_{DQ} is the dc quiescent diode current.

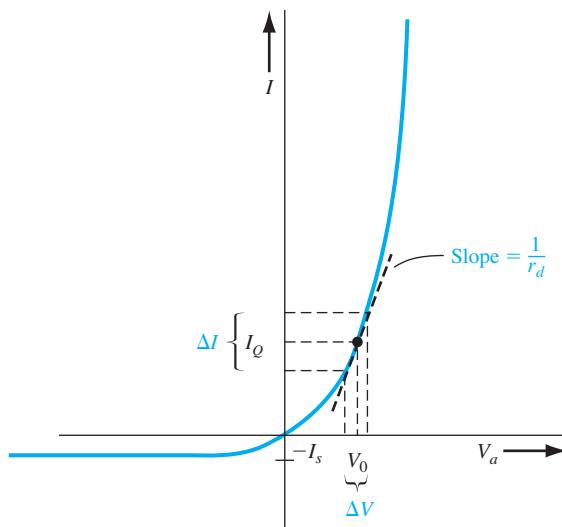


Figure 8.18 | Curve showing the concept of the small-signal diffusion resistance.

If we assume that the diode is biased sufficiently far in the forward-bias region, then the (-1) term can be neglected and the incremental conductance becomes

$$g_d = \left. \frac{dI_D}{dV_a} \right|_{V_a=V_0} = \left(\frac{e}{kT} \right) I_s \exp\left(\frac{eV_0}{kT}\right) \approx \frac{I_{DQ}}{V_t} \quad (8.67)$$

The small-signal incremental resistance is then the reciprocal function, or

$$r_d = \frac{V_t}{I_{DQ}} \quad (8.68)$$

The incremental resistance decreases as the bias current increases, and is inversely proportional to the slope of the I - V characteristic as shown in Figure 8.18. The incremental resistance is also known as the **diffusion resistance**.

8.3.2 Small-Signal Admittance

In the last chapter, we considered the pn junction capacitance as a function of the reverse-biased voltage. When the pn junction diode is forward-biased, another capacitance becomes a factor in the diode admittance. The small-signal admittance, or impedance, of the pn junction under forward bias is derived using the minority carrier diffusion current relations we have already considered.

Qualitative Analysis Before we delve into the mathematical analysis, we can qualitatively understand the physical processes that lead to a diffusion capacitance, which is one component of the junction admittance. Figure 8.19a schematically shows a pn junction forward biased with a dc voltage. A small ac voltage is also

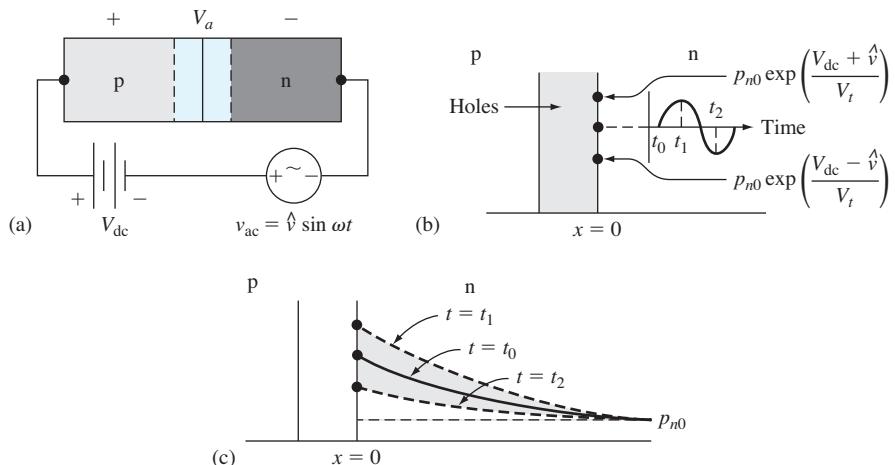


Figure 8.19 | (a) A pn junction with an ac voltage superimposed on a forward-biased dc value; (b) the hole concentration versus time at the space charge edge; (c) the hole concentration versus distance in the n region at three different times.

superimposed on the dc voltage so that the total forward-biased voltage can be written as $V_a = V_{dc} + \hat{v} \sin \omega t$.

As the voltage across the junction changes, the number of holes injected across the space charge region into the n region also changes. Figure 8.19b shows the hole concentration at the space charge edge as a function of time. At $t = t_0$, the ac voltage is zero so that the concentration of holes at $x = 0$ is just given by $p_n(0) = p_{n0} \exp(V_{dc}/V_i)$, which is what we have seen previously.

Now, as the ac voltage increases during its positive half cycle, the concentration of holes at $x = 0$ will increase and reach a peak value at $t = t_1$, which corresponds to the peak value of the ac voltage. When the ac voltage is on its negative half cycle, the total voltage across the junction decreases so that the concentration of holes at $x = 0$ decreases. The concentration reaches a minimum value at $t = t_2$, which corresponds to the time that the ac voltage reaches its maximum negative value. The minority carrier hole concentration at $x = 0$, then, has an ac component superimposed on the dc value as indicated in Figure 8.19b.

As previously discussed, the holes at the space charge edge ($x = 0$) diffuse into the n region where they recombine with the majority carrier electrons. We assume that the period of the ac voltage is large compared to the time it takes carriers to diffuse into the n region. The hole concentration as a function of distance into the n region can then be treated as a steady-state distribution. Figure 8.19c shows the steady-state hole concentrations at three different times. At $t = t_0$, the ac voltage is zero, so the $t = t_0$ curve corresponds to the hole distribution established by the dc voltage. The $t = t_1$ curve corresponds to the distribution established when the ac voltage has reached its peak positive value, and the $t = t_2$ curve corresponds to the distribution established when the ac voltage has reached its maximum negative value. The shaded areas represent the charge ΔQ that is alternately charged and discharged during the ac voltage cycle.

Exactly the same process is occurring in the p region with the electron concentration. The mechanism of charging and discharging of holes in the n region and electrons in the p region leads to a capacitance. This capacitance is called *diffusion capacitance*. The physical mechanism of this diffusion capacitance is different from that of the junction capacitance discussed in the last chapter. We show that the magnitude of the diffusion capacitance in a forward-biased pn junction is usually substantially larger than the junction capacitance.

Mathematical Analysis The minority carrier distribution in the pn junction will be derived for the case when a small sinusoidal voltage is superimposed on the dc junction voltage. We can then determine small signal, or ac, diffusion currents from these minority carrier functions. Figure 8.20 shows the minority carrier distribution in a pn junction when a forward-biased dc voltage is applied. The origin, $x = 0$, is set at the edge of the space charge region on the n side for convenience. The minority carrier hole concentration at $x = 0$ is given by Equation (8.7) as $p_n(0) = p_{n0} \exp(eV_a/kT)$, where V_a is the applied voltage across the junction.

Now let

$$V_a = V_0 + v_1(t) \quad (8.69)$$

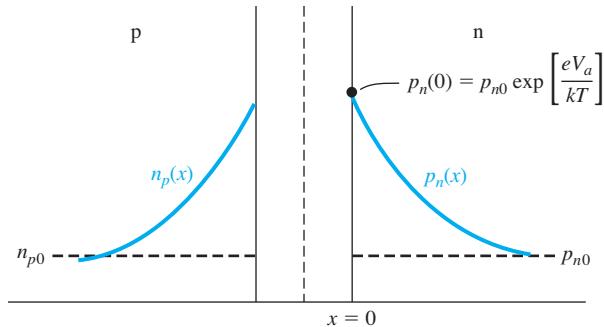


Figure 8.20 | The dc characteristics of a forward-biased pn junction used in the small-signal admittance calculations.

where V_0 is the dc quiescent bias voltage and $v_1(t)$ is the ac signal voltage that is superimposed on this dc level. We may now write

$$p_n(x = 0) = p_{n0} \exp \left\{ \frac{e[V_0 + v_1(t)]}{kT} \right\} = p_n(0, t) \quad (8.70)$$

Equation (8.70) may be written as

$$p_n(0, t) = p_{dc} \exp \left[\frac{ev_1(t)}{kT} \right] \quad (8.71)$$

where

$$p_{dc} = p_{n0} \exp \left(\frac{eV_0}{kT} \right) \quad (8.72)$$

If we assume that $|v_1(t)| \ll (kT/e) = V_t$, then the exponential term in Equation (8.71) may be expanded into a Taylor series retaining only the linear terms, and the minority carrier hole concentration at $x = 0$ can be written as

$$p_n(0, t) \approx p_{dc} \left[1 + \frac{v_1(t)}{V_t} \right] \quad (8.73)$$

If we assume that the time-varying voltage $v_1(t)$ is a sinusoidal signal, we can write Equation (8.73) as

$$p_n(0, t) = p_{dc} \left(1 + \frac{\hat{V}_1}{V_t} e^{j\omega t} \right) \quad (8.74)$$

where \hat{V}_1 is the phasor of the applied sinusoidal voltage. Equation (8.74) will be used as the boundary condition in the solution of the time-dependent diffusion equation for the minority carrier holes in the n region.

In the neutral n region ($x > 0$), the electric field is assumed to be zero; thus, the behavior of the excess minority carrier holes is determined from the equation

$$D_p \frac{\partial^2 (\delta p_n)}{\partial x^2} - \frac{\delta p_n}{\tau_{p0}} = \frac{\partial (\delta p_n)}{\partial t} \quad (8.75)$$

where δp_n is the excess hole concentration in the n region. We are assuming that the ac signal voltage $v_1(t)$ is sinusoidal. We then expect the steady-state solution for δp_n to be of the form of a sinusoidal solution superimposed on the dc solution, or

$$\delta p_n(x, t) = \delta p_0(x) + p_1(x)e^{j\omega t} \quad (8.76)$$

where $\delta p_0(x)$ is the dc excess carrier concentration and $p_1(x)$ is the magnitude of the ac component of the excess carrier concentration. The expression for $\delta p_0(x)$ is the same as that given in Equation (8.14).

Substituting Equation (8.76) into the differential Equation (8.75), we obtain

$$D_p \left\{ \frac{\partial^2 [\delta p_0(x)]}{\partial x^2} + \frac{\partial^2 p_1(x)}{\partial x^2} e^{j\omega t} \right\} - \frac{\delta p_0(x) + p_1(x)e^{j\omega t}}{\tau_{p0}} = j\omega p_1(x)e^{j\omega t} \quad (8.77)$$

We may rewrite this equation, combining the time-dependent and time-independent terms, as

$$\left\{ D_p \frac{\partial^2 [\delta p_0(x)]}{\partial x^2} - \frac{\delta p_0(x)}{\tau_{p0}} \right\} + \left[D_p \frac{\partial^2 p_1(x)}{\partial x^2} - \frac{p_1(x)}{\tau_{p0}} - j\omega p_1(x) \right] e^{j\omega t} = 0 \quad (8.78)$$

If the ac component, $p_1(x)$, is zero, then the first bracketed term is just the differential Equation (8.10), which is identically zero. Then we have, from the second bracketed term,

$$D_p \frac{d^2 p_1(x)}{dx^2} - \frac{p_1(x)}{\tau_{p0}} - j\omega p_1(x) = 0 \quad (8.79)$$

Noting that $L_p^2 = D_p \tau_{p0}$, Equation (8.79) may be rewritten in the form

$$\frac{d^2 p_1(x)}{dx^2} - \frac{(1 + j\omega \tau_{p0})}{L_p^2} p_1(x) = 0 \quad (8.80)$$

or

$$\frac{d^2 p_1(x)}{dx^2} - C_p^2 p_1(x) = 0 \quad (8.81)$$

where

$$C_p^2 = \frac{(1 + j\omega \tau_{p0})}{L_p^2} \quad (8.82)$$

The general solution to Equation (8.81) is

$$p_1(x) = K_1 e^{-C_p x} + K_2 e^{+C_p x} \quad (8.83)$$

One boundary condition is that $p_1(x \rightarrow +\infty) = 0$, which implies that the coefficient $K_2 = 0$. Then

$$p_1(x) = K_1 e^{-C_p x} \quad (8.84)$$

Applying the boundary condition at $x = 0$ from Equation (8.74), we obtain

$$p_1(0) = K_1 = p_{dc} \left(\frac{\hat{V}_1}{V_t} \right) \quad (8.85)$$

The hole diffusion current density can be calculated at $x = 0$. This is given by

$$J_p = -eD_p \frac{\partial p_n}{\partial x} \Big|_{x=0} \quad (8.86)$$

If we consider a homogeneous semiconductor, the derivative of the hole concentration will be just the derivative of the excess hole concentration. Then

$$J_p = -eD_p \frac{\partial(\delta p_n)}{\partial x} \Big|_{x=0} = -eD_p \frac{\partial[\delta p_0(x)]}{\partial x} \Big|_{x=0} - eD_p \frac{\partial p_1(x)}{\partial x} \Big|_{x=0} e^{j\omega t} \quad (8.87)$$

We can write this equation in the form

$$J_p = J_{p0} + j_p(t) \quad (8.88)$$

where

$$J_{p0} = -eD_p \frac{\partial[\delta p_0(x)]}{\partial x} \Big|_{x=0} = \frac{eD_p p_{n0}}{L_p} \left[\exp\left(\frac{eV_0}{kT}\right) - 1 \right] \quad (8.89)$$

Equation (8.89) is the dc component of the hole diffusion current density and is exactly the same as in the ideal I - V relation derived previously.

The sinusoidal component of the diffusion current density is then found from

$$j_p(t) = \hat{J}_p e^{j\omega t} = -eD_p \frac{\partial p_1(x)}{\partial x} e^{j\omega t} \Big|_{x=0} \quad (8.90)$$

where \hat{J}_p is the current density phasor. Combining Equations (8.90), (8.84), and (8.85), we have

$$\hat{J}_p = -eD_p(-C_p) \left[p_{dc} \left(\frac{\hat{V}_1}{V_t} \right) \right] e^{-c_p x} \Big|_{x=0} \quad (8.91)$$

We can write the total ac hole current phasor as

$$\hat{I}_p = A \hat{J}_p = eAD_p C_p p_{dc} \left(\frac{\hat{V}_1}{V_t} \right) \quad (8.92)$$

where A is the cross-sectional area of the pn junction. Substituting the expression for C_p , we obtain

$$\hat{I}_p = \frac{eAD_p p_{dc}}{L_p} \sqrt{1 + j\omega\tau_{p0}} \left(\frac{\hat{V}_1}{V_t} \right) \quad (8.93)$$

If we define

$$I_{p0} = \frac{eAD_p p_{dc}}{L_p} = \frac{eAD_p p_{n0}}{L_p} \exp\left(\frac{eV_0}{kT}\right) \quad (8.94)$$

then Equation (8.93) becomes

$$\hat{I}_p = I_{p0} \sqrt{1 + j\omega\tau_{p0}} \left(\frac{\hat{V}_1}{V_t} \right) \quad (8.95)$$

Going through the same type of analysis for the minority carrier electrons in the p region, we obtain

$$\hat{I}_n = I_{n0} \sqrt{1 + j\omega\tau_{n0}} \left(\frac{\hat{V}_1}{V_t} \right) \quad (8.96)$$

where

$$I_{n0} = \frac{eAD_n n_{p0}}{L_n} \exp\left(\frac{eV_0}{kT}\right) \quad (8.97)$$

The total ac current phasor is the sum of \hat{I}_p and \hat{I}_n . The pn junction admittance is the total ac current phasor divided by the ac voltage phasor, or

$$Y = \frac{\hat{I}}{\hat{V}_1} = \frac{\hat{I}_p + \hat{I}_n}{\hat{V}_1} = \left(\frac{1}{V_t}\right) [I_{p0}\sqrt{1+j\omega\tau_{p0}} + I_{n0}\sqrt{1+j\omega\tau_{n0}}] \quad (8.98)$$

There is not a linear, lumped, finite, passive, bilateral network that can be synthesized to give this admittance function. However, we may make the following approximations. Assume that

$$\omega\tau_{p0} \ll 1 \quad (8.99a)$$

and

$$\omega\tau_{n0} \ll 1 \quad (8.99b)$$

These two assumptions imply that the frequency of the ac signal is not too large. Then we may write

$$\sqrt{1+j\omega\tau_{p0}} \approx 1 + \frac{j\omega\tau_{p0}}{2} \quad (8.100a)$$

and

$$\sqrt{1+j\omega\tau_{n0}} \approx 1 + \frac{j\omega\tau_{n0}}{2} \quad (8.100b)$$

Substituting Equations (8.100a) and (8.100b) into the admittance Equation (8.98), we obtain

$$Y = \left(\frac{1}{V_t}\right) \left[I_{p0} \left(1 + \frac{j\omega\tau_{p0}}{2}\right) + I_{n0} \left(1 + \frac{j\omega\tau_{n0}}{2}\right) \right] \quad (8.101)$$

If we combine the real and imaginary portions, we get

$$Y = \left(\frac{1}{V_t}\right) (I_{p0} + I_{n0}) + j\omega \left[\left(\frac{1}{2V_t}\right) (I_{p0}\tau_{p0} + I_{n0}\tau_{n0})\right] \quad (8.102)$$

Equation (8.102) may be written in the form

$$Y = g_d + j\omega C_d \quad (8.103)$$

The parameter g_d is called the **diffusion conductance** and is given by

$$g_d = \left(\frac{1}{V_t}\right) (I_{p0} + I_{n0}) = \frac{I_{DQ}}{V_t} \quad (8.104)$$

where I_{DQ} is the dc bias current. Equation (8.104) is exactly the same conductance as we obtained previously in Equation (8.67). The parameter C_d is called the **diffusion capacitance** and is given by

$$C_d = \left(\frac{1}{2V_t}\right) (I_{p0}\tau_{p0} + I_{n0}\tau_{n0}) \quad (8.105)$$

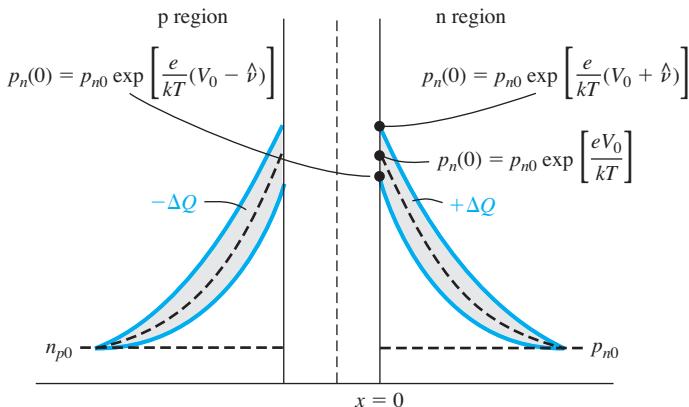


Figure 8.21 | Minority carrier concentration changes with changing forward-bias voltage.

The physics of the diffusion capacitance may be seen in Figure 8.21. The dc values of the minority carrier concentrations are shown along with the changes due to the ac component of voltage. The ΔQ charge is alternately being charged and discharged through the junction as the voltage across the junction changes. The change in the stored minority carrier charge as a function of the change in voltage is the diffusion capacitance. One consequence of the approximations $\omega\tau_{p0} \ll 1$ and $\omega\tau_{n0} \ll 1$ is that there are no “wiggles” in the minority carrier curves. The sinusoidal frequency is low enough so that the exponential curves are maintained at all times.

EXAMPLE 8.7

Objective: Calculate the small-signal admittance parameters of a pn junction diode.

This example is intended to give an indication of the magnitude of the diffusion capacitance as compared with the junction capacitance considered in the last chapter. The diffusion resistance will also be calculated. Assume that $N_a \gg N_d$ so that $p_{n0} \gg n_{p0}$. This assumption implies that $I_{p0} \gg I_{n0}$. Let $T = 300$ K, $\tau_{p0} = 10^{-7}$ s, and $I_{p0} = I_{DQ} = 1$ mA.

■ Solution

The diffusion capacitance, with these assumptions, is given by

$$C_d \approx \left(\frac{1}{2V_t}\right)(I_{p0}\tau_{p0}) = \frac{1}{(2)(0.0259)}(10^{-3})(10^{-7}) = 1.93 \times 10^{-9} \text{ F}$$

The diffusion resistance is

$$r_d = \frac{V_t}{I_{DQ}} = \frac{0.0259 \text{ V}}{1 \text{ mA}} = 25.9 \Omega$$

■ Comment

The value of 1.93 nF for the diffusion capacitance of a forward-biased pn junction is three to four orders of magnitude larger than the junction capacitance of the reverse-biased pn junction, which we calculated in Example 7.5.

EXERCISE PROBLEM

Ex 8.7 A silicon pn junction diode at $T = 300$ K has the following parameters: $N_d = 8 \times 10^{16} \text{ cm}^{-3}$, $N_a = 2 \times 10^{15} \text{ cm}^{-3}$, $D_n = 25 \text{ cm}^2/\text{s}$, $D_p = 10 \text{ cm}^2/\text{s}$, $\tau_{n0} = 5 \times 10^{-7} \text{ s}$, and $\tau_{p0} = 10^{-7} \text{ s}$. The cross-sectional area is $A = 10^{-3} \text{ cm}^2$. Determine the diffusion resistance and diffusion capacitance if the diode is forward biased at (a) $V_a = 0.550 \text{ V}$ and (b) $V_a = 0.610 \text{ V}$.

[Ans. (a) $r_d = 118 \Omega$, $C_d = 2.07 \text{ nF}$; (b) $r_d = 11.6 \Omega$, $C_d = 20.9 \text{ nF}$

The diffusion capacitance tends to dominate the capacitance terms in a forward-biased pn junction. The small-signal diffusion resistance can be fairly small if the diode current is a fairly large value. As the diode current decreases, the diffusion resistance increases. We will consider the impedance of forward-biased pn junctions again when we discuss bipolar transistors.

8.3.3 Equivalent Circuit

The small-signal equivalent circuit of the forward-biased pn junction is derived from Equation (8.103). This circuit is shown in Figure 8.22a. We need to add the junction capacitance, which will be in parallel with the diffusion resistance and diffusion capacitance. The last element we add, to complete the equivalent circuit, is a series resistance. The neutral n and p regions have finite resistances so the actual pn junction will include a series resistance. The complete equivalent circuit is given in Figure 8.22b.

The voltage across the actual junction is V_a and the total voltage applied to the pn diode is given by V_{app} . The junction voltage V_a is the voltage in the ideal current–voltage expression. We can write the expression

$$V_{\text{app}} = V_a + I r_s \quad (8.106)$$

Figure 8.23 is a plot of the current–voltage characteristic from Equation (8.106) showing the effect of the series resistance. A larger applied voltage is required to

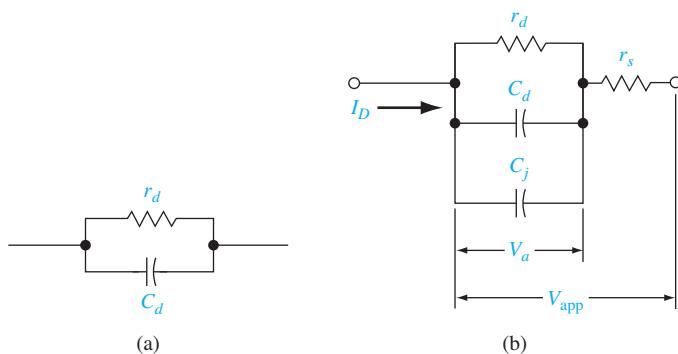


Figure 8.22 | (a) Small-signal equivalent circuit of ideal forward-biased pn junction diode; (b) complete small-signal equivalent circuit of pn junction.

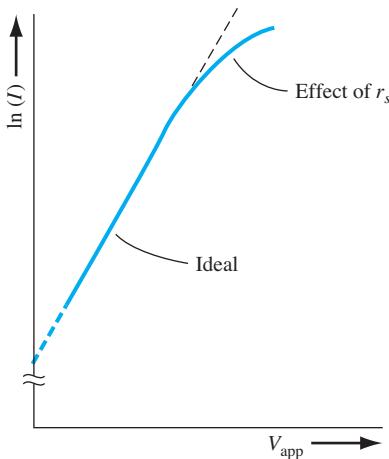


Figure 8.23 | Forward-biased I - V characteristics of a pn junction diode showing the effect of series resistance.

achieve the same current value when a series resistance is included. In most diodes, the series resistance will be negligible. In some semiconductor devices with pn junctions, however, the series resistance will be in a feedback loop; in these cases, the resistance is multiplied by a gain factor and becomes non-negligible.

TEST YOUR UNDERSTANDING

- TYU 8.5** A GaAs pn junction diode at $T = 300$ K has the same parameters given in Ex 8.7 except that $D_n = 207 \text{ cm}^2/\text{s}$ and $D_p = 9.80 \text{ cm}^2/\text{s}$. Determine the diffusion resistance and diffusion capacitance if the diode is forward biased at (a) $V_a = 0.970$ V and (b) $V_a = 1.045$ V.

[Ans. (a) $r_d = 263 \Omega$, $C_d = 0.940 \text{ nF}$; (b) $r_d = 14.6 \Omega$, $C_d = 17.0 \text{ nF}$

- TYU 8.6** A silicon pn junction diode at $T = 300$ K has the same parameters as those described in Ex 8.7. The neutral n-region and neutral p-region lengths are 0.01 cm. Estimate the series resistance of the diode (neglect ohmic contacts).
(Ans. $R_s = 99 \Omega$)

*8.4 | CHARGE STORAGE AND DIODE TRANSIENTS

The pn junction is typically used as an electrical switch. In forward bias, referred to as the *on* state, a relatively large current can be produced by a small applied voltage; in reverse bias, referred to as the *off* state, only a very small current will exist. Of primary interest in circuit applications is the speed of the pn junction diode in switching states. We qualitatively discuss the transients that occur and the charge storage effects. We simply state the equations that describe the switching times without any mathematical derivations.

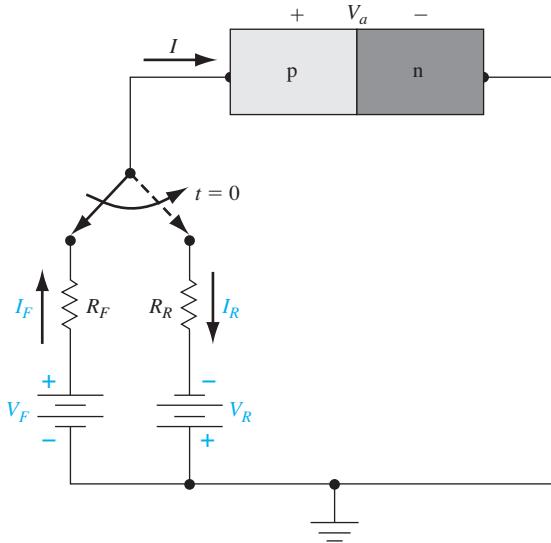


Figure 8.24 | Simple circuit for switching a diode from forward to reverse bias.

8.4.1 The Turn-off Transient

Suppose we want to switch a diode from the forward bias on state to the reverse-biased off state. Figure 8.24 shows a simple circuit that will switch the applied bias at $t = 0$. For $t < 0$, the forward-bias current is

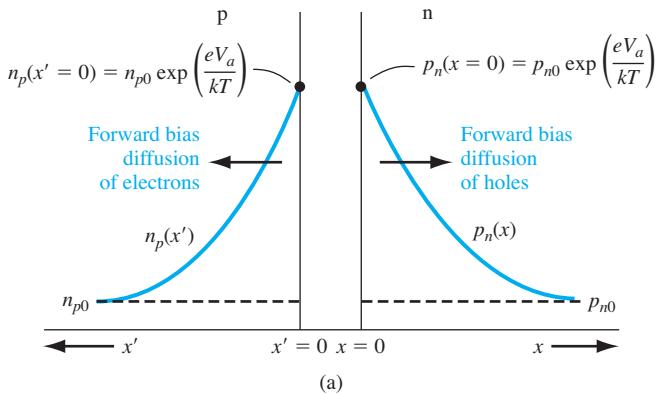
$$I = I_F = \frac{V_F - V_a}{R_F} \quad (8.107)$$

The minority carrier concentrations in the device, for the applied forward voltage V_F , are shown in Figure 8.25a. There is excess minority carrier charge stored in both the p and n regions of the diode. The excess minority carrier concentrations at the space charge edges are supported by the forward-bias junction voltage V_a . When the voltage is switched from the forward- to the reverse-biased state, the excess minority carrier concentrations at the space charge edges can no longer be supported and they start to decrease, as shown in Figure 8.25b.

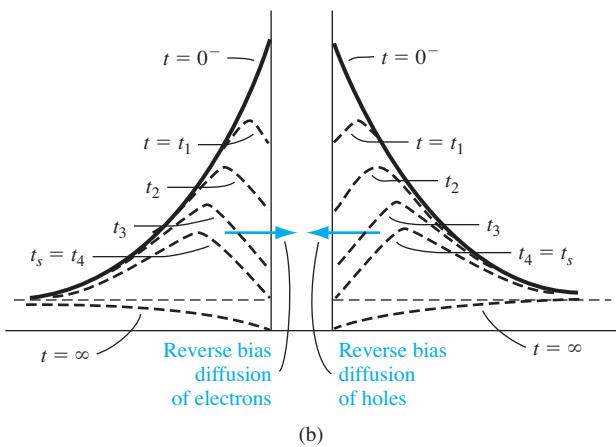
The collapse of the minority carrier concentrations at the edges of the space charge region leads to large concentration gradients and diffusion currents in the reverse-biased direction. If we assume, for the moment, that the voltage across the diode junction is small compared with V_R , then the reverse-biased current is limited to approximately

$$I = -I_R \approx \frac{-V_R}{R_R} \quad (8.108)$$

The junction capacitances do not allow the junction voltage to change instantaneously. If the current I_R were larger than this value, there would be a forward-bias voltage across the junction, which would violate our assumption of a reverse-biased current.



(a)



(b)

Figure 8.25 | (a) Steady-state forward-bias minority carrier concentrations; (b) minority carrier concentrations at various times during switching.

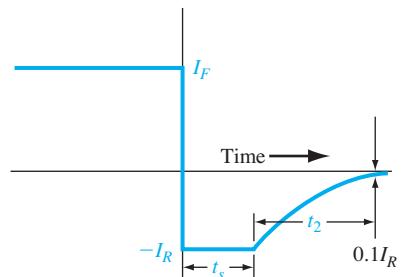


Figure 8.26 | Current characteristic versus time during diode switching.

If the current I_R were smaller than this value, there would be a reverse-biased voltage across the junction, which means that the junction voltage would have changed instantaneously. Since the reverse current is limited to the value given by Equation (8.108), the reverse-biased density gradient is constant; thus, the minority carrier concentrations at the space charge edge decrease with time as shown in Figure 8.25b.

This reverse current I_R will be approximately constant for $0^+ \leq t \leq t_s$, where t_s is called the *storage time*. The storage time is the length of time required for the minority carrier concentrations at the space charge edge to reach the thermal-equilibrium values. After this time, the voltage across the junction will begin to change. The current characteristic is shown in Figure 8.26. The reverse current is the flow of the stored minority carrier charge, which is the difference between the minority carrier concentrations at $t = 0^-$ and $t = \infty$, as shown in Figure 8.25b.

The storage time t_s can be determined by solving the time-dependent continuity equation. If we consider a one-sided p⁺n junction, the storage time is determined from the equation

$$\operatorname{erf} \sqrt{\frac{t_s}{\tau_{p0}}} = \frac{I_F}{I_F + I_R} \quad (8.109)$$

where erf (x) is known as the error function. An approximate solution for the storage time can be obtained as

$$t_s \approx \tau_{p0} \ln \left(1 + \frac{I_F}{I_R} \right) \quad (8.110)$$

The recovery phase for $t > t_s$ is the time required for the junction to reach its steady-state reverse-biased condition. The remainder of the excess charge is being removed and the space charge width is increasing to the reverse-biased value. The decay time t_2 is determined from

$$\operatorname{erf} \sqrt{\frac{t_2}{\tau_{p0}}} + \frac{\exp(-t_2/\tau_{p0})}{\sqrt{\pi t_2/\tau_{p0}}} = 1 + 0.1 \left(\frac{I_R}{I_F} \right) \quad (8.111)$$

The total turn-off time is the sum of t_s and t_2 .

To switch the diode quickly, we need to be able to produce a large reverse current as well as have a small minority carrier lifetime. In the design of diode circuits, then, the designer must provide a path for the transient reverse-biased current pulse in order to be able to switch the diode quickly. These same effects will be considered when we discuss the switching of bipolar transistors.

8.4.2 The Turn-on Transient

The turn-on transient occurs when the diode is switched from its “off” state into the forward-bias “on” state. The turn-on can be accomplished by applying a forward-bias current pulse. The first stage of turn-on occurs very quickly and is the length of time required to narrow the space charge width from the reverse-biased value to its thermal-equilibrium value when $V_a = 0$. During this time, ionized donors and acceptors are neutralized as the space charge width narrows.

The second stage of the turn-on process is the time required to establish the minority carrier distributions. During this time the voltage across the junction is increasing toward its steady-state value. A small turn-on time is achieved if the minority carrier lifetime is small and if the forward-bias current is small.

TEST YOUR UNDERSTANDING

- TYU 8.7** A one-sided p⁺n silicon diode, which has a forward-bias current of $I_F = 1.75$ mA, is switched to reverse bias with an effective reverse-biased voltage of $V_R = 2$ V and an effective series resistance of $R_R = 4$ kΩ. The minority carrier hole lifetime is 10^{-7} s. (a) Determine the storage time t_s . (b) Calculate the decay time t_2 . (c) What is the turn-off time of the diode?

[Ans. (a) 0.746×10^{-7} s; (b) 1.25×10^{-7} s; (c) $\approx 2 \times 10^{-7}$ s]

*8.5 | THE TUNNEL DIODE

The *tunnel diode* is a pn junction in which both the n and p regions are degenerately doped. As we discuss the operation of this device, we will find a region that exhibits a negative differential resistance. The tunnel diode was used in oscillator circuits in the past, but other types of solid-state devices are now used as high-frequency oscillators; thus, the tunnel diode is really only of academic interest. Nevertheless, this device does demonstrate the phenomenon of tunneling we discussed in Chapter 2.

Recall the degenerately doped semiconductors we discussed in Chapter 4: the Fermi level is in the conduction band of a degenerately doped n-type material and in the valence band of a degenerately doped p-type material. Then, even at $T = 0\text{ K}$, electrons will exist in the conduction band of the n-type material, and holes (empty states) will exist in the p-type material.

Figure 8.27 shows the energy-band diagram of a pn junction in thermal equilibrium for the case when both the n and p regions are degenerately doped. The depletion region width decreases as the doping increases and may be on the order of approximately 100 \AA for the case shown in Figure 8.27. The potential barrier at the junction can be approximated by a triangular potential barrier, as shown in Figure 8.28. This potential barrier is similar to the potential barrier used in Chapter 2 to illustrate the tunneling phenomenon. The barrier width is small and the electric field in the space charge region is quite large; thus, a finite probability exists that an electron may tunnel through the forbidden band from one side of the junction to the other.

We may qualitatively determine the current–voltage characteristics of the tunnel diode by considering the simplified energy-band diagrams in Figure 8.29.

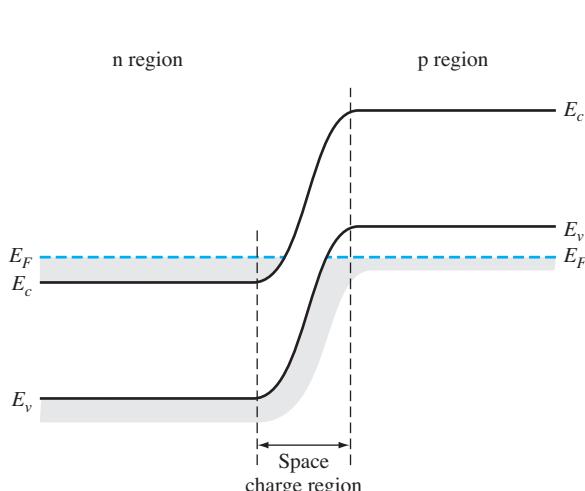


Figure 8.27 | Energy-band diagram of a pn junction in thermal equilibrium in which both the n and p regions are degenerately doped.

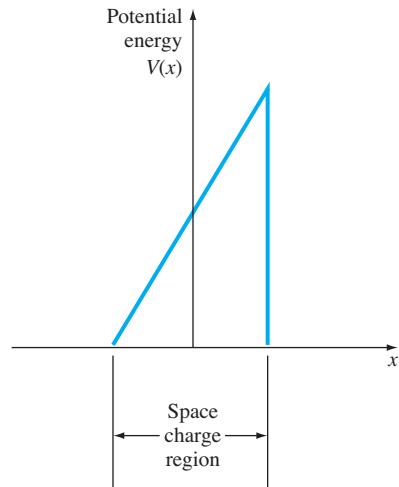


Figure 8.28 | Triangular potential barrier approximation of the potential barrier in the tunnel diode.

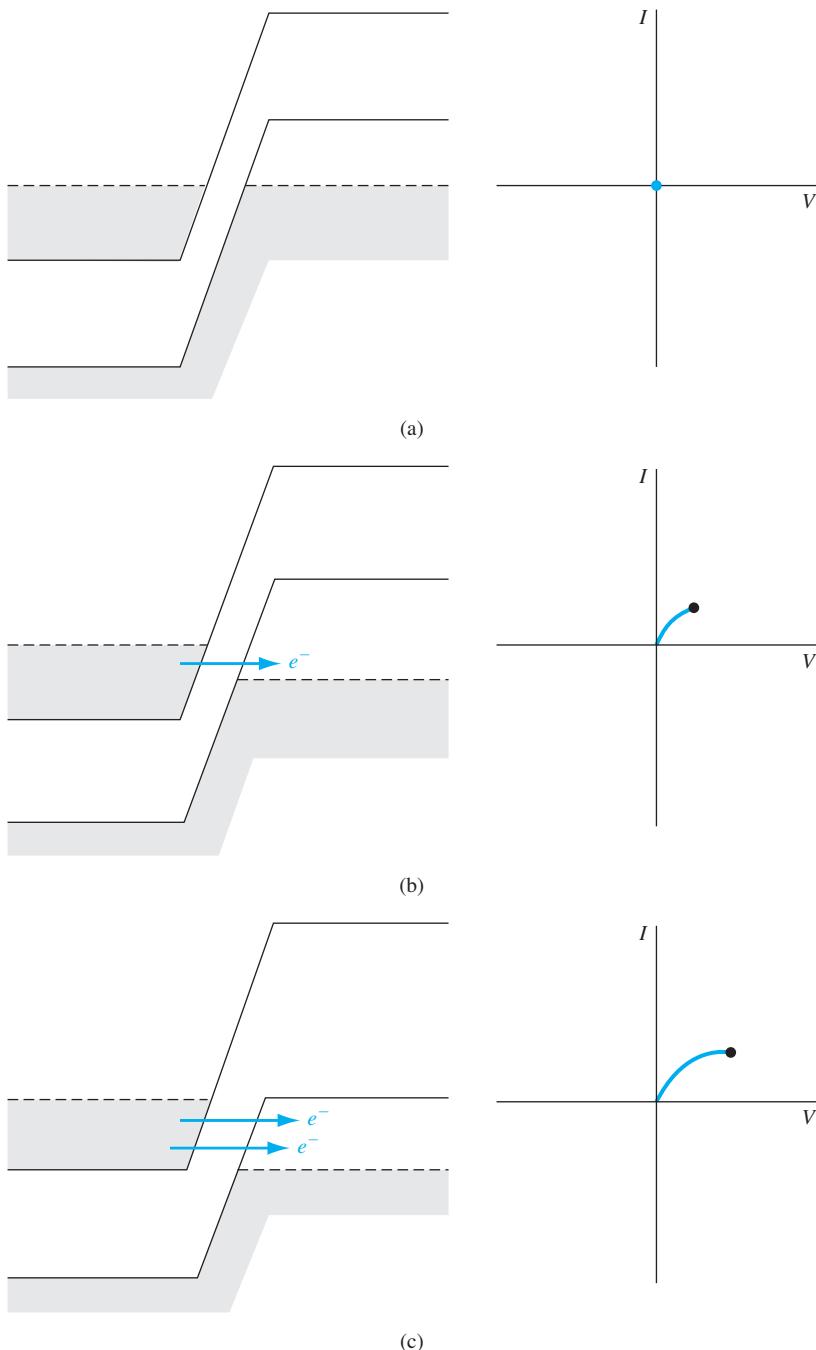


Figure 8.29 | Simplified energy-band diagrams and I - V characteristics of the tunnel diode at (a) zero bias; (b) a slight forward bias; (c) a forward bias producing maximum tunneling current. (*continued*)

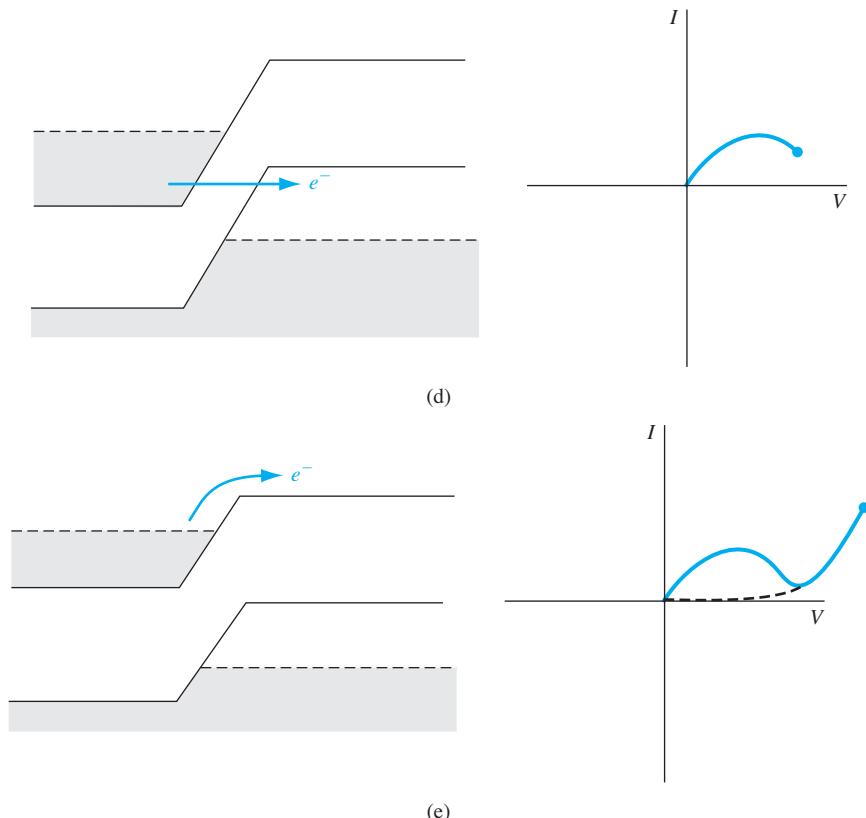


Figure 8.29 | (concluded) (d) A higher forward bias showing less tunneling current; (e) a forward bias for which the diffusion current dominates.

Figure 8.29a shows the energy-band diagram at zero bias, which produces zero current on the I - V diagram. If we assume, for simplicity, that we are near 0 K, then all energy states are filled below E_F on both sides of the junction.

Figure 8.29b shows the situation when a small forward-bias voltage is applied to the junction. Electrons in the conduction band of the n region are directly opposite to empty states in the valence band of the p region. There is a finite probability that some of these electrons will tunnel directly into the empty states, producing a forward-bias tunneling current as shown. With a slightly larger forward-bias voltage, as in Figure 8.29c, the maximum number of electrons in the n region will be opposite the maximum number of empty states in the p region; this will produce a maximum tunneling current.

As the forward-bias voltage continues to increase, the number of electrons on the n side directly opposite empty states on the p side decreases, as in Figure 8.29d, and the tunneling current will decrease. In Figure 8.29e, there are no electrons on the n side directly opposite to available empty states on the p side. For this forward-bias voltage, the tunneling current will be zero and the normal ideal diffusion current will exist in the device as shown in the I - V characteristics.

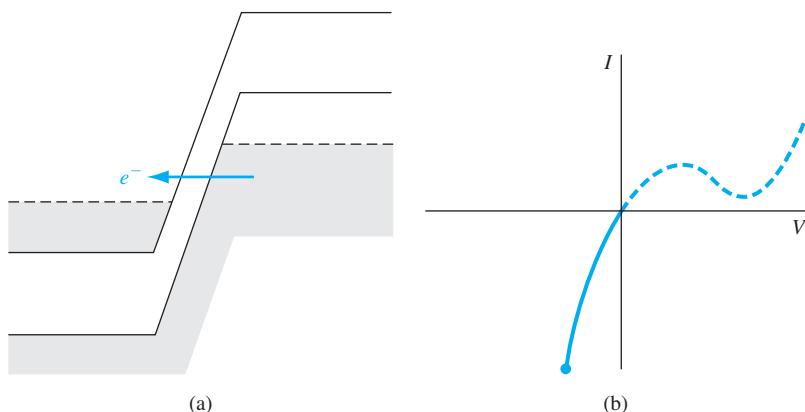


Figure 8.30 | (a) Simplified energy-band diagram of a tunnel diode with a reverse-biased voltage; (b) I - V characteristic of a tunnel diode with a reverse-biased voltage.

The portion of the curve showing a decrease in current with an increase in voltage is the region of differential negative resistance. The range of voltage and current for this region is quite small; thus, any power generated from an oscillator using this negative resistance property would also be fairly small.

A simplified energy-band diagram of the tunnel diode with an applied reverse-biased voltage is shown in Figure 8.30a. Electrons in the valence band on the p side are directly opposite empty states in the conduction band on the n side, so electrons can now tunnel directly from the p region into the n region, resulting in a large reverse-biased tunneling current. This tunneling current will exist for any reverse-biased voltage. The reverse-biased current will increase monotonically and rapidly with reverse-biased voltage as shown in Figure 8.30b.

8.6 | SUMMARY

- When a forward-bias voltage is applied across a pn junction (p region positive with respect to the n region), the potential barrier is lowered so that holes from the p region and electrons from the n region can flow across the junction.
- The boundary conditions relating the minority carrier hole concentration in the n region at the space charge edge and the minority carrier electron concentration in the p region at the space charge edge were derived.
- The holes that are injected into the n region and the electrons that are injected into the p region now become excess minority carriers. The behavior of the excess minority carrier is described by the ambipolar transport equation developed and described in Chapter 6. Solving the ambipolar transport equation and using the boundary conditions, the steady-state minority carrier hole and electron concentrations in the n region and p region, respectively, were derived.
- Gradients exist in the minority carrier hole and electron concentrations so that minority carrier diffusion currents exist in the pn junction. These diffusion currents yield the ideal current–voltage relationship of the pn junction diode.

- Excess carriers are generated in the space charge region of a reverse-biased pn junction. These carriers are swept out by the electric field and create the reverse-biased generation current that is another component of the reverse-biased diode current. Excess carriers recombine in the space charge region of a forward-biased pn junction. This recombination process creates the forward-bias recombination current that is another component of the forward-bias diode current.
- The small-signal equivalent circuit of the pn junction diode was developed. The two parameters of interest are the diffusion resistance and the diffusion capacitance.
- When a pn junction is switched from forward bias to reverse bias, the stored excess minority carrier charge must be removed from the junction. The time required to remove this charge is called the storage time and is a limiting factor in the switching speed of a diode.
- The I - V characteristics of a tunnel diode were developed showing a region of negative differential resistance.

GLOSSARY OF IMPORTANT TERMS

carrier injection The flow of carriers across the space charge region of a pn junction when a voltage is applied.

diffusion capacitance The capacitance of a forward-biased pn junction due to minority carrier storage effects.

diffusion conductance The ratio of a low-frequency, small-signal sinusoidal current to voltage in a forward-biased pn junction.

diffusion resistance The inverse of diffusion conductance.

forward bias The condition in which a positive voltage is applied to the p region with respect to the n region of a pn junction so that the potential barrier between the two regions is lowered below the thermal-equilibrium value.

generation current The reverse-biased pn junction current produced by the thermal generation of electron-hole pairs within the space charge region.

high-level injection The condition in which the excess carrier concentration becomes comparable to or greater than the majority carrier concentration.

“long” diode A pn junction diode in which both the neutral p and n regions are long compared with the respective minority carrier diffusion lengths.

recombination current The forward-bias pn junction current produced as a result of the flow of electrons and holes that recombine within the space charge region.

reverse saturation current The ideal reverse-biased current in a pn junction.

“short” diode A pn junction diode in which at least one of the neutral p or n regions is short compared to the respective minority carrier diffusion length.

storage time The time required for the excess minority carrier concentrations at the space charge edge to go from their steady-state values to zero when the diode is switched from forward to reverse bias.

CHECKPOINT

After studying this chapter, the reader should have the ability to:

- Describe the mechanism of charge flow across the space charge region of a pn junction when a forward-bias voltage is applied.

- State the boundary conditions for the minority carrier concentrations at the edge of the space charge region.
- Derive the expressions for the steady-state minority carrier concentrations in the pn junction.
- Derive the ideal current–voltage relationship for a pn junction diode.
- Describe the characteristics of a “short” diode.
- Describe generation and recombination currents in a pn junction.
- Define high-level injection and describe its effect on the diode I – V characteristics.
- Describe what is meant by diffusion resistance and diffusion capacitance.
- Describe the turn-off transient response in a pn junction.

REVIEW QUESTIONS

1. Sketch the energy bands in a zero-biased, reverse-biased, and forward-biased pn junction.
2. Write the boundary conditions for the excess minority carriers in a pn junction (a) under forward bias and (b) under reverse bias.
3. Sketch the steady-state minority carrier concentrations in a forward-biased pn junction.
4. Explain the procedure that is used in deriving the ideal current–voltage relationship in a pn junction diode.
5. Sketch the electron and hole currents through a forward-biased pn junction diode. Are currents near the junction primarily due to drift or diffusion? What about currents far from the junction?
6. What is the temperature dependence of the ideal reverse-saturation current?
7. What is meant by a “short” diode?
8. Explain the physical mechanism of the (a) generation current and (b) recombination current.
9. Sketch the forward-bias I – V characteristics of a pn junction diode showing the effects of recombination and high-level injection.
10. (a) Explain the physical mechanism of diffusion capacitance. (b) What is diffusion resistance?
11. If a forward-biased pn junction is switched off, explain what happens to the stored minority carriers. In which direction is the current immediately after the diode is switched off?

PROBLEMS

[Note: In the following problems, assume $T = 300$ K and the following parameters unless otherwise stated. For silicon pn junctions: $D_n = 25 \text{ cm}^2/\text{s}$, $D_p = 10 \text{ cm}^2/\text{s}$, $\tau_{n0} = 5 \times 10^{-7} \text{ s}$, $\tau_{p0} = 10^{-7} \text{ s}$. For GaAs pn junctions: $D_n = 205 \text{ cm}^2/\text{s}$, $D_p = 9.8 \text{ cm}^2/\text{s}$, $\tau_{n0} = 5 \times 10^{-8} \text{ s}$, $\tau_{p0} = 10^{-8} \text{ s}$.]

Section 8.1 pn Junction Current

- 8.1** (a) Consider an ideal pn junction diode at $T = 300$ K operating in the forward-bias region. Calculate the change in diode voltage that will cause a factor of 10 increase in current. (b) Repeat part (a) for a factor of 100 increase in current.

- 8.2** A silicon pn junction has impurity doping concentrations of $N_d = 2 \times 10^{15} \text{ cm}^{-3}$ and $N_a = 8 \times 10^{15} \text{ cm}^{-3}$. Determine the minority carrier concentrations at the edges of the space charge region for (a) $V_a = 0.45 \text{ V}$, (b) $V_a = 0.55 \text{ V}$, and (c) $V_a = -0.55 \text{ V}$.
- 8.3** The doping concentrations in a GaAs pn junction are $N_d = 10^{16} \text{ cm}^{-3}$ and $N_a = 4 \times 10^{16} \text{ cm}^{-3}$. Find the minority carrier concentrations at the edges of the space charge region for (a) $V_a = 0.90 \text{ V}$, (b) $V_a = 1.10 \text{ V}$, and (c) $V_a = -0.95 \text{ V}$.
- 8.4** (a) The doping concentrations in a silicon pn junction are $N_d = 5 \times 10^{15} \text{ cm}^{-3}$ and $N_a = 5 \times 10^{16} \text{ cm}^{-3}$. The minority carrier concentration at either space charge edge is to be no larger than 10 percent of the respective majority carrier concentration. (i) Determine the maximum forward-bias voltage that can be applied to the junction and still meet the required specifications. (ii) Is the n-region or p-region concentration the factor that limits the forward-bias voltage? (b) Repeat part (a) if the doping concentrations are $N_d = 3 \times 10^{16} \text{ cm}^{-3}$ and $N_a = 7 \times 10^{15} \text{ cm}^{-3}$.
- 8.5** Consider a GaAs pn junction with doping concentrations $N_a = 5 \times 10^{16} \text{ cm}^{-3}$ and $N_d = 10^{16} \text{ cm}^{-3}$. The junction cross-sectional area is $A = 10^{-3} \text{ cm}^2$ and the applied forward-bias voltage is $V_a = 1.10 \text{ V}$. Calculate the (a) minority electron diffusion current at the edge of the space charge region, (b) minority hole diffusion current at the edge of the space charge region, and (c) total current in the pn junction diode.
- 8.6** An n⁺p silicon diode at $T = 300 \text{ K}$ has the following parameters: $N_d = 10^{18} \text{ cm}^{-3}$, $N_a = 10^{16} \text{ cm}^{-3}$, $D_n = 25 \text{ cm}^2/\text{s}$, $D_p = 10 \text{ cm}^2/\text{s}$, $\tau_{n0} = \tau_{p0} = 1 \mu\text{s}$, and $A = 10^{-4} \text{ cm}^2$. Determine the diode current for (a) a forward-bias voltage of 0.5 V and (b) a reverse-biased voltage of 0.5 V.
- 8.7** An ideal germanium pn junction diode has the following parameters: $N_a = 4 \times 10^{15} \text{ cm}^{-3}$, $N_d = 2 \times 10^{17} \text{ cm}^{-3}$, $D_p = 48 \text{ cm}^2/\text{s}$, $D_n = 90 \text{ cm}^2/\text{s}$, $\tau_{p0} = \tau_{n0} = 2 \times 10^{-6} \text{ s}$, and $A = 10^{-4} \text{ cm}^2$. Determine the diode current for (a) a forward-bias voltage of 0.25 V and (b) a reverse-biased voltage of 0.25 V.
- 8.8** A one-sided p⁺n silicon diode has doping concentrations of $N_a = 5 \times 10^{17} \text{ cm}^{-3}$ and $N_d = 8 \times 10^{15} \text{ cm}^{-3}$. The minority carrier lifetimes are $\tau_{n0} = 10^{-7} \text{ s}$ and $\tau_{p0} = 8 \times 10^{-8} \text{ s}$. The cross-sectional area is $A = 2 \times 10^{-4} \text{ cm}^2$. Calculate the (a) reverse-biased saturation current, and (b) the forward-bias current at (i) $V_a = 0.45 \text{ V}$, (ii) $V_a = 0.55 \text{ V}$, and (iii) $V_a = 0.65 \text{ V}$.
- 8.9** Calculate the applied reverse-biased voltage at which the ideal reverse current in a pn junction diode at $T = 300 \text{ K}$ reaches 90 percent of its reverse-saturation current value.
- 8.10** Fill in the missing data in the following table.
- | Case | V_a (V) | I (mA) | I_s (mA) | J_s (mA/cm ²) | A (cm ²) |
|------|-----------|----------|---------------------|-----------------------------|------------------------|
| 1 | 0.65 | 0.50 | | | 2×10^{-4} |
| 2 | 0.70 | | 2×10^{-12} | | 1×10^{-3} |
| 3 | | 0.80 | | 1×10^{-7} | 1×10^{-4} |
| 4 | 0.72 | 1.20 | | | 2×10^{-8} |
- 8.11** Consider an ideal silicon pn junction diode. (a) What must be the ratio of N_d/N_a so that 90 percent of the current in the depletion region is due to the flow of electrons? (b) Repeat part (a) if 80 percent of the current in the depletion region is due to the flow of holes.
- 8.12** A silicon pn junction diode is to be designed to operate at $T = 300 \text{ K}$ such that the diode current is $I = 10 \text{ mA}$ at a diode voltage of $V_D = 0.65 \text{ V}$. The ratio of electron current to total current is to be 0.10 and the maximum current density is to be no more than 20 A/cm^2 . Use the semiconductor parameters given in Example 8.2.

- 8.13** An ideal silicon pn junction at $T = 300$ K is under forward bias. The minority carrier lifetimes are $\tau_{n0} = 10^{-6}$ s and $\tau_{p0} = 10^{-7}$ s. The doping concentration in the n region is $N_d = 10^{16}$ cm $^{-3}$. Plot the ratio of hole current to the total current crossing the space charge region as the p-region doping concentration varies over the range $10^{15} \leq N_a \leq 10^{18}$ cm $^{-3}$. (Use a log scale for the doping concentrations.)
- 8.14** For a silicon pn junction at $T = 300$ K, assume $\tau_{p0} = 0.1\tau_{n0}$ and $\mu_n = 2.4\mu_p$. The ratio of electron current crossing the depletion region to the total current is defined as the electron injection efficiency. Determine the expression for the electron injection efficiency as a function of (a) N_d/N_a and (b) the ratio of n-type conductivity to p-type conductivity.
- 8.15** A silicon pn junction with a cross-sectional area of 10^{-4} cm 2 has the following properties at $T = 300$ K:

n region	p region
$N_d = 10^{17}$ cm $^{-3}$	$N_a = 5 \times 10^{15}$ cm $^{-3}$
$\tau_{p0} = 10^{-7}$ s	$\tau_{n0} = 10^{-6}$ s
$\mu_n = 850$ cm 2 /V-s	$\mu_n = 1250$ cm 2 /V-s
$\mu_p = 320$ cm 2 /V-s	$\mu_p = 420$ cm 2 /V-s

- (a) Sketch the thermal equilibrium energy-band diagram of the pn junction, including the values of the Fermi level with respect to the intrinsic level on each side of the junction. (b) Calculate the reverse-saturation current I_s and determine the forward-bias current I at a forward-bias voltage of 0.5 V. (c) Determine the ratio of hole current to total current at the space charge edge x_n .
- 8.16** Consider an ideal silicon pn junction diode with the geometry shown in Figure P8.16. The doping concentrations are $N_a = 5 \times 10^{16}$ cm $^{-3}$ and $N_d = 1.5 \times 10^{16}$ cm $^{-3}$, and the minority carrier lifetimes are $\tau_{n0} = 2 \times 10^{-7}$ s and $\tau_{p0} = 8 \times 10^{-8}$ s. The cross-sectional area is $A = 5 \times 10^{-4}$ cm 2 . Calculate (a) the ideal reverse-saturation current due to holes, (b) the ideal reverse-saturation current due to electrons, (c) the hole concentration at $x = x_n$ for $V_a = 0.8V_{bi}$, (d) the electron current at $x = x_n$ for $V_a = 0.8 V_{bi}$, and (e) the electron current at $x = x_n + (1/2)L_p$ for $V_a = 0.8 V_{bi}$.
- 8.17** Consider the ideal long silicon pn junction shown in Figure P8.17. $T = 300$ K. The n region is doped with 10^{16} donor atoms per cm 3 and the p region is doped with 5×10^{16} acceptor atoms per cm 3 . The minority carrier lifetimes are $\tau_{n0} = 0.05$ μ s and $\tau_{p0} = 0.01$ μ s. The minority carrier diffusion coefficients are $D_n = 23$ cm 2 /s and $D_p = 8$ cm 2 /s. The forward-bias voltage is $V_a = 0.610$ V. Calculate (a) the excess hole concentration as a function of x for $x \geq 0$, (b) the hole diffusion current density at $x = 3 \times 10^{-4}$ cm, and (c) the electron current density at $x = 3 \times 10^{-4}$ cm.
- 8.18** The limit of low injection is normally defined to be when the minority carrier concentration at the edge of the space charge region in the low-doped region becomes equal

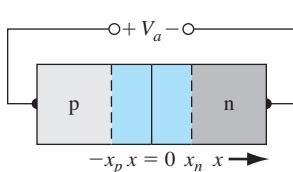


Figure P8.16 | Figure for Problem 8.16.

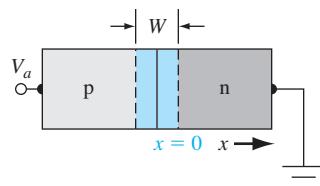


Figure P8.17 | Figure for Problem 8.17.

to one-tenth the majority carrier concentration in this region. Determine the value of the forward-bias voltage at which the limit of low injection is reached for the diode described in (a) Problem 8.7 and (b) Problem 8.8.

- 8.19** The cross-sectional area of a silicon pn junction is 10^{-3} cm^2 . The temperature of the diode is $T = 300 \text{ K}$, and the doping concentrations are $N_d = 10^{16} \text{ cm}^{-3}$ and $N_a = 8 \times 10^{15} \text{ cm}^{-3}$. Assume minority carrier lifetimes of $\tau_{n0} = 10^{-6} \text{ s}$ and $\tau_{p0} = 10^{-7} \text{ s}$. Calculate the total number of excess electrons in the p region and the total number of excess holes in the n region for (a) $V_a = 0.3 \text{ V}$, (b) $V_a = 0.4 \text{ V}$, and (c) $V_a = 0.5 \text{ V}$.
- 8.20** Consider two ideal pn junctions at $T = 300 \text{ K}$, having exactly the same electrical and physical parameters except for the bandgap energy of the semiconductor materials. The first pn junction has a bandgap energy of 0.525 eV and a forward-bias current of 10 mA with $V_a = 0.255 \text{ V}$. For the second pn junction, “design” the bandgap energy so that a forward-bias voltage of $V_a = 0.32 \text{ V}$ will produce a current of $10 \mu\text{A}$.
- 8.21** The reverse-biased saturation current is a function of temperature. (a) Assuming that I_s varies with temperature only from the intrinsic carrier concentration, show that we can write $I_s = CT^3 \exp(-E_g/kT)$ where C is a constant and a function only of the diode parameters. (b) Determine the increase in I_s as the temperature increases from $T = 300 \text{ K}$ to $T = 400 \text{ K}$ for a (i) germanium diode and (ii) silicon diode.
- 8.22** Assume that the mobilities, diffusion coefficients, and minority carrier lifetime parameters are independent of temperature (use the $T = 300 \text{ K}$ values). Assume that $\tau_{n0} = 10^{-6} \text{ s}$, $\tau_{p0} = 10^{-7} \text{ s}$, $N_d = 5 \times 10^{15} \text{ cm}^{-3}$, and $N_a = 5 \times 10^{16} \text{ cm}^{-3}$. Plot the ideal reverse saturation current density from $T = 200 \text{ K}$ to $T = 500 \text{ K}$ for (a) silicon, (b) germanium, and (c) gallium arsenide ideal pn junctions. (Use a log scale for the current density.)
- 8.23** An ideal silicon pn junction diode has a cross-sectional area of $A = 5 \times 10^{-4} \text{ cm}^2$. The doping concentrations are $N_a = 4 \times 10^{15} \text{ cm}^{-3}$ and $N_d = 2 \times 10^{17} \text{ cm}^{-3}$. Assume that $E_g = 1.12 \text{ eV}$ as well as the diffusion coefficients and lifetimes are independent of temperature. The ratio of the magnitude of forward- to reverse-biased currents is to be no less than 2×10^4 with forward- and reverse-biased voltages of 0.50 V , and the maximum reverse-biased current is to be limited to $1.2 \mu\text{A}$. Determine the maximum temperature at which the diode will meet these specifications and determine which specification is the limiting factor.
- 8.24** (a) A silicon pn junction diode has the geometry shown in Figure 8.11 in which the n region is “short” with a length $W_n = 0.7 \mu\text{m}$. The doping concentrations are $N_a = 2 \times 10^{17} \text{ cm}^{-3}$ and $N_d = 2 \times 10^{15} \text{ cm}^{-3}$. The cross-sectional area is $A = 10^{-3} \text{ cm}^2$. Determine (i) the maximum forward-bias voltage such that low injection is still valid, and (ii) the resulting current at this forward-bias voltage. (b) Repeat part (a) if the doping concentrations are reversed such that $N_a = 2 \times 10^{15} \text{ cm}^{-3}$ and $N_d = 2 \times 10^{17} \text{ cm}^{-3}$.
- *8.25** A p+n silicon diode is fabricated with a narrow n region as shown in Figure 8.11, in which $W_n < L_p$. Assume the boundary condition of $p_n = p_{n0}$ at $x = x_n + W_n$. (a) Derive the expression for the excess hole concentration $\delta p_n(x)$ as given by Equation (8.31). (b) Using the results of part (a), show that the current density in the diode is given by

$$J = \frac{eD_p p_{n0}}{L_p} \coth\left(\frac{W_n}{L_p}\right) \left[\exp\left(\frac{eV}{kT}\right) - 1 \right]$$

*Asterisks next to problems indicate problems that are more difficult.

- 8.26** A silicon diode can be used to measure temperature by operating the diode at a fixed forward-bias current. The forward-bias voltage is then a function of temperature. At $T = 300$ K, the diode voltage is found to be 0.60 V. Determine the diode voltage at (a) $T = 310$ K and (b) $T = 320$ K.
- 8.27** A forward-biased silicon diode is to be used as a temperature sensor. The diode is forward biased with a constant current source and V_a is measured as a function of temperature. (a) Derive an expression for $V_a(T)$ assuming that D/L for electrons and holes, and E_g are independent of temperature. (b) If the diode is biased at $I_D = 0.1$ mA and if $I_s = 10^{-15}$ A at $T = 300$ K, plot V_a versus T for $20^\circ\text{C} < T < 200^\circ\text{C}$. (c) Repeat part (b) if $I_D = 1$ mA. (d) Determine any changes in the results of parts (a) through (c) if the change in bandgap energy with temperature is taken into account.

Section 8.2 Generation–Recombination Currents

- 8.28** Consider a silicon pn junction diode with an applied reverse-biased voltage of $V_R = 5$ V. The doping concentrations are $N_a = N_d = 4 \times 10^{16} \text{ cm}^{-3}$ and the cross-sectional area is $A = 10^{-4} \text{ cm}^2$. Assume minority carrier lifetimes of $\tau_0 = \tau_{n0} = \tau_{p0} = 10^{-7} \text{ s}$. Calculate the (a) ideal reverse-saturation current, (b) reverse-biased generation current, and (c) the ratio of the generation current to ideal saturation current.
- 8.29** Consider the diode described in Problem 8.28. Assume that all parameters except n_i are independent of temperature. (a) Determine the temperature at which I_s and I_{gen} will be equal. What are the values of I_s and I_{gen} at this temperature? (b) Calculate the forward-bias voltage at which the ideal diffusion current is equal to the recombination current at $T = 300$ K.
- 8.30** Consider a GaAs pn junction diode with a cross-sectional area of $A = 2 \times 10^{-4} \text{ cm}^2$ and doping concentrations of $N_a = N_d = 7 \times 10^{16} \text{ cm}^{-3}$. The electron and hole mobility values are $\mu_n = 5500 \text{ cm}^2/\text{V}\cdot\text{s}$ and $\mu_p = 220 \text{ cm}^2/\text{V}\cdot\text{s}$, respectively, and the lifetime values are $\tau_0 = \tau_{n0} = \tau_{p0} = 2 \times 10^{-8} \text{ s}$. (a) Calculate the ideal diode current at a (i) reverse-biased voltage of $V_R = 3$ V, (ii) forward-bias voltage of $V_a = 0.6$ V, (iii) forward-bias voltage of $V_a = 0.8$ V, and (iv) forward-bias voltage of $V_a = 1.0$ V. (b) (i) Calculate the generation current at $V_R = 3$ V. Assuming the recombination current extrapolated to $V_a = 0$ is $I_{ro} = 6 \times 10^{-14}$ A, determine the generation current at (ii) $V_a = 0.6$ V, (iii) $V_a = 0.8$ V, and (iv) $V_a = 1.0$ V.
- 8.31** Consider the pn junction diode described in Problem 8.30. Plot the diode recombination current and the ideal diode current (on a log scale) versus forward-bias voltage over the range $0.1 \leq V_a \leq 1.0$ V.
- 8.32** A silicon pn junction diode at $T = 300$ K has the following parameters: $N_a = N_d = 10^{16} \text{ cm}^{-3}$, $\tau_{p0} = \tau_{n0} = \tau_0 = 5 \times 10^{-7} \text{ s}$, $D_p = 10 \text{ cm}^2/\text{s}$, $D_n = 25 \text{ cm}^2/\text{s}$, and a cross-sectional area of 10^{-4} cm^2 . Plot the diode recombination current and the ideal diode current (on a log scale) versus forward-bias voltage over the range $0.1 \leq V_a \leq 0.6$ V.
- 8.33** Consider a GaAs pn diode at $T = 300$ K with $N_a = N_d = 10^{17} \text{ cm}^{-3}$ and with a cross-sectional area of $5 \times 10^{-3} \text{ cm}^2$. The minority carrier mobilities are $\mu_n = 3500 \text{ cm}^2/\text{V}\cdot\text{s}$ and $\mu_p = 220 \text{ cm}^2/\text{V}\cdot\text{s}$. The electron–hole lifetimes are $\tau_{n0} = \tau_{p0} = \tau_0 = 10^{-8} \text{ s}$. Plot the diode forward-bias current including recombination current between diode voltages of $0.1 \leq V_D \leq 1.0$ V. Compare this plot to that for an ideal diode.
- *8.34** Starting with Equation (8.44) and using the suitable approximations, show that the maximum recombination rate in a forward-biased pn junction is given by Equation (8.52).

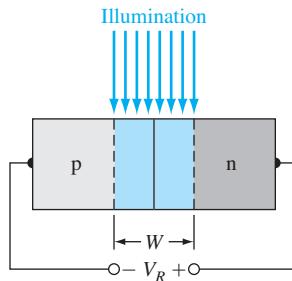


Figure P8.35 | Figure for Problem 8.35 and 8.36.

- 8.35** Consider, as shown in Figure P8.35, a uniformly doped silicon pn junction at $T = 300$ K with impurity doping concentrations of $N_a = N_d = 5 \times 10^{15} \text{ cm}^{-3}$ and minority carrier lifetimes of $\tau_{n0} = \tau_{p0} = \tau_0 = 10^{-7} \text{ s}$. A reverse-biased voltage of $V_R = 10$ V is applied. A light source is incident only on the space charge region, producing an excess carrier generation rate of $g' = 4 \times 10^{19} \text{ cm}^{-3} \text{ s}^{-1}$. Calculate the generation current density.
- 8.36** A long silicon pn junction diode has the following parameters: $N_d = 10^{18} \text{ cm}^{-3}$, $N_a = 3 \times 10^{16} \text{ cm}^{-3}$, $\tau_{n0} = \tau_{p0} = \tau_0 = 10^{-7} \text{ s}$, $D_n = 18 \text{ cm}^2/\text{s}$, and $D_p = 6 \text{ cm}^2/\text{s}$. A light source is incident on the space charge region such as shown in Figure P8.35, producing a generation current density of $J_G = 25 \text{ mA/cm}^2$. The diode is open circuited. The generation current density forward biases the junction, inducing a forward-bias current in the opposite direction to the generation current. A steady-state condition is reached when the generation current density and forward-bias current density are equal in magnitude. What is the induced forward-bias voltage at this steady-state condition?

Section 8.3 Small-Signal Model of the pn Junction

- 8.37** (a) Calculate the small-signal diffusion capacitance and diffusion resistance of a silicon pn junction diode biased at $I_{DQ} = 1.2$ mA. Assume the minority carrier lifetimes are $0.5 \mu\text{s}$ in both the n and p regions. (b) Repeat part (a) for the case when the diode is biased at $I_{DQ} = 0.12$ mA.
- 8.38** Consider the diode described in Problem 8.37. A sinusoidal signal voltage with a peak value of 50 mV is superimposed on the dc forward-bias voltage. Determine the magnitude of charge that is being alternately charged and discharged in the n region.
- 8.39** Consider a p^+n silicon diode at $T = 300$ K. The diode is forward biased at a current of 1 mA. The hole lifetime in the n region is 10^{-7} s. Neglecting the depletion capacitance, calculate the diode impedance at frequencies of 10 kHz, 100 kHz, 1 MHz, and 10 MHz.
- 8.40** Consider a silicon pn junction with parameters as described in Problem 8.8. (a) Calculate and plot the depletion capacitance and diffusion capacitance over the voltage range $-10 \leq V_a \leq 0.75$ V. (b) Determine the voltage at which the two capacitances are equal.
- 8.41** Consider a p^+n silicon diode at $T = 300$ K. The slope of the diffusion capacitance versus forward-bias current is $2.5 \times 10^{-6} \text{ F/A}$. Determine the hole lifetime and the diffusion capacitance at a forward-bias current of 1 mA.

- 8.42** A one-sided p⁺n silicon diode has doping concentrations of $N_a = 4 \times 10^{17} \text{ cm}^{-3}$ and $N_d = 8 \times 10^{15} \text{ cm}^{-3}$. The diode cross-sectional area is $A = 5 \times 10^{-4} \text{ cm}^2$. (a) The maximum diffusion capacitance is to be limited to 1 nF. Determine (i) the maximum current through the diode, (ii) the maximum forward-bias voltage, and (iii) the diffusion resistance. (b) Repeat part (a) if the maximum diffusion capacitance is limited to 0.25 nF.
- 8.43** A silicon pn junction diode at $T = 300 \text{ K}$ has a cross-sectional area of 10^{-2} cm^2 . The length of the p region is 0.2 cm and the length of the n region is 0.1 cm. The doping concentrations are $N_d = 10^{15} \text{ cm}^{-3}$ and $N_a = 10^{16} \text{ cm}^{-3}$. Determine (a) approximately the series resistance of the diode and (b) the current through the diode that will produce a 0.1 V drop across this series resistance.
- 8.44** We want to consider the effect of a series resistance on the forward-bias voltage required to achieve a particular diode current. (a) Assume the reverse-saturation current in a diode is $I_s = 10^{-10} \text{ A}$ at $T = 300 \text{ K}$. The resistivity of the n region is $0.2 \Omega\text{-cm}$ and the resistivity of the p region is $0.1 \Omega\text{-cm}$. Assume the length of each neutral region is 10^{-2} cm and the cross-sectional area is $2 \times 10^{-5} \text{ cm}^2$. Determine the required applied voltage to achieve a current of (i) 1 mA and (ii) 10 mA. (b) Repeat part (a) neglecting the series resistance.
- 8.45** (a) The reverse-saturation current in a diode is $I_s = 5 \times 10^{-12} \text{ A}$. The maximum small-signal diffusion resistance is to be $r_d = 32 \Omega$. Determine the minimum forward-bias voltage that can be applied to meet this specification. (b) Repeat part (a) if the maximum small-signal diffusion resistance is to be $r_d = 60 \Omega$.
- 8.46** (a) An ideal silicon pn junction diode at $T = 300 \text{ K}$ is forward biased at $V_a = +20 \text{ mV}$. The reverse-saturation current is $I_s = 10^{-13} \text{ A}$. Calculate the small-signal diffusion resistance. (b) Repeat part (a) for an applied reverse-biased voltage of $V_a = -20 \text{ mV}$.

Section 8.4 Charge Storage and Diode Transients

- 8.47** (a) In switching a pn junction from forward to reverse bias, assume that the ratio of reverse current, I_R , to forward current, I_F , is 0.2. Determine the ratio of storage time to minority carrier lifetime, t_s/τ_{p0} . (b) Repeat part (a) if the ratio of I_R to I_F is 1.0.
- 8.48** A pn junction is switched from forward to reverse bias. The storage time is specified to be $\tau_s = 0.3 \tau_{p0}$. (a) Determine the required ratio of I_R to I_F to meet this specification. (b) Determine t_2/τ_{p0} .
- 8.49** Consider a diode with a junction capacitance of 18 pF at zero bias and 4.2 pF at a reverse-biased voltage of $V_R = 10 \text{ V}$. The minority carrier lifetimes are 10^{-7} s . The diode is switched from a forward bias with a current of 2 mA to a reverse-biased voltage of 10 V applied through a 10 kΩ resistor. Estimate the turn-off time.

Section 8.5 The Tunnel Diode

- 8.50** Consider a silicon pn junction at $T = 300 \text{ K}$ with doping concentration of $N_d = N_a = 5 \times 10^{19} \text{ cm}^{-3}$. Assuming the abrupt junction approximation is valid, determine the space charge width at a forward-bias voltage of $V_a = 0.40 \text{ V}$.
- 8.51** Sketch the energy-band diagram of an abrupt pn junction under zero bias in which the p region is degenerately doped and $E_C = E_F$ in the n region. Sketch the forward- and reverse-biased current-voltage characteristics. This diode is sometimes called a *backward diode*. Why?

Summary and Review

- 8.52** (a) Explain physically why the diffusion capacitance is not important in a reverse-biased pn junction. (b) Consider a silicon, a germanium, and gallium arsenide pn junction. If the total current density is the same in each diode under forward bias, discuss the expected relative values of electron and hole current densities.
- *8.53** A silicon p⁺n junction diode is to be designed to have a breakdown voltage of at least 60 V and to have a forward-bias current of $I_D = 50$ mA while still operating under low injection. The minority carrier lifetimes are $\tau_0 = \tau_{n0} = \tau_{p0} = 2 \times 10^{-7}$ s. Determine the doping concentrations and the minimum cross-sectional area.
- *8.54** The donor and acceptor concentrations on either side of a silicon step junction are equal. (a) Derive an expression for the breakdown voltage in terms of the critical electric field and doping concentration. (b) If the breakdown voltage is to be $V_B = 50$ V, specify the range of allowed doping concentrations.

READING LIST

1. Dimitrijev, S. *Principles of Semiconductor Devices*. New York: Oxford University Press, 2006.
2. Kano, K. *Semiconductor Devices*. Upper Saddle River, NJ: Prentice Hall, 1998.
3. Muller, R. S., and T. I. Kamins. *Device Electronics for Integrated Circuits*, 2nd ed. New York: John Wiley and Sons, 1986.
4. Neudeck, G. W. *The PN Junction Diode*. Vol. 2 of the *Modular Series on Solid State Devices*. 2nd ed. Reading, MA: Addison-Wesley, 1989.
- *5. Ng, K. K. *Complete Guide to Semiconductor Devices*. New York: McGraw-Hill, 1995.
6. Pierret, R. F. *Semiconductor Device Fundamentals*. Reading, MA: Addison-Wesley Publishing Co., 1996.
7. Roulston, D. J. *An Introduction to the Physics of Semiconductor Devices*. New York: Oxford University Press, 1999.
8. Shur, M. *Introduction to Electronic Devices*. New York: John Wiley and Sons, 1996.
- *9. Shur, M. *Physics of Semiconductor Devices*. Englewood Cliffs, NJ: Prentice Hall, 1990.
10. Streetman, B. G., and S. K. Banerjee. *Solid State Electronic Devices*, 6th ed. Upper Saddle River, NJ: Pearson Prentice Hall, 2006.
11. Sze, S. M. and K. K. Ng. *Physics of Semiconductor Devices*, 3rd ed. Hoboken, NJ: John Wiley and Sons, 2007.
12. Sze, S. M. *Semiconductor Devices: Physics and Technology*, 2nd ed. New York: John Wiley and Sons, 2001.
- *13. Wang, S. *Fundamentals of Semiconductor Theory and Device Physics*. Englewood Cliffs, NJ: Prentice Hall, 1989.
14. Yang, E. S. *Microelectronic Devices*. New York: McGraw-Hill, 1988.

*Indicates references that are at an advanced level compared to this text.

C H A P T E R

14

Optical Devices

In previous chapters, we have considered the basic physics of transistors that are used to amplify or switch electrical signals. Semiconductor devices can be designed to convert optical energy into electrical energy, and to convert electrical signals into optical signals. These devices are used in broadband communications and data transmission over optical fibers. The general classification of these devices is called *optoelectronics*.

In this chapter, we discuss the basic principles of solar cells, several photodetectors, light emitting diodes, and laser diodes. Solar cells and photodetectors convert optical energy into electrical energy; light emitting diodes and laser diodes convert electrical signals into optical signals. ■

14.0 | PREVIEW

In this chapter, we will:

- Discuss and analyze photon absorption in a semiconductor and present absorption coefficient data for several semiconductor materials.
- Consider the basic principles of solar cells, analyze their I - V characteristics, and discuss the conversion efficiency.
- Present various types of solar cells, including homojunction, heterojunction, and amorphous silicon solar cells.
- Discuss the basic principles of photodetectors, including photoconductors, photodiodes, and phototransistors.
- Derive the output current characteristics of the various photodetectors.
- Present and analyze the basic operation of the Light Emitting Diode (LED).
- Discuss the basic principles and operation of the laser diode.

14.1 | OPTICAL ABSORPTION

In Chapter 2, we discussed the wave–particle duality principle and indicated that light waves could be treated as particles, which are referred to as photons. The energy of a photon is $E = h\nu$ where h is Plank's constant and ν is the frequency. We can also relate the wavelength and energy by

$$\lambda = \frac{c}{\nu} = \frac{hc}{E} = \frac{1.24}{E} \mu\text{m} \quad (14.1)$$

where E is the photon energy in eV and c is the speed of light.

There are several possible photon–semiconductor interaction mechanisms. For example, photons can interact with the semiconductor lattice whereby the photon energy is converted into heat. Photons can also interact with impurity atoms, either donors or acceptors, or they can interact with defects within the semiconductor. However, the basic photon interaction process of greatest interest is the interaction with valence electrons. When a photon collides with a valence electron, enough energy may be imparted to elevate the electron into the conduction band. Such a process generates electron–hole pairs and creates excess carrier concentrations. The behavior of excess carriers in a semiconductor was considered in Chapter 6.

14.1.1 Photon Absorption Coefficient

When a semiconductor is illuminated with light, the photons may be absorbed or they may propagate through the semiconductor, depending on the photon energy and on the bandgap energy E_g . If the photon energy is less than E_g , the photons are not readily absorbed. In this case, the light is transmitted through the material and the semiconductor appears to be transparent.

If $E = h\nu > E_g$, the photon can interact with a valence electron and elevate the electron into the conduction band. The valence band contains many electrons and the conduction band contains many empty states, so the probability of this interaction is high when $h\nu > E_g$. This interaction creates an electron in the conduction band and a hole in the valence band—an electron–hole pair. The basic absorption processes for different values of $h\nu$ are shown in Figure 14.1. When $h\nu > E_g$, an electron–hole

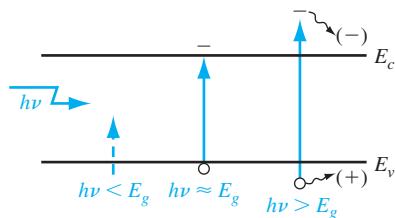


Figure 14.1 | Optically generated electron–hole pair formation in a semiconductor.

pair is created and the excess energy may give the electron or hole additional kinetic energy, which will be dissipated as heat in the semiconductor.

The intensity of the photon flux is denoted by $I_\nu(x)$ and is expressed in terms of energy/cm²-s. Figure 14.2 shows an incident photon intensity at a position x and the photon flux emerging at a distance $x + dx$. The energy absorbed per unit time in the distance dx is given by

$$\alpha I_\nu(x) dx \quad (14.2)$$

where α is the absorption coefficient. The absorption coefficient is the relative number of photons absorbed per unit distance, given in units of cm⁻¹.

From Figure 14.2, we can write

$$I_\nu(x + dx) - I_\nu(x) = \frac{dI_\nu(x)}{dx} \cdot dx = -\alpha I_\nu(x) dx \quad (14.3)$$

or

$$\frac{dI_\nu(x)}{dx} = -\alpha I_\nu(x) \quad (14.4)$$

If the initial condition is given as $I_\nu(0) = I_{\nu0}$, then the solution to the differential equation, Equation (14.4), is

$$I_\nu(x) = I_{\nu0} e^{-\alpha x} \quad (14.5)$$

The intensity of the photon flux decreases exponentially with distance through the semiconductor material. The photon intensity as a function of x for two general values of absorption coefficient is shown in Figure 14.3. If the absorption coefficient is large, the photons are absorbed over a relatively short distance.

The absorption coefficient in the semiconductor is a very strong function of photon energy and bandgap energy. Figure 14.4 shows the absorption coefficient α plotted as a function of wavelength for several semiconductor materials. The absorption coefficient increases very rapidly for $h\nu > E_g$, or for $\lambda < 1.24/E_g$. The absorption

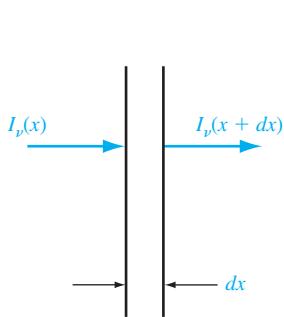


Figure 14.2 | Optical absorption in a differential length.

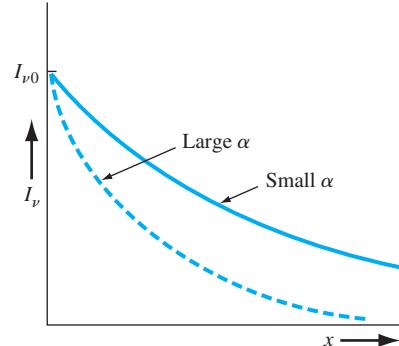


Figure 14.3 | Photon intensity versus distance for two absorption coefficients.

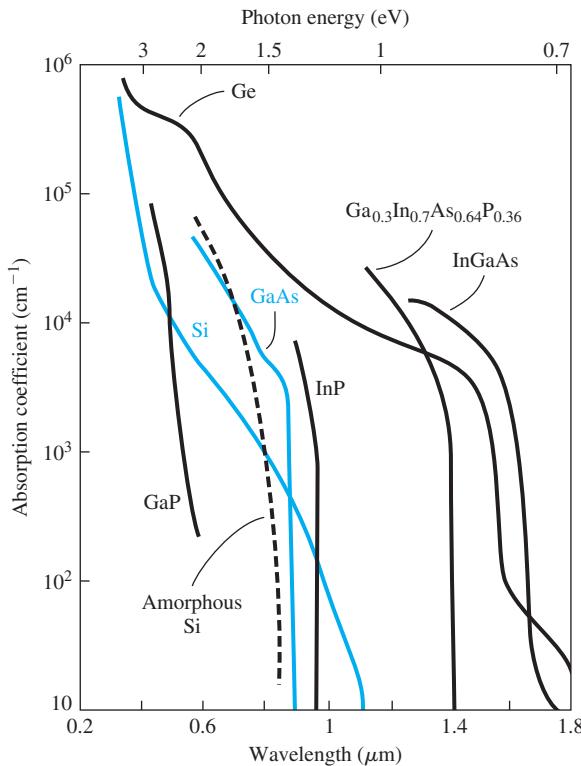


Figure 14.4 | Absorption coefficient as a function of wavelength for several semiconductors.
(From Shur [13].)

coefficients are very small for $h\nu < E_g$, so the semiconductor appears transparent to photons in this energy range.

Objective: Calculate the thickness of a semiconductor that will absorb 90 percent of the incident photon energy.

EXAMPLE 14.1

Consider silicon and assume that in the first case the incident wavelength is $\lambda = 1.0 \mu\text{m}$ and in the second case, the incident wavelength is $\lambda = 0.5 \mu\text{m}$.

■ Solution

From Figure 14.4, the absorption coefficient is $\alpha \approx 10^2 \text{ cm}^{-1}$ for $\lambda = 1.0 \mu\text{m}$. If 90 percent of the incident flux is to be absorbed in a distance d , then the flux emerging at $x = d$ will be 10 percent of the incident flux. We can write

$$\frac{I_\nu(d)}{I_{\nu 0}} = 0.1 = e^{-\alpha d}$$

Solving for the distance d , we have

$$d = \frac{1}{\alpha} \ln \left(\frac{1}{0.1} \right) = \frac{1}{10^2} \ln (10) = 0.0230 \text{ cm}$$

In the second case, the absorption coefficient is $\alpha \approx 10^4 \text{ cm}^{-1}$ for $\lambda = 0.5 \mu\text{m}$. The distance d , then, in which 90 percent of the incident flux is absorbed, is

$$d = \frac{1}{10^4} \ln \left(\frac{1}{0.1} \right) = 2.30 \times 10^{-4} \text{ cm} = 2.30 \mu\text{m}$$

Comment

As the incident photon energy increases, the absorption coefficient increases rapidly, so that the photon energy can be totally absorbed in a very narrow region at the surface of the semiconductor.

EXERCISE PROBLEM

- Ex 14.1** Consider a slab of silicon 5 μm thick. Determine the percentage of photon energy that will pass through the slab if the photon wavelength is (a) $\lambda = 0.8 \mu\text{m}$ and (b) $\lambda = 0.6 \mu\text{m}$.

[Ans. (a) 10.09%; (b) 5.2%]

The relation between the bandgap energies of some of the common semiconductor materials and the light spectrum is shown in Figure 14.5. We may note that silicon and gallium arsenide will absorb all of the visible spectrum, whereas gallium phosphide, for example, will be transparent to the red spectrum.

14.1.2 Electron–Hole Pair Generation Rate

We have shown that photons with energy greater than E_g can be absorbed in a semiconductor, thereby creating electron–hole pairs. The intensity $I_\nu(x)$ is in units of

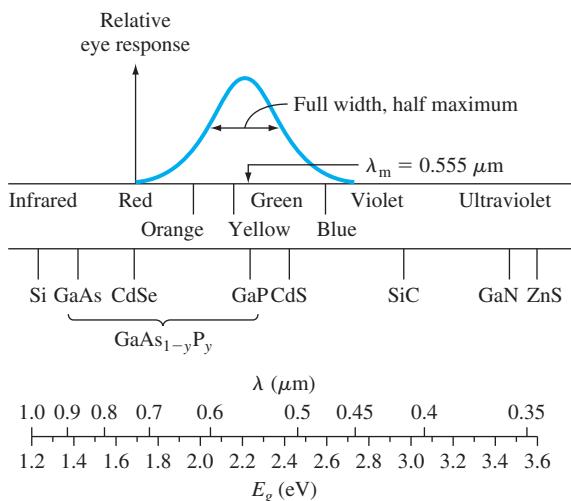


Figure 14.5 | Light spectrum versus wavelength and energy. Figure includes relative response of the human eye. (From Sze [18].)

energy/cm²-s and $\alpha I_\nu(x)$ is the rate at which energy is absorbed per unit volume. If we assume that one absorbed photon at an energy $h\nu$ creates one electron–hole pair, then the generation rate of electron–hole pairs is

$$g' = \frac{\alpha I_\nu(x)}{h\nu} \quad (14.6)$$

which is in units of #/cm³-s. We may note that the ratio $I_\nu(x)/h\nu$ is the photon flux. If, on the average, one absorbed photon produces less than one electron–hole pair, then Equation (14.6) must be multiplied by an efficiency factor.

Objective: Calculate the generation rate of electron–hole pairs given an incident intensity of photons.

EXAMPLE 14.2

Consider gallium arsenide at $T = 300$ K. Assume the photon intensity at a particular point is $I_\nu(x) = 0.05$ W/cm² at a wavelength of $\lambda = 0.75$ μm. This intensity is typical of sunlight, for example.

Solution

The absorption coefficient for gallium arsenide at this wavelength is $\alpha \approx 0.9 \times 10^4$ cm⁻¹. The photon energy, using Equation (14.1), is

$$E = h\nu = \frac{1.24}{0.75} = 1.65 \text{ eV}$$

Then, from Equation (14.6) and including the conversion factor between joules and eV, we have, for a unity efficiency factor,

$$g' = \frac{\alpha I_\nu(x)}{h\nu} = \frac{(0.9 \times 10^4)(0.05)}{(1.6 \times 10^{-19})(1.65)} = 1.70 \times 10^{21} \text{ cm}^{-3}\text{-s}^{-1}$$

If the incident photon intensity is a steady-state intensity, then, from Chapter 6, the steady-state excess carrier concentration is $\delta n = g'\tau$, where τ is the excess minority carrier lifetime. If $\tau = 10^{-7}$ s, for example, then

$$\delta n = (1.70 \times 10^{21})(10^{-7}) = 1.70 \times 10^{14} \text{ cm}^{-3}$$

Comment

This example gives an indication of the magnitude of the electron–hole generation rate and the magnitude of the excess carrier concentration. Obviously, as the photon intensity decreases with distance in the semiconductor, the generation rate also decreases.

EXERCISE PROBLEM

Ex 14.2 A photon flux with an intensity of $I_{\nu 0} = 0.10$ W/cm² and at a wavelength of $\lambda = 1$ μm is incident on the surface of silicon. Neglecting any reflection from the surface, determine the generation rate of electron–hole pairs at a depth of
 (a) $x = 5$ μm and (b) $x = 20$ μm from the surface.

[Ans. (a) 4.79×10^{19} cm⁻³s⁻¹; (b) 4.13×10^{19} cm⁻³s⁻¹]

TEST YOUR UNDERSTANDING

- TYU 14.1** (a) A photon flux with an intensity of $I_{r0} = 0.10 \text{ W/cm}^2$ is incident on the surface of silicon. The wavelength of the incident photon signal is $\lambda = 1 \mu\text{m}$. Neglecting any reflection from the surface, determine the photon flux intensity at a depth of (i) $x = 5 \mu\text{m}$ and (ii) $x = 20 \mu\text{m}$ from the surface. (b) Repeat part (a) for a wavelength of $\lambda = 0.60 \mu\text{m}$. [Ans. (i) 0.0951 W/cm^2 , (ii) 0.0135 W/cm^2 ; (b) (i) 0.01819 W/cm^2 , (ii) $3.35 \times 10^{-5} \text{ W/cm}^2$]

14.2 | SOLAR CELLS

A solar cell is a pn junction device with no voltage directly applied across the junction. The solar cell converts photon power into electrical power and delivers this power to a load. These devices have long been used for the power supply of satellites and space vehicles, and also as the power supply to some calculators. We will first consider the simple pn junction solar cell with uniform generation of excess carriers. We will also discuss briefly the heterojunction and amorphous silicon solar cells.

14.2.1 The pn Junction Solar Cell

Consider the pn junction shown in Figure 14.6 with a resistive load. Even with zero bias applied to the junction, an electric field exists in the space charge region as shown in the figure. Incident photon illumination can create electron–hole pairs in the space charge region that will be swept out producing the photocurrent I_L in the reverse-biased direction as shown.

The photocurrent I_L produces a voltage drop across the resistive load which forward biases the pn junction. The forward-bias voltage produces a forward-bias current I_F as indicated in the figure. The net pn junction current, in the reverse-biased direction, is

$$I = I_L - I_F = I_L - I_S \left[\exp \left(\frac{eV}{kT} \right) - 1 \right] \quad (14.7)$$

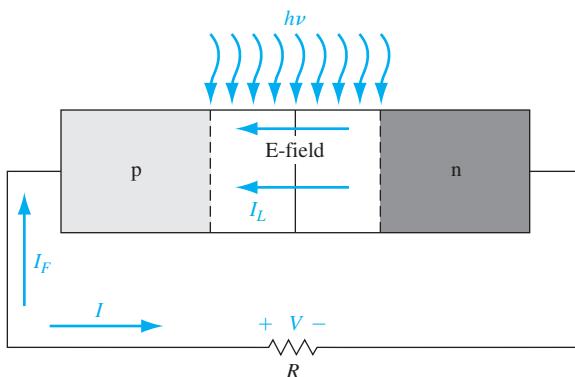


Figure 14.6 | A pn junction solar cell with resistive load.

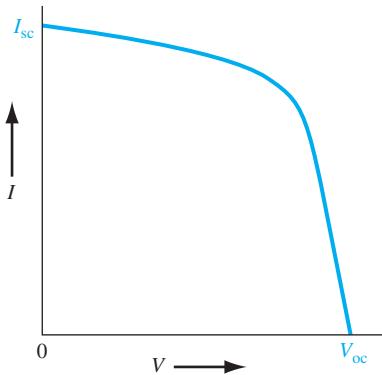


Figure 14.7 | I - V characteristics of a pn junction solar cell.

where the ideal diode equation has been used. As the diode becomes forward biased, the magnitude of the electric field in the space charge region decreases, but does not go to zero or change direction. The photocurrent is always in the reverse-biased direction and the net solar cell current is also always in the reverse-biased direction.

There are two limiting cases of interest. The short-circuit condition occurs when $R = 0$ so that $V = 0$. The current in this case is referred to as the *short-circuit current*, or

$$I = I_{sc} = I_L \quad (14.8)$$

The second limiting case is the open-circuit condition and occurs when $R \rightarrow \infty$. The net current is zero and the voltage produced is the *open-circuit voltage*. The photocurrent is just balanced by the forward-biased junction current, so we have

$$I = 0 = I_L - I_S \left[\exp\left(\frac{eV_{oc}}{kT}\right) - 1 \right] \quad (14.9)$$

We can find the open circuit voltage V_{oc} as

$$V_{oc} = V_t \ln \left(1 + \frac{I_L}{I_S} \right) \quad (14.10)$$

A plot of the diode current I as a function of the diode voltage V from Equation (14.7) is shown in Figure 14.7. We may note the short-circuit current and open-circuit voltage points on the figure.

Objective: Calculate the open-circuit voltage of a silicon pn junction solar cell.

Consider a silicon pn junction at $T = 300$ K with the following parameters:

$$\begin{aligned} N_a &= 5 \times 10^{18} \text{ cm}^{-3} & N_d &= 10^{16} \text{ cm}^{-3} \\ D_n &= 25 \text{ cm}^2/\text{s} & D_p &= 10 \text{ cm}^2/\text{s} \\ \tau_{n0} &= 5 \times 10^{-7} \text{ s} & \tau_{p0} &= 10^{-7} \text{ s} \end{aligned}$$

Let the photocurrent density be $J_L = I_L/A = 15 \text{ mA/cm}^2$.

EXAMPLE 14.3

■ Solution

We have that

$$J_s = \frac{I_s}{A} = \left(\frac{eD_n n_{p0}}{L_n} + \frac{eD_p p_{n0}}{L_p} \right) = en_i^2 \left(\frac{D_n}{L_n N_a} + \frac{D_p}{L_p N_d} \right)$$

We may calculate

$$L_n = \sqrt{D_n \tau_{n0}} = \sqrt{(25)(5 \times 10^{-7})} = 35.4 \mu\text{m}$$

and

$$L_p = \sqrt{D_p \tau_{p0}} = \sqrt{(10)(10^{-7})} = 10.0 \mu\text{m}$$

Then

$$\begin{aligned} J_s &= (1.6 \times 10^{-19})(1.5 \times 10^{10})^2 \times \left[\frac{25}{(35.4 \times 10^{-4})(5 \times 10^{18})} + \frac{10}{(10 \times 10^{-4})(10^{16})} \right] \\ &= 3.6 \times 10^{-11} \text{ A/cm}^2 \end{aligned}$$

Then from Equation (14.10), we can find

$$V_{oc} = V_t \ln \left(1 + \frac{I_L}{I_s} \right) = V_t \ln \left(1 + \frac{J_L}{J_s} \right) = (0.0259) \ln \left(1 + \frac{15 \times 10^{-3}}{3.6 \times 10^{-11}} \right) = 0.514 \text{ V}$$

■ Comment

We may determine the built-in potential barrier of this junction to be $V_{bi} = 0.8556$ V. Taking the ratio of the open-circuit voltage to the built-in potential barrier, we find that $V_{oc}/V_{bi} = 0.60$. The open-circuit voltage will always be less than the built-in potential barrier.

■ EXERCISE PROBLEM

Ex 14.3 Consider a GaAs pn junction solar cell with the following parameters:

$N_a = 10^{17} \text{ cm}^{-3}$, $N_d = 2 \times 10^{16} \text{ cm}^{-3}$, $D_n = 190 \text{ cm}^2/\text{s}$, $D_p = 10 \text{ cm}^2/\text{s}$, $\tau_{n0} = 10^{-7} \text{ s}$, and $\tau_{p0} = 10^{-8} \text{ s}$. Assume a photocurrent density of $J_L = 20 \text{ mA/cm}^2$ is generated in the solar cell. (a) Calculate the open-circuit voltage and (b) determine the ratio of open-circuit voltage to built-in potential barrier.

$$[\Sigma 8L'0 = ^{14}\Lambda / ^{30}\Lambda (q) : \Lambda 1L'0 = ^{30}\Lambda (v) \cdot \text{suV}]$$

The power delivered to the load is

$$P = I \cdot V = I_L \cdot V - I_s \left[\exp \left(\frac{eV}{kT} \right) - 1 \right] \cdot V \quad (14.11)$$

We may find the current and voltage which will deliver the maximum power to the load by setting the derivative equal to zero, or $dP/dV = 0$. Using Equation (14.11), we find

$$\frac{dP}{dV} = 0 = I_L - I_s \left[\exp \left(\frac{eV_m}{kT} \right) - 1 \right] - I_s V_m \left(\frac{e}{kT} \right) \exp \left(\frac{eV_m}{kT} \right) \quad (14.12)$$

where V_m is the voltage that produces the maximum power. We may rewrite Equation (14.12) in the form

$$\left(1 + \frac{V_m}{V_t} \right) \exp \left(\frac{eV_m}{kT} \right) = 1 + \frac{I_L}{I_s} \quad (14.13)$$

The value of V_m may be determined by trial and error. Figure 14.8 shows the maximum power rectangle where I_m is the current when $V = V_m$.

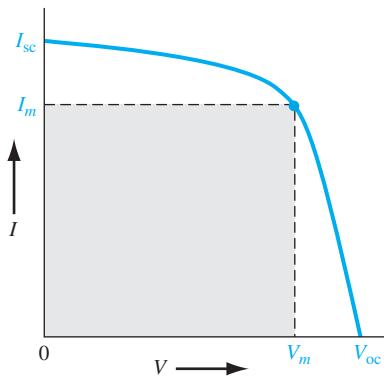


Figure 14.8 | Maximum power rectangle of the solar cell I - V characteristics.

14.2.2 Conversion Efficiency and Solar Concentration

The conversion efficiency of a solar cell is defined as the ratio of output electrical power to incident optical power. For the maximum power output, we can write

$$\eta = \frac{P_m}{P_{in}} \times 100\% = \frac{I_m V_m}{P_{in}} \times 100\% \quad (14.14)$$

The maximum possible current and the maximum possible voltage in the solar cell are I_{sc} and V_{oc} , respectively. The ratio $I_m V_m / I_{sc} V_{oc}$ is called the fill factor and is a measure of the realizable power from a solar cell. Typically, the fill factor is between 0.7 and 0.8.

The conventional pn junction solar cell has a single semiconductor bandgap energy. When the cell is exposed to the solar spectrum, a photon with energy less than E_g will have no effect on the electrical output power of the solar cell. A photon with energy greater than E_g will contribute to the solar cell output power, but the fraction of photon energy that is greater than E_g will eventually only be dissipated as heat. Figure 14.9 shows the solar spectral irradiance (power per unit area per unit wavelength) where air mass zero represents the solar spectrum outside the earth's atmosphere and air mass one is the solar spectrum at the earth's surface at noon. The maximum efficiency of a silicon pn junction solar cell is approximately 28 percent. Nonideal factors, such as series resistance and reflection from the semiconductor surface, will lower the conversion efficiency typically to the range of 10 to 15 percent.

A large optical lens can be used to concentrate sunlight onto a solar cell so that the light intensity can be increased up to several hundred times. The short-circuit current increases linearly with light concentration while the open-circuit voltage increases only slightly with concentration. Figure 14.10 shows the ideal solar cell efficiency at 300 K for two values of solar concentration. We can see that the conversion efficiency increases only slightly with optical concentration. The primary advantage of using concentration techniques is to reduce the overall system cost since an optical lens is less expensive than an equivalent area of solar cells.

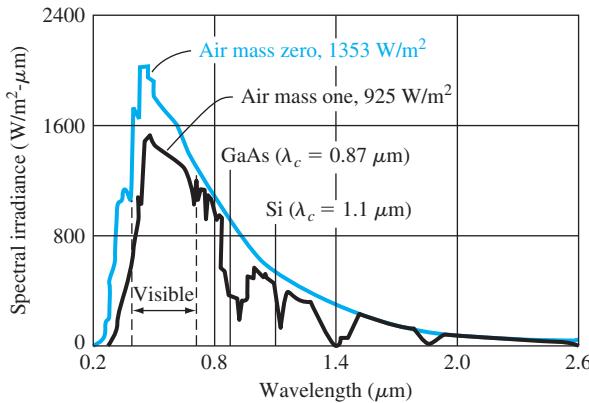


Figure 14.9 | Solar spectral irradiance.

(From Sze [18].)

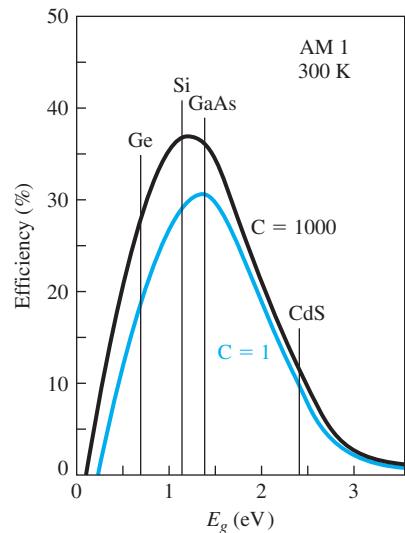


Figure 14.10 | Ideal solar cell efficiency at $T = 300$ K for $C = 1$ sun and for a $C = 1000$ sun concentrations as a function of bandgap energy.
(From Sze [18].)

14.2.3 Nonuniform Absorption Effects

We have seen from the previous section that the photon absorption coefficient in a semiconductor is a very strong function of the incident photon energy or wavelength. Figure 14.4 shows the absorption coefficient as a function of wavelength for several semiconductor materials. As the absorption coefficient increases, more photon energy will be absorbed near the surface than deeper into the semiconductor. In this case, then, we will not have uniform excess carrier generation in a solar cell.

The number of photons absorbed per cm^3 per second as a function of distance x from the surface can be written as

$$\alpha \Phi_0 e^{-\alpha x} \quad (14.15)$$

where Φ_0 is the incident photon flux ($\text{cm}^{-2} \text{s}^{-1}$) on the surface of the semiconductor. We can also take into account the reflection of photons from the surface. Let $R(\lambda)$ be the fraction of photons that are reflected. (For bare silicon, $R \approx 35$ percent.) If we assume that each photon absorbed creates one electron–hole pair, then the generation rate of electron–hole pairs as a function of distance x from the surface is

$$G_L = \alpha(\lambda)\Phi_0(\lambda)[1 - R(\lambda)]e^{-\alpha(\lambda)x} \quad (14.16)$$

where each parameter may be a function of the incident wavelength. Figure 14.11 shows the excess minority carrier concentrations in this pn solar cell for two values of wavelength and for the case when $s = 0$ at the surface.

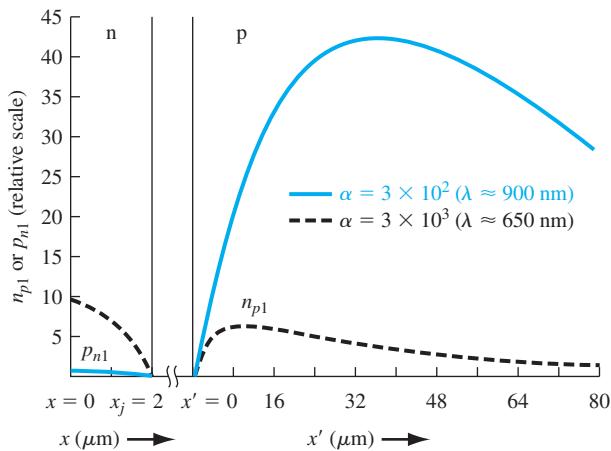


Figure 14.11 | Steady-state, photon-induced normalized minority carrier concentration in the pn junction solar cell for two values of incident photon wavelength ($x_j = 2 \mu\text{m}$, $W = 1 \mu\text{m}$, $L_p = L_n = 40 \mu\text{m}$).

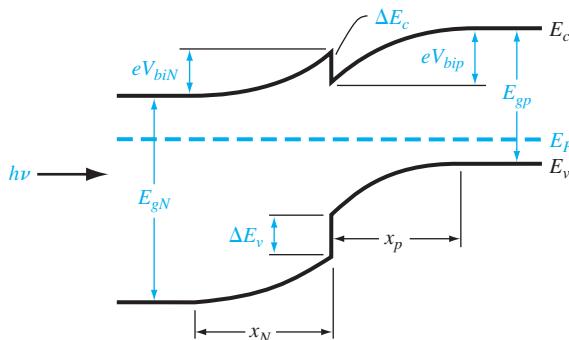


Figure 14.12 | The energy-band diagram of a pN heterojunction in thermal equilibrium.

14.2.4 The Heterojunction Solar Cell

As we have mentioned in previous chapters, a heterojunction is formed between two semiconductors with different bandgap energies. A typical pN heterojunction energy-band diagram in thermal equilibrium is shown in Figure 14.12. Assume that photons are incident on the wide-bandgap material. Photons with energy less than E_{gN} will pass through the wide-bandgap material, which acts as an optical window, and photons with energies greater than E_{gp} will be absorbed in the narrow bandgap material. On the average, excess carriers created in the depletion region and within a diffusion length of the junction will be collected and will contribute to the photocurrent. Photons with an energy greater than E_{gN} will be absorbed in the wide-bandgap

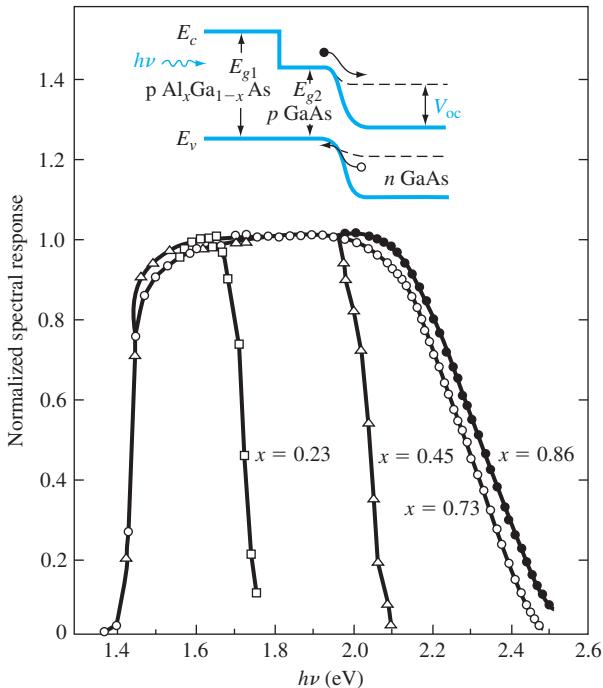


Figure 14.13 | The normalized spectral response of several AlGaAs/GaAs solar cells with different compositions.
(From Sze [17].)

material, and excess carriers generated within one diffusion length of the junction will be collected. If E_{gN} is large enough, then the high-energy photons will be absorbed in the space charge region of the narrow-bandgap material. This heterojunction solar cell should have better characteristics than a homojunction cell, especially at the shorter wavelengths.

A variation of the heterojunction is shown in Figure 14.13. A pn homojunction is formed and then a wide-bandgap material is grown on top. Again, the wide-bandgap material acts as an optical window for photon energies $h\nu < E_{g1}$. Photons with energies $E_{g2} < h\nu < E_{g1}$ will create excess carriers in the homojunction and photons with energies $h\nu > E_{g1}$ will create excess carriers in the window type material. If the absorption coefficient in the narrow bandgap material is high, then essentially all of the excess carriers will be generated within a diffusion length of the junction, so the collection efficiency will be very high. Figure 14.13 also shows the normalized spectral response for various mole fractions x in the $\text{Al}_x\text{Ga}_{1-x}\text{As}$.

14.2.5 Amorphous Silicon Solar Cells

Single-crystal silicon solar cells tend to be expensive and are limited to approximately 6 inches in diameter. A system powered by solar cells requires, in general,

a very large area solar cell array to generate the required power. Amorphous silicon solar cells provide the possibility of fabricating large area and relatively inexpensive solar cell systems.

When silicon is deposited by CVD techniques at temperatures below 600°C, an amorphous film is formed regardless of the type of substrate. In amorphous silicon, there is only very short range order, and no crystalline regions are observed. Hydrogen may be incorporated in the silicon to reduce the number of dangling bonds, creating a material called hydrogenated amorphous silicon.

The density of states versus energy for amorphous silicon is shown in Figure 14.14. Amorphous silicon contains large numbers of electronic energy states within the normal bandgap of single-crystal silicon. However, because of the short-range order, the effective mobility is quite small, typically in the range between 10^{-6} and $10^{-3} \text{ cm}^2/\text{V}\cdot\text{s}$. The mobilities in the states above E_c and below E_v are between 1 and $10 \text{ cm}^2/\text{V}\cdot\text{s}$. Consequently, conduction through the energy states between E_c and E_v is negligible because of the low mobility. Because of the difference in mobility values, E_c and E_v are referred to as the mobility edges and the energy between E_c and E_v is referred to as the mobility gap. The mobility gap can be modified by adding specific types of impurities. Typically, the mobility gap is on the order of 1.7 eV.

Amorphous silicon has a very high optical absorption coefficient, so most sunlight is absorbed within approximately 1 μm of the surface. Consequently, only a

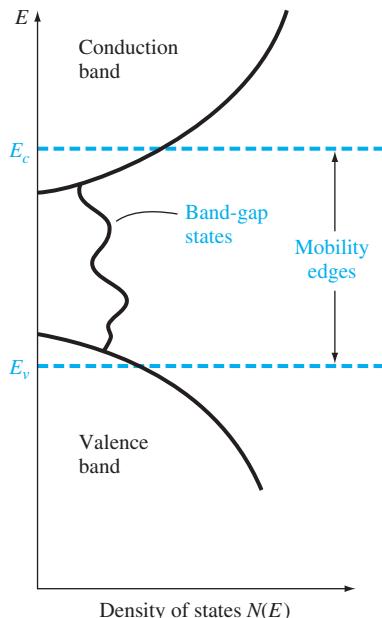


Figure 14.14 | Density of states versus energy of amorphous silicon.
(From Yang [22].)

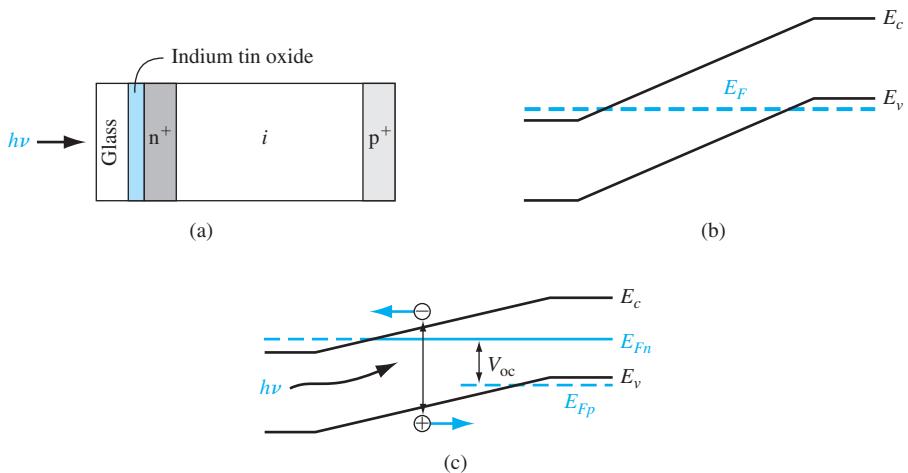


Figure 14.15 | The (a) cross section, (b) energy-band diagram at thermal equilibrium, and (c) energy-band diagram under photon illumination of an amorphous silicon PIN solar cell. (From Yang [22].)

very thin layer of amorphous silicon is required for a solar cell. A typical amorphous silicon solar cell is a PIN device shown in Figure 14.15. The amorphous silicon is deposited on an optically transparent indium tin oxide-coated glass substrate. If aluminum is used as the back contact, it will reflect any transmitted photons back through the PIN device. The n⁺ and p⁺ regions can be quite thin while the intrinsic region may be in the range of 0.5 to 1.0 μm thick. The energy-band diagram for the thermal equilibrium case is shown in the figure. Excess carriers generated in the intrinsic region are separated by the electric field and produce the photocurrent, as we have discussed. Conversion efficiencies are smaller than in single-crystal silicon, but the reduced cost makes this technology attractive. Amorphous silicon solar cells approximately 40 cm wide and many meters long have been fabricated.

TEST YOUR UNDERSTANDING

- TYU 14.2** Consider a silicon pn junction solar cell with the parameters given in Example 14.3. Determine the required photocurrent density to produce an open-circuit voltage of $V_{\text{oc}} = 0.60 \text{ V}$.
(Ans. $J_L = 0.414 \text{ A/cm}^2$)
- TYU 14.3** Consider the silicon pn junction solar cell described in Example 14.3. Let the solar intensity increase by a factor of 10. Calculate the open-circuit voltage.
(Ans. $V_{\text{oc}} = 0.574 \text{ V}$)
- TYU 14.4** The silicon pn junction solar cell described in TYU 14.2 has a cross-sectional area of 1 cm^2 . Determine the maximum power that can be delivered to a load.
(Ans. 0.205 W)

14.3 | PHOTODETECTORS

There are several semiconductor devices that can be used to detect the presence of photons. These devices are known as photodetectors; they convert optical signals into electrical signals. When excess electrons and holes are generated in a semiconductor, there is an increase in the conductivity of the material. This change in conductivity is the basis of the photoconductor, perhaps the simplest type of photodetector. If electrons and holes are generated within the space charge region of a pn junction, then they will be separated by the electric field and a current will be produced. The pn junction is the basis of several photodetector devices including the photodiode and the phototransistor.

14.3.1 Photoconductor

Figure 14.16 shows a bar of semiconductor material with ohmic contacts at each end and a voltage applied between the terminals. The initial thermal-equilibrium conductivity is

$$\sigma_0 = e(\mu_n n_0 + \mu_p p_0) \quad (14.17)$$

If excess carriers are generated in the semiconductor, the conductivity becomes

$$\sigma = e[\mu_n(n_0 + \delta n) + \mu_p(p_0 + \delta p)] \quad (14.18)$$

where δn and δp are the excess electron and hole concentrations, respectively. If we consider an n-type semiconductor, then, from charge neutrality, we can assume that

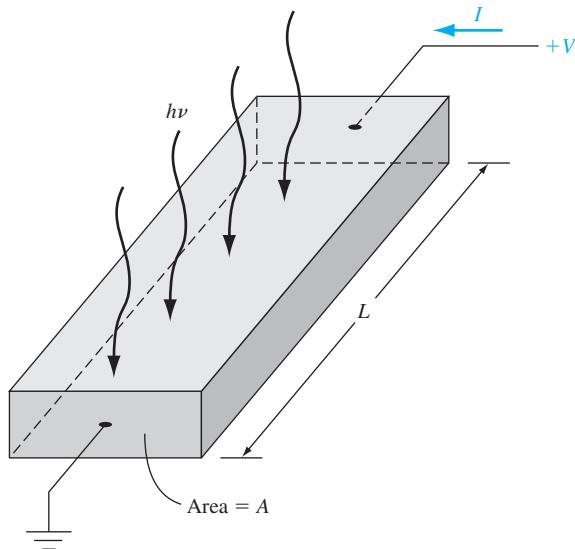


Figure 14.16 | A photoconductor.

$\delta n = \delta p \equiv \delta p$. We will use δp as the concentration of excess carriers. In steady state, the excess carrier concentration is given by $\delta p = G_L \tau_p$, where G_L is the generation rate of excess carriers ($\text{cm}^{-3}\text{-s}^{-1}$) and τ_p is the excess minority carrier lifetime.

The conductivity from Equation (14.18) can be rewritten as

$$\sigma = e(\mu_n n_0 + \mu_p p_0) + e(\delta p)(\mu_n + \mu_p) \quad (14.19)$$

The change in conductivity due to the optical excitation, known as the *photoconductivity*, is then

$$\Delta\sigma = e(\delta p)(\mu_n + \mu_p) \quad (14.20)$$

An electric field is induced in the semiconductor by the applied voltage, which produces a current. The current density can be written as

$$J = (J_0 + J_L) = (\sigma_0 + \Delta\sigma)E \quad (14.21)$$

where J_0 is the current density in the semiconductor prior to optical excitation and J_L is the photocurrent density. The photocurrent density is $J_L = \Delta\sigma \cdot E$. If the excess electrons and holes are generated uniformly throughout the semiconductor, then the photocurrent is given by

$$I_L = J_L \cdot A = \Delta\sigma \cdot AE = eG_L \tau_p (\mu_n + \mu_p) AE \quad (14.22)$$

where A is the cross-sectional area of the device. The photocurrent is directly proportional to the excess carrier generation rate, which in turn is proportional to the incident photon flux.

If excess electrons and holes are not generated uniformly throughout the semiconductor material, then the total photocurrent is found by integrating the photoconductivity over the cross-sectional area.

Since $\mu_n E$ is the electron drift velocity, the electron transit time, that is, the time required for an electron to flow through the photoconductor, is

$$t_n = \frac{L}{\mu_n E} \quad (14.23)$$

The photocurrent, from Equation (14.22), can be rewritten as

$$I_L = eG_L \left(\frac{\tau_p}{t_n} \right) \left(1 + \frac{\mu_p}{\mu_n} \right) AL \quad (14.24)$$

We may define a photoconductor gain, Γ_{ph} , as the ratio of the rate at which charge is collected by the contacts to the rate at which charge is generated within the photoconductor. We can write the gain as

$$\Gamma_{\text{ph}} = \frac{I_L}{eG_L AL} \quad (14.25)$$

which, using Equation (14.24), can be written

$$\Gamma_{\text{ph}} = \frac{\tau_p}{t_n} \left(1 + \frac{\mu_p}{\mu_n} \right) \quad (14.26)$$

Objective: Calculate the photoconductor gain of a silicon photoconductor.

Consider an n-type silicon photoconductor with a length $L = 100 \mu\text{m}$, cross-sectional area $A = 10^{-7} \text{ cm}^2$, and minority carrier lifetime $\tau_p = 10^{-6} \text{ s}$. Let the applied voltage be $V = 10 \text{ volts}$.

EXAMPLE 14.4

■ Solution

The electron transit time is determined as

$$t_n = \frac{L}{\mu_n E} = \frac{L^2}{\mu_n V} = \frac{(100 \times 10^{-4})^2}{(1350)(10)} = 7.41 \times 10^{-9} \text{ s}$$

The photoconductor gain is then

$$\Gamma_{ph} = \frac{\tau_p}{t_n} \left(1 + \frac{\mu_p}{\mu_n}\right) = \frac{10^{-6}}{7.41 \times 10^{-9}} \left(1 + \frac{480}{1350}\right) = 1.83 \times 10^2$$

■ Comment

The fact that a photoconductor—a bar of semiconductor material—has a gain may be surprising.

■ EXERCISE PROBLEM

Ex 14.4 Consider the photoconductor described in Example 14.4. Determine the photocurrent if $G_L = 10^{21} \text{ cm}^{-3} \text{ s}^{-1}$ and $E = 10 \text{ V/cm}$. Also assume that $\mu_n = 1000 \text{ cm}^2/\text{V-s}$ and $\mu_p = 400 \text{ cm}^2/\text{V-s}$.

$$\text{Ans. } I_L = 0.224 \mu\text{A}$$

Let's consider physically what happens to a photon-generated electron, for example. After the excess electron is generated, it drifts very quickly out of the photoconductor at the anode terminal. In order to maintain charge neutrality throughout the photoconductor, another electron immediately enters the photoconductor at the cathode and drifts toward the anode. This process will continue during a time period equal to the mean carrier lifetime. At the end of this period, on the average, the photoelectron will recombine with a hole.

The electron transit time, using the parameters from Example 14.4, is $t_n = 7.41 \times 10^{-9} \text{ s}$. In a simplistic sense, the photoelectron will circulate around the photoconductor circuit 135 times during the 10^{-6} s time duration, which is the mean carrier lifetime. If we take into account the photon-generated hole, the total number of charges collected at the photoconductor contacts for every electron generated is 183.

When the optical signal ends, the photocurrent will decay exponentially with a time constant equal to the minority carrier lifetime. From the photoconductor gain expression, we would like a large minority carrier lifetime, but the switching speed is enhanced by a small minority carrier lifetime. There is obviously a trade-off between gain and speed. In general, the performance of a photodiode, which we will discuss next, is superior to that of a photoconductor.

14.3.2 Photodiode

A photodiode is a pn junction diode operated with an applied reverse-biased voltage. We will initially consider a long diode in which excess carriers are generated

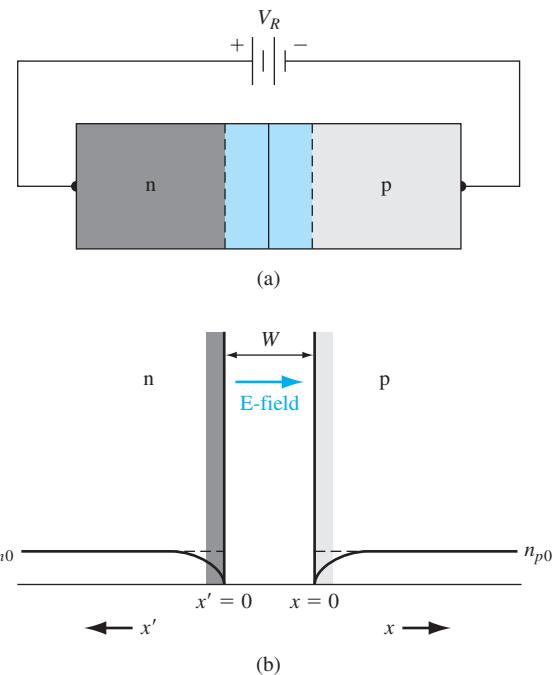


Figure 14.17 | (a) A reverse-biased pn junction. (b) Minority carrier concentration in the reverse-biased pn junction.

uniformly throughout the semiconductor device. Figure 14.17a shows the reverse-biased diode and Figure 14.17b shows the minority carrier distribution in the reverse-biased junction prior to photon illumination.

Let G_L be the generation rate of excess carriers. The excess carriers generated within the space charge region are swept out of the depletion region very quickly by the electric field; the electrons are swept into the n region and the holes into the p region. The photon-generated current density from the space charge region is given by

$$J_{L1} = e \int G_L dx \quad (14.27)$$

where the integral is over the space charge region width. If G_L is constant throughout the space charge volume, then

$$J_{L1} = eG_L W \quad (14.28)$$

where W is the space charge width. We may note that J_{L1} is in the reverse-biased direction through the pn junction. This component of photocurrent responds very quickly to the photon illumination and is known as the prompt photocurrent.

We may note, by comparing Equations (14.28) and (14.25), that the photodiode gain is unity. The speed of the photodiode is limited by the carrier transport through

the space charge region. If we assume that the saturation drift velocity is 10^7 cm/s and the depletion width is $2 \mu\text{m}$, the transit time is $\tau_t = 20$ ps. The ideal modulating frequency has a period of $2\tau_t$, so the frequency is $f = 25$ GHz. This frequency response is substantially higher than that of photoconductors.

Excess carriers are also generated within the neutral n and p regions of the diode. The excess minority carrier electron distribution in the p region is found from the ambipolar transport equation, which is

$$D_n \frac{\partial^2(\delta n_p)}{\partial x^2} + G_L - \frac{\delta n_p}{\tau_{n0}} = \frac{\partial(\delta n_p)}{\partial t} \quad (14.29)$$

We will assume that the E-field is zero in the neutral regions. In steady state, $\partial(\delta n_p)/\partial t = 0$, so that Equation (14.29) can be written as

$$\frac{d^2(\delta n_p)}{dx^2} - \frac{\delta n_p}{L_n^2} = -\frac{G_L}{D_n} \quad (14.30)$$

where $L_n^2 = D_n\tau_{n0}$.

The solution to Equation (14.30) can be found as the sum of the homogeneous and particular solutions. The homogeneous solution is found from the equation

$$\frac{d^2(\delta n_{ph})}{dx^2} - \frac{\delta n_{ph}}{L_n^2} = 0 \quad (14.31)$$

where δn_{ph} is the homogeneous solution and is given by

$$\delta n_{ph} = Ae^{-x/L_n} + Be^{+x/L_n} \quad (x \geq 0) \quad (14.32)$$

One boundary condition is that δn_{ph} must remain finite, which implies that $B \equiv 0$ for the “long” diode.

The particular solution is found from

$$-\frac{\delta n_{pp}}{L_n^2} = -\frac{G_L}{D_n} \quad (14.33)$$

which yields

$$\delta n_{pp} = \frac{G_L L_n^2}{D_n} = \frac{G_L (D_n \tau_{n0})}{D_n} = G_L \tau_{n0} \quad (14.34)$$

The total steady-state solution for the excess minority carrier electron concentration in the p region is then

$$\delta n_p = Ae^{-x/L_n} + G_L \tau_{n0} \quad (14.35)$$

The total electron concentration is zero at $x = 0$ for the reverse-biased junction. The excess electron concentration $x = 0$ is then

$$\delta n_p(x = 0) = -n_{p0} \quad (14.36)$$

Using the boundary condition from Equation (14.36), the electron concentration given by Equation (14.35) becomes

$$\delta n_p = G_L \tau_{n0} - (G_L \tau_{n0} + n_{p0})e^{-x/L_n} \quad (14.37)$$

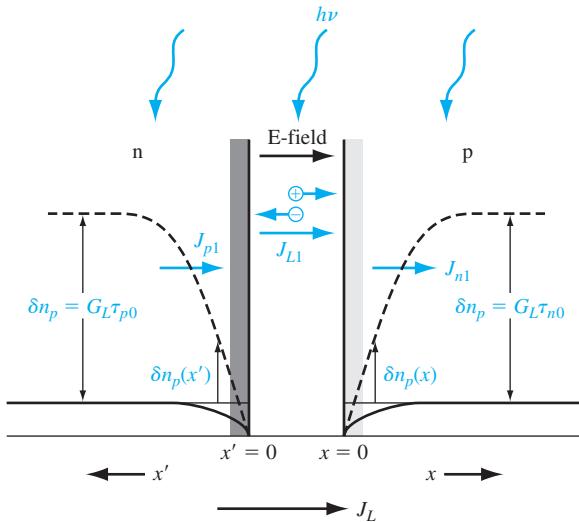


Figure 14.18 | Steady-state, photoinduced minority carrier concentrations and photocurrents in a “long” reverse-biased pn junction.

We can find the excess minority carrier hole concentration in the n region using the same type of analysis. Using the x' notation shown in Figure 14.17, we can write

$$\delta p_n = G_L \tau_{p0} - (G_L \tau_{p0} + p_{n0}) e^{-x'/L_p} \quad (14.38)$$

Equations (14.37) and (14.38) are plotted in Figure 14.18. We may note that the steady-state values far from the space charge region are the same as were given previously.

The gradient in the minority carrier concentrations will produce diffusion currents in the pn junction. The diffusion current density at $x = 0$ due to minority carrier electrons is

$$\begin{aligned} J_{n1} &= e D_n \frac{d(\delta n_p)}{dx} \Big|_{x=0} = e D_n \frac{d}{dx} [G_L \tau_{n0} - (G_L \tau_{n0} + n_{p0}) e^{-x/L_n}] \Big|_{x=0} \\ &= \frac{e D_n}{L_n} (G_L \tau_{n0} + n_{p0}) \end{aligned} \quad (14.39)$$

Equation (14.39) can be written as

$$J_{n1} = e G_L L_n + \frac{e D_n n_{p0}}{L_n} \quad (14.40)$$

The first term in Equation (14.40) is the steady-state photocurrent density while the second term is the ideal reverse saturation current density due to the minority carrier electrons.

The diffusion current density (in the x direction) at $x' = 0$ due to the minority carrier holes is

$$J_{p1} = e G_L L_p + \frac{e D_p p_{n0}}{L_p} \quad (14.41)$$

Similarly, the first term is the steady-state photocurrent density and the second term is the ideal reverse saturation current density.

The total steady-state diode photocurrent density for the long diode is now

$$J_L = eG_L W + eG_L L_n + eG_L L_p = e(W + L_n + L_p)G_L \quad (14.42)$$

Again note that the photocurrent is in the reverse-biased direction through the diode. The photocurrent given by Equation (14.42) is the result of assuming uniform generation of excess carriers throughout the structure, a long diode, and steady state.

The time response of the diffusion components of the photocurrent is relatively slow, since these currents are the results of the diffusion of minority carriers toward the depletion region. The diffusion components of photocurrent are referred to as the delayed photocurrent.

Objective: Calculate the steady-state photocurrent density in a reverse-biased, long pn diode.

Consider a silicon pn diode at $T = 300$ K with the following parameters:

$$\begin{aligned} N_a &= 10^{16} \text{ cm}^{-3} & N_d &= 10^{16} \text{ cm}^{-3} \\ D_n &= 25 \text{ cm}^2/\text{s} & D_p &= 10 \text{ cm}^2/\text{s} \\ \tau_{n0} &= 5 \times 10^{-7} \text{ s} & \tau_{p0} &= 10^{-7} \text{ s} \end{aligned}$$

Assume that a reverse-biased voltage of $V_R = 5$ volts is applied and let $G_L = 10^{21} \text{ cm}^{-3}\text{-s}^{-1}$.

EXAMPLE 14.5

Solution

We may calculate various parameters as follows:

$$L_n = \sqrt{D_n \tau_{n0}} = \sqrt{(25)(5 \times 10^{-7})} = 35.4 \mu\text{m}$$

$$L_p = \sqrt{D_p \tau_{p0}} = \sqrt{(10)(10^{-7})} = 10.0 \mu\text{m}$$

$$V_{bi} = V_t \ln \left(\frac{N_a N_d}{n_i^2} \right) = (0.0259) \ln \left[\frac{(10^{16})(10^{16})}{(1.5 \times 10^{10})^2} \right] = 0.695 \text{ V}$$

$$\begin{aligned} W &= \left\{ \frac{2\epsilon_s}{e} \left(\frac{N_a + N_d}{N_a N_d} \right) (V_{bi} + V_R) \right\}^{1/2} \\ &= \left\{ \frac{2(11.7)(8.85 \times 10^{-14})}{1.6 \times 10^{-19}} \cdot \frac{(2 \times 10^{16})}{(10^{16})(10^{16})} \cdot (0.695 + 5) \right\}^{1/2} = 1.21 \mu\text{m} \end{aligned}$$

Finally, the steady-state photocurrent density is

$$\begin{aligned} J_L &= e(W + L_n + L_p)G_L \\ &= (1.6 \times 10^{-19})(1.21 + 35.4 + 10.0) \times 10^{-4}(10^{21}) = 0.75 \text{ A/cm}^2 \end{aligned}$$

Comment

Again, keep in mind that this photocurrent is in the reverse-biased direction through the diode and is many orders of magnitude larger than the reverse-biased saturation current density in the pn junction diode.

■ EXERCISE PROBLEM

Ex 14.5 The doping concentrations of the photodiode described in Example 14.5 are changed to $N_a = N_d = 10^{15} \text{ cm}^{-3}$. (a) Determine the steady-state photocurrent density. (b) Calculate the ratio of prompt photocurrent to steady-state photocurrent.

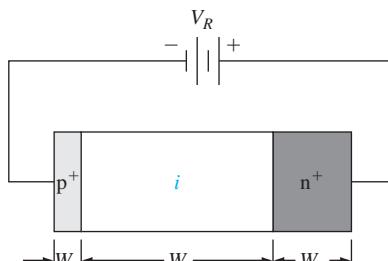
$$[E770'0 = I_p/I_s(q); A/cm^2; I_p(v) = 787.0 \text{ A/s}]$$

In this example calculation, $L_n \gg W$ and $L_p \gg W$. In many pn junction structures, the assumption of a long diode will not be valid, so the photocurrent expression will have to be modified. In addition, the photon energy absorption may not be uniform throughout the pn structure. The effect of nonuniform absorption will be considered in the next section.

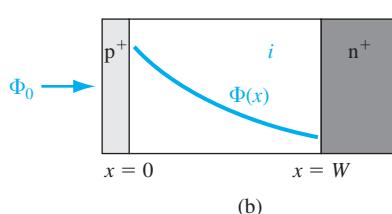
14.3.3 PIN Photodiode

In many photodetector applications, the speed of response is important; therefore, the prompt photocurrent generated in the space charge region is the only photocurrent of interest. To increase the photodetector sensitivity, the depletion region width should be made as large as possible. This can be achieved in a PIN photodiode.

The PIN diode consists of a p region and an n region separated by an intrinsic region. A sketch of a PIN diode is shown in Figure 14.19a. The intrinsic region width W is much larger than the space charge width of a normal pn junction. If a reverse bias is applied to the PIN diode, the space charge region extends completely through the intrinsic region.



(a)



(b)

Figure 14.19 | (a) A reverse-biased PIN photodiode. (b) Geometry showing nonuniform photon absorption.

Assume that a photon flux Φ_0 is incident on the p⁺ region. If we assume that the p⁺ region width W_p is very thin, then the photon flux, as a function of distance, in the intrinsic region is $\Phi(x) = \Phi_0 e^{-\alpha x}$, where α is the photon absorption coefficient. This nonlinear photon absorption is shown in Figure 14.19b. The photocurrent density generated in the intrinsic region can be found as

$$J_L = e \int_0^W G_L dx = e \int_0^W \Phi_0 \alpha e^{-\alpha x} dx = e \Phi_0 (1 - e^{-\alpha W}) \quad (14.43)$$

This equation assumes that there is no electron–hole recombination within the space charge region and also that each photon absorbed creates one electron–hole pair.

Objective: Calculate the photocurrent density in a PIN photodiode.

EXAMPLE 14.6

Consider a silicon PIN diode with an intrinsic region width of $W = 20 \mu\text{m}$. Assume that the photon flux is $10^{17} \text{ cm}^{-2}\text{-s}^{-1}$ and the absorption coefficient is $\alpha = 10^3 \text{ cm}^{-1}$.

■ Solution

The generation rate of electron–hole pairs at the front edge of the intrinsic region is

$$G_{L1} = \alpha \Phi_0 = (10^3)(10^{17}) = 10^{20} \text{ cm}^{-3}\text{-s}^{-1}$$

and the generation rate at the back edge of the intrinsic region is

$$\begin{aligned} G_{L2} &= \alpha \Phi_0 e^{-\alpha W} = (10^3)(10^{17}) \exp [-(10^3)(20 \times 10^{-4})] \\ &= 0.135 \times 10^{20} \text{ cm}^{-3}\text{-s}^{-1} \end{aligned}$$

The generation rate is obviously not uniform throughout the intrinsic region. The photocurrent density is then

$$\begin{aligned} J_L &= e \Phi_0 (1 - e^{-\alpha W}) \\ &= (1.6 \times 10^{-19})(10^{17}) \{1 - \exp [-(10^3)(20 \times 10^{-4})]\} \\ &= 13.8 \text{ mA/cm}^2 \end{aligned}$$

■ Comment

The prompt photocurrent density of a PIN photodiode will be larger than that of a regular photodiode since the space charge region is larger in a PIN photodiode.

■ EXERCISE PROBLEM

Ex 14.6 Repeat Example 14.6 for photon absorption coefficients of (a) $\alpha = 10^2 \text{ cm}^{-1}$ and (b) $\alpha = 10^4 \text{ cm}^{-1}$.

[Ans. (a) $J_L = 2.96 \text{ mA/cm}^2$; (b) $J_L = 16.0 \text{ mA/cm}^2$]

In most situations, we will not have a long diode; thus, the steady-state photocurrent described by Equation (14.42) will not apply for most photodiodes.

14.3.4 Avalanche Photodiode

The avalanche photodiode is similar to the pn or PIN photodiode except that the bias applied to the avalanche photodiode is sufficiently large to cause impact ionization.

Electron–hole pairs are generated in the space charge region by photon absorption, as we have discussed previously. The photon-generated electrons and holes now generate additional electron–hole pairs through impact ionization. The avalanche photodiode now has a current gain introduced by the avalanche multiplication factor.

The electron–hole pairs generated by photon absorption and by impact ionization are swept out of the space charge region very quickly. If the saturation velocity is 10^7 cm/s in a depletion region that is $10\ \mu\text{m}$ wide, then the transit time is

$$\tau_t = \frac{10^7}{10 \times 10^{-4}} = 100\ \text{ps}$$

The period of a modulation signal would be $2\tau_t$, so that the frequency would be

$$f = \frac{1}{2\tau_t} = \frac{1}{200 \times 10^{-12}} = 5\ \text{GHz}$$

If the avalanche photodiode current gain is 20, then the gain-bandwidth product is 100 GHz. The avalanche photodiode could respond to light waves modulated at microwave frequencies.

14.3.5 Phototransistor

A bipolar transistor can also be used as a photodetector. The phototransistor can have high gain through the transistor action. An npn bipolar phototransistor is shown in Figure 14.20a. This device has a large base–collector junction area and is usually operated with the base open circuited. Figure 14.20b shows the block diagram of the phototransistor. Electrons and holes generated in the reverse-biased B–C junction are swept out of the space charge region, producing a photocurrent I_L . Holes are swept

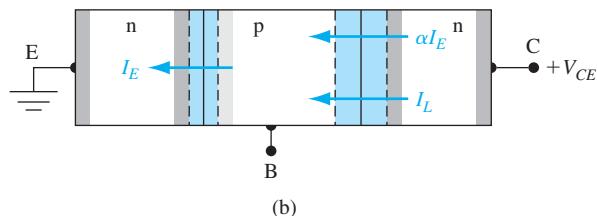
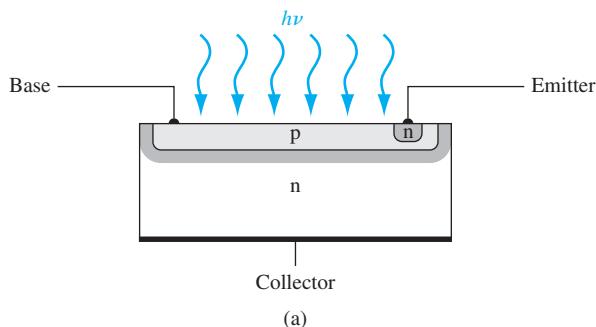


Figure 14.20 | (a) A bipolar phototransistor. (b) Block diagram of the open-base phototransistor.

into the p-type base, making the base positive with respect to the emitter. Since the B-E becomes forward-biased, electrons will be injected from the emitter back into the base, leading to the normal transistor action.

From Figure 14.20b, we see that

$$I_E = \alpha I_E + I_L \quad (14.44)$$

where I_L is the photon-generated current and α is the common base current gain. Since the base is an open circuit, we have $I_C = I_E$, so Equation (14.44) can be written as

$$I_C = \alpha I_C + I_L \quad (14.45)$$

Solving for I_C , we find

$$I_C = \frac{I_L}{1 - \alpha} \quad (14.46)$$

Relating α to β , the dc common emitter current gain, Equation (14.46) becomes

$$I_C = (1 + \beta)I_L \quad (14.47)$$

Equation (14.47) shows that the basic B-C photocurrent is multiplied by the factor $(1 + \beta)$. The phototransistor, then, amplifies the basic photocurrent.

With the relatively large B-C junction area, the frequency response of the phototransistor is limited by the B-C junction capacitance. Since the base is essentially the input to the device, the large B-C capacitance is multiplied by the Miller effect, so the frequency response of the phototransistor is further reduced. The phototransistor, however, is a lower-noise device than the avalanche photodiode.

Phototransistors can also be fabricated in heterostructures. The injection efficiency is increased as a result of the bandgap differences, as we discussed in Chapter 12. With the bandgap difference, the lightly doped base restriction no longer applies. A fairly heavily doped, narrow-base device can be fabricated with a high blocking voltage and a high gain.

TEST YOUR UNDERSTANDING

- TYU 14.5** Consider a long silicon pn junction photodiode with the parameters given in Example 14.5. The cross-sectional area is $A = 10^{-3} \text{ cm}^2$. Assume the photodiode is reverse biased by a 5-volt battery in series with a $5 \text{ k}\Omega$ load resistor. An optical signal at a wavelength of $\lambda = 1 \mu\text{m}$ is incident on the photodiode producing a uniform generation rate of excess carriers throughout the entire device. Determine the incident intensity such that the voltage across the load resistor is 0.5 V.
 (Ans. $I^* = 0.269 \text{ W/cm}^2$)

14.4 | PHOTOLUMINESCENCE AND ELECTROLUMINESCENCE

In the first section of this chapter, we have discussed the creation of excess electron-hole pairs by photon absorption. Eventually, excess electrons and holes recombine, and in direct bandgap materials the recombination process may result in the emission of a photon. The general property of light emission is referred to as luminescence.

When excess electrons and holes are created by photon absorption, photon emission from the recombination process is called photoluminescence.

Electroluminescence is the process of generating photon emission when the excitation of excess carriers is a result of an electric current caused by an applied electric field. We are mainly concerned here with injection electroluminescence, the result of injecting carriers across a pn junction. The light emitting diode and the pn junction laser diode are examples of this phenomenon. In these devices, electric energy, in the form of a current, is converted directly into photon energy.

14.4.1 Basic Transitions

Once electron–hole pairs are formed, there are several possible processes by which the electrons and holes can recombine. Some recombination processes may result in photon emission from direct bandgap materials, whereas other recombination processes in the same material may not.

Figure 14.21a shows the basic interband transitions. Curve (i) corresponds to an intrinsic emission very close to the bandgap energy of the material. Curves (ii) and (iii) correspond to energetic electrons or holes. If either of these recombinations result in the emission of a photon, the energy of the emitted photon will be slightly larger than the bandgap energy. There will then be an emission spectrum and a bandwidth associated with the emission.

The possible recombination processes involving impurity or defect states are shown in Figure 14.21b. Curve (i) is the conduction band to acceptor transition, curve (ii) is the donor to valence-band transition, curve (iii) is the donor to acceptor transition, and curve (iv) is the recombination due to a deep trap. Curve (iv) is a nonradiative process corresponding to the Shockley–Read–Hall recombination process discussed in Chapter 6. The other recombination processes may or may not result in the emission of a photon.

Figure 14.21c shows the Auger recombination process, which can become important in direct bandgap materials with high doping concentrations. The Auger recombination process is a nonradiative process. The Auger recombination, in one case, shown in curve (i), is a recombination between an electron and hole, accompanied by the transfer of energy to another free hole. Similarly, in the second case, the recombination between an electron and hole can result in the transfer of energy to a free electron as shown in curve (ii). The third particle involved in this process will eventually lose its energy to the lattice in the form of heat. The process involving two holes and an electron would occur predominantly in heavily doped p-type materials, and the process involving two electrons and a hole would occur primarily in a heavily doped n-type material.

The recombination processes shown in Figure 14.21a indicate that the emission of a photon is not necessarily at a single, discrete energy, but can occur over a range of energies. The spontaneous emission rate generally has the form

$$I(\nu) \propto \nu^2 (h\nu - E_g)^{1/2} \exp\left[\frac{-(h\nu - E_g)}{kT}\right] \quad (14.48)$$

where E_g is the bandgap energy. Figure 14.22 shows the emission spectra from gallium arsenide. The peak photon energy decreases with temperature because the

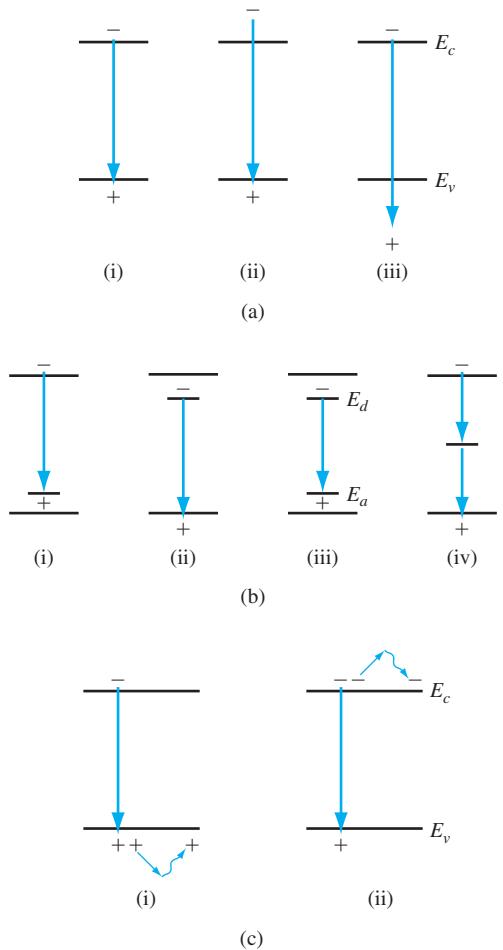


Figure 14.21 | Basic transitions in a semiconductor.

bandgap energy decreases with temperature. We will show that the bandwidth of the emission spectra can be greatly reduced in a laser diode by using an optical resonator.

14.4.2 Luminescent Efficiency

We have shown that not all recombination processes are radiative. An efficient luminescent material is one in which radiative transitions predominate. The quantum efficiency is defined as the ratio of the radiative recombination rate to the total recombination rate for all processes. We can write

$$\eta_q = \frac{R_r}{R} \quad (14.49)$$

where η_q is the quantum efficiency, R_r is the radiative recombination rate, and R is the total recombination rate of the excess carriers. Since the recombination rate is

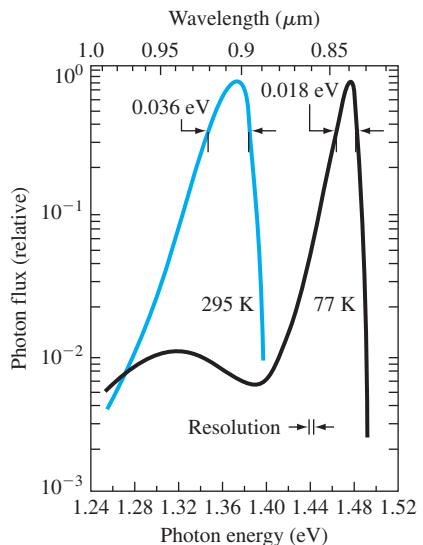


Figure 14.22 | GaAs diode emission spectra at $T = 300$ K and $T = 77$ K.
(From Sze and Ng [17].)

inversely proportional to lifetime, we can write the quantum efficiency in terms of lifetimes as

$$\eta_q = \frac{\tau_{nr}}{\tau_{nr} + \tau_r} \quad (14.50)$$

where τ_{nr} is the nonradiative lifetime and τ_r is the radiative lifetime. For a high luminescent efficiency, the nonradiative lifetimes must be large; thus, the probability of a nonradiative recombination is small compared to the radiative recombination.

The interband recombination rate of electrons and holes will be directly proportional to the number of electrons available and directly proportional to the number of available empty states (holes). We can write

$$R_r = Bnp \quad (14.51)$$

where R_r is the band-to-band radiative recombination rate and B is the constant of proportionality. The values of B for direct-bandgap materials are on the order of 10^6 larger than for indirect bandgap materials. The probability of a direct band-to-band radiative recombination transition in an indirect bandgap material is very unlikely.

One problem encountered with the emission of photons from a direct bandgap material is the reabsorption of the emitted photons. In general, the emitted photons will have energies $h\nu > E_g$, which means that the absorption coefficient is not zero for this energy. In order to generate a light output from a light emitting device, the process must take place near the surface. One possible solution to the reabsorption problem is to use heterojunction devices. These are discussed in later sections.

14.4.3 Materials

An important direct bandgap semiconductor material for optical devices is gallium arsenide. Another compound material that is of great interest is $\text{Al}_x\text{Ga}_{1-x}\text{As}$. This material is a compound semiconductor in which the ratio of aluminum atoms to gallium atoms can be varied to achieve specific characteristics. Figure 14.23 shows the bandgap energy as a function of the mole fraction between aluminum and gallium. We can note from the figure that for $0 < x < 0.45$, the alloy material is a direct bandgap material. For $x > 0.45$, the material becomes an indirect bandgap material, not suitable for optical devices. For $0 < x < 0.35$, the bandgap energy can be expressed as

$$E_g = 1.424 + 1.247x \text{ eV} \quad (14.52)$$

Another compound semiconductor used for optical devices is the $\text{GaAs}_{1-x}\text{P}_x$ system. Figure 14.24a shows the bandgap energy as a function of the mole fraction x . For $0 \leq x \leq 0.45$, this material is also a direct bandgap material, and for $x > 0.45$, the bandgap becomes indirect. Figure 14.24b is the E versus k diagram, showing how the bandgap changes from direct to indirect as the mole fraction changes.

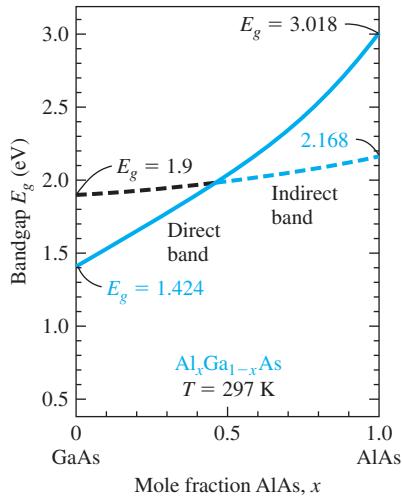


Figure 14.23 | Bandgap energy of $\text{Al}_x\text{Ga}_{1-x}\text{As}$ as a function of the mole fraction x .
(From Sze [18].)

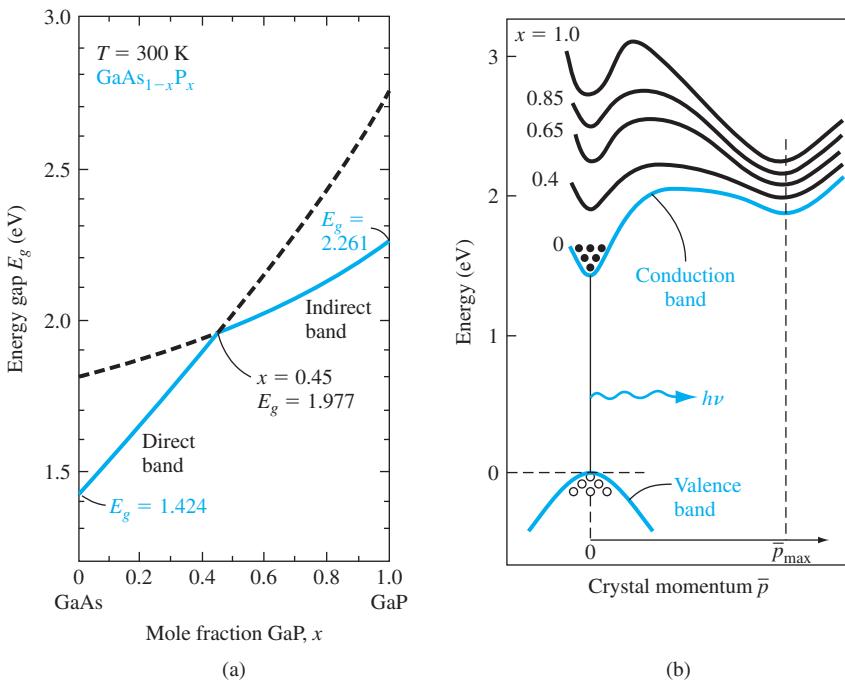


Figure 14.24 | (a) Bandgap energy of $\text{GaAs}_{1-x}\text{P}_x$ as a function of mole fraction x .
(b) E versus k diagram of $\text{GaAs}_{1-x}\text{P}_x$ for various values of x .
(From Sze [18].)

EXAMPLE 14.7

Objective: Determine the output wavelength of a $\text{GaAs}_{1-x}\text{P}_x$ material for two different mole fractions.

Consider first GaAs and then $\text{GaAs}_{1-x}\text{P}_x$.

■ Solution

GaAs has a bandgap energy of $E_g = 1.42 \text{ eV}$. This material would produce a photon output at a wavelength of

$$\lambda = \frac{1.24}{E} = \frac{1.24}{1.42} = 0.873 \mu\text{m}$$

This wavelength is in the infrared range and not in the visible range. If we desire a visible output with a wavelength of $\lambda = 0.653 \mu\text{m}$, for example, the bandgap energy would have to be

$$E = \frac{1.24}{\lambda} = \frac{1.24}{0.653} = 1.90 \text{ eV}$$

This bandgap energy would correspond to a mole fraction of approximately $x = 0.4$.

■ Comment

By changing the mole fraction in the $\text{GaAs}_{1-x}\text{P}_x$ system, the output can change from the infrared to the red spectrum.

■ EXERCISE PROBLEM

Ex 14.7 Determine the output wavelength of a $\text{GaAs}_{1-x}\text{P}_x$ material for mole fractions of

- (a) $x = 0.15$ and (b) $x = 0.30$.

[unravel 20 = γ(q) : unravel 220 = γ(v) · sun]

14.5 | LIGHT EMITTING DIODES

Photodetectors and solar cells convert optical energy into electrical energy—the photons generate excess electrons and holes, which produce an electric current. We might also apply a voltage across a pn junction resulting in a diode current, which in turn can produce photons and a light output. This inverse mechanism is called injection electroluminescence. This device is known as a **Light Emitting Diode (LED)**. The spectral output of an LED may have a relatively wide wavelength bandwidth of between 30 and 40 nm. However, this emission spectrum is narrow enough so that a particular color is observed, provided the output is in the visible range.

14.5.1 Generation of Light

As we have discussed previously, photons may be emitted if an electron and hole recombine by a direct band-to-band recombination process in a direct bandgap material. The emission wavelength, from Equation (14.1), is

$$\lambda = \frac{hc}{E_g} = \frac{1.24}{E_g} \mu\text{m} \quad (14.53)$$

where E_g is the bandgap energy measured in electron-volts.

When a voltage is applied across a pn junction, electrons and holes are injected across the space charge region where they become excess minority carriers. These excess minority carriers diffuse into the neutral semiconductor regions where they recombine with majority carriers. If this recombination process is a direct band-to-band process, photons are emitted. The diode diffusion current is directly proportional to the recombination rate, so the output photon intensity will also be proportional to the ideal diode diffusion current. In gallium arsenide, electroluminescence originates primarily on the p side of the junction because the efficiency for electron injection is higher than that for hole injection.

14.5.2 Internal Quantum Efficiency

The *internal quantum efficiency* of an LED is the fraction of diode current that produces luminescence. The internal quantum efficiency is a function of the injection efficiency and a function of the percentage of radiative recombination events compared with the total number of recombination events.

The three current components in a forward-biased diode are the minority carrier electron diffusion current, the minority carrier hole diffusion current, and the space charge recombination current. These current densities can be written, respectively, as

$$J_n = \frac{eD_n n_{p0}}{L_n} \left[\exp\left(\frac{eV}{kT}\right) - 1 \right] \quad (14.54a)$$

$$J_p = \frac{eD_p p_{n0}}{L_p} \left[\exp\left(\frac{eV}{kT}\right) - 1 \right] \quad (14.54b)$$

and

$$J_R = \frac{en_i W}{2\tau_0} \left[\exp\left(\frac{eV}{2kT}\right) - 1 \right] \quad (14.54c)$$

The recombination of electrons and holes within the space charge region is, in general, through traps near midgap and is a nonradiative process. Since luminescence is due primarily to the recombination of minority carrier electrons in GaAs, we can define an injection efficiency as the fraction of electron current to total current. Then

$$\gamma = \frac{J_n}{J_n + J_p + J_R} \quad (14.55)$$

where γ is the injection efficiency. We can make γ approach unity by using an n⁺p diode so that J_p is a small fraction of the diode current and by forward biasing the diode sufficiently so that J_R is a small fraction of the total diode current.

Once the electrons are injected into the p region, not all electrons will recombine radiatively. We can define the radiative and nonradiative recombination rates as

$$R_r = \frac{\delta n}{\tau_r} \quad (14.56a)$$

and

$$R_{nr} = \frac{\delta n}{\tau_{nr}} \quad (14.56b)$$

where τ_r and τ_{nr} are the radiative and nonradiative recombination lifetimes, respectively, and δn is the excess carrier concentration. The total recombination rate is

$$R = R_r + R_{nr} = \frac{\delta n}{\tau} = \frac{\delta n}{\tau_r} + \frac{\delta n}{\tau_{nr}} \quad (14.57)$$

where τ is the net excess carrier lifetime.

The radiative efficiency is defined as the fraction of recombinations that are radiative. We can write

$$\eta = \frac{R_r}{R_r + R_{nr}} = \frac{\frac{1}{\tau_r}}{\frac{1}{\tau_r} + \frac{1}{\tau_{nr}}} = \frac{\tau}{\tau_r + \tau_{nr}} \quad (14.58)$$

where η is the radiative efficiency. The nonradiative recombination rate is proportional to N_t , which is the density of nonradiative trapping sites within the forbidden bandgap. Obviously, the radiative efficiency increases as N_t is reduced.

The internal quantum efficiency is now written as

$$\eta_i = \gamma \eta \quad (14.59)$$

The radiative recombination rate is proportional to the p-type doping. As the p-type doping increases, the radiative recombination rate increases. However, the injection efficiency decreases as the p-type doping increases; therefore, there is an optimum doping that maximizes the internal quantum efficiency.

14.5.3 External Quantum Efficiency

One very important parameter of the LED is the *external quantum efficiency*: the fraction of generated photons that are actually emitted from the semiconductor. The external quantum efficiency is normally a much smaller number than the internal quantum efficiency. Once a photon has been produced in the semiconductor, there are three loss mechanisms the photon may encounter: photon absorption within the semiconductor, Fresnel loss, and critical angle loss.

Figure 14.25 shows a pn junction LED. Photons can be emitted in any direction. Since the emitted photon energy must be $h\nu \geq E_g$, these emitted photons can be reabsorbed within the semiconductor material. The majority of photons will actually be emitted away from the surface and reabsorbed in the semiconductor.

Photons must be emitted from the semiconductor into air; thus, the photons must be transmitted across a dielectric interface. Figure 14.26 shows the incident, reflected, and transmitted waves. The parameter \bar{n}_2 is the index of refraction for the semiconductor and \bar{n}_1 is the index of refraction for air. The reflection coefficient is

$$\Gamma = \left(\frac{\bar{n}_2 - \bar{n}_1}{\bar{n}_2 + \bar{n}_1} \right)^2 \quad (14.60)$$

This effect is called Fresnel loss. The reflection coefficient Γ is the fraction of incident photons that are reflected back into the semiconductor.

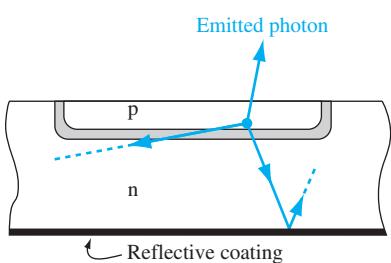


Figure 14.25 | Schematic of photon emission at the pn junction of an LED.

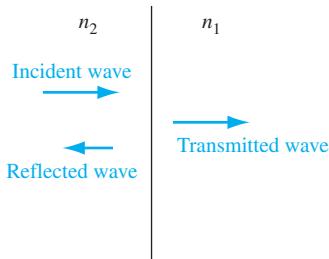


Figure 14.26 | Schematic of incident, reflected, and transmitted photons at a dielectric interface.

Objective: Calculate the reflection coefficient at a semiconductor–air interface.

Consider the interface between a GaAs semiconductor and air.

EXAMPLE 14.8

■ Solution

The index of refraction for GaAs is $\bar{n}_2 = 3.8$ at a wavelength of $\lambda = 0.70 \mu\text{m}$ and the index of refraction for air is $\bar{n}_1 = 1.0$. The reflection coefficient is

$$\Gamma = \left(\frac{\bar{n}_2 - \bar{n}_1}{\bar{n}_2 + \bar{n}_1} \right)^2 = \left(\frac{3.8 - 1.0}{3.8 + 1.0} \right)^2 = 0.34$$

■ Comment

A reflection coefficient of $\Gamma = 0.34$ means that 34 percent of the photons incident from the gallium arsenide onto the semiconductor–air interface are reflected back into the semiconductor.

■ EXERCISE PROBLEM

- Ex 14.8** At a wavelength of $\lambda = 0.70 \mu\text{m}$, the index of refraction for GaAs is $\bar{n}_2 = 3.8$ and that for GaP is $\bar{n}_2 = 3.2$. Consider a $\text{GaAs}_{1-x}\text{P}_x$ material with a mole fraction $x = 0.40$. Assuming the index of refraction is a linear function of the mole fraction, determine the reflection coefficient, Γ , at the $\text{GaAs}_{0.6}\text{P}_{0.4}$ –air interface.

(Ans. $\Gamma = 0.315$)

Photons incident on the semiconductor–air interface at an angle are refracted as shown in Figure 14.27. If the photons are incident on the interface at an angle greater than the critical angle θ_c , the photons experience total internal reflection. The critical angle is determined from Snell's law and is given by

$$\theta_c = \sin^{-1} \left(\frac{\bar{n}_1}{\bar{n}_2} \right) \quad (14.61)$$

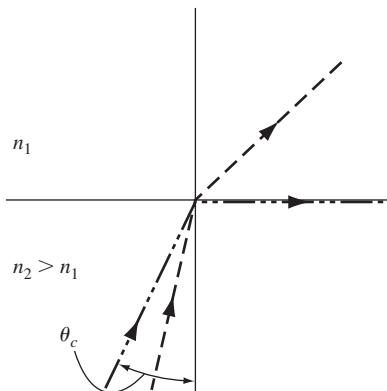


Figure 14.27 | Schematic showing refraction and total internal reflection at the critical angle at a dielectric interface.

EXAMPLE 14.9

Objective: Calculate the critical angle at a semiconductor–air interface.

Consider the interface between GaAs and air.

■ Solution

For GaAs, $\bar{n}_2 = 3.8$ at a wavelength of $\lambda = 0.70 \mu\text{m}$ and for air, $\bar{n}_1 = 1.0$. The critical angle is

$$\theta_c = \sin^{-1}\left(\frac{\bar{n}_1}{\bar{n}_2}\right) = \sin^{-1}\left(\frac{1.0}{3.8}\right) = 15.3^\circ$$

■ Comment

Any photon that is incident at an angle greater than 15.3° will be reflected back into the semiconductor.

■ EXERCISE PROBLEM

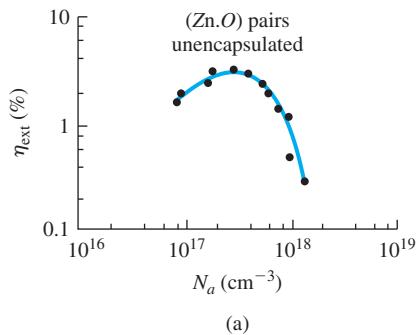
Ex 14.9 Repeat Example 14.9 for GaAs_{0.6}P_{0.4}. See Exercise Problem Ex 14.8 for a discussion of the dielectric constant.

$$(\text{Ans. } 16.3^\circ)$$

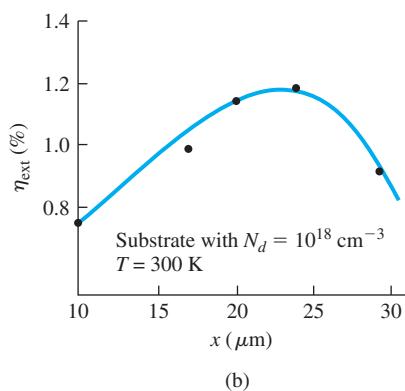
Figure 14.28a shows the external quantum efficiency plotted as a function of the p-type doping concentration and Figure 14.28b is a plot of the external efficiency as a function of junction depth below the surface. Both figures show that the external quantum efficiency is in the range of 1 to 3 percent.

14.5.4 LED Devices

The wavelength of the output signal of an LED is determined by the bandgap energy of the semiconductor. Gallium arsenide, a direct bandgap material, has a bandgap energy of $E_g = 1.42 \text{ eV}$, which yields a wavelength of $\lambda = 0.873 \mu\text{m}$. Comparing this wavelength to the visible spectrum, which is shown in Figure 14.5, the output



(a)



(b)

Figure 14.28 | (a) External quantum efficiency of a GaP LED versus acceptor doping. (b) External quantum efficiency of a GaAs LED versus junction depth. (From Yang [22].)

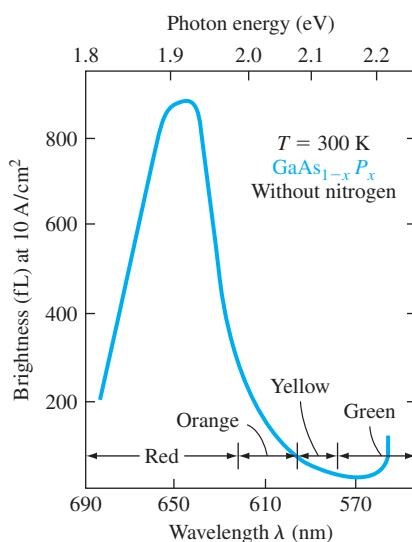


Figure 14.29 | Brightness of GaAsP diodes versus wavelength (or versus bandgap energy). (From Yang [22].)

of a GaAs LED is not in the visible range. For a visible output, the wavelength of the signal should be in the range of 0.4 to 0.72 μm . This range of wavelengths corresponds to bandgap energies between approximately 1.7 and 3.1 eV.

$\text{GaAs}_{1-x}\text{P}_x$ is a direct bandgap material for $0 \leq x \leq 0.45$, as shown in Figure 14.24. At $x = 0.40$, the bandgap energy is approximately $E_g = 1.9$ eV, which would produce an optical output in the red range. Figure 14.29 shows the brightness of $\text{GaAs}_{1-x}\text{P}_x$ diodes for different values of x . The peak also occurs in the red range. By using planar technology, $\text{GaAs}_{0.6}\text{P}_{0.4}$ monolithic arrays have been fabricated for numeric and alphanumeric displays. When the mole fraction x is greater than 0.45, the material changes to an indirect bandgap semiconductor so that the quantum efficiency is greatly reduced.

$\text{GaAl}_x\text{As}_{1-x}$ can be used in a heterojunction structure to form an LED. A device structure is shown in Figure 14.30. Electrons are injected from the wide-bandgap N - $\text{GaAl}_{0.7}\text{As}_{0.3}$ into the narrow-bandgap p - $\text{GaAl}_{0.6}\text{As}_{0.4}$. The minority carrier electrons

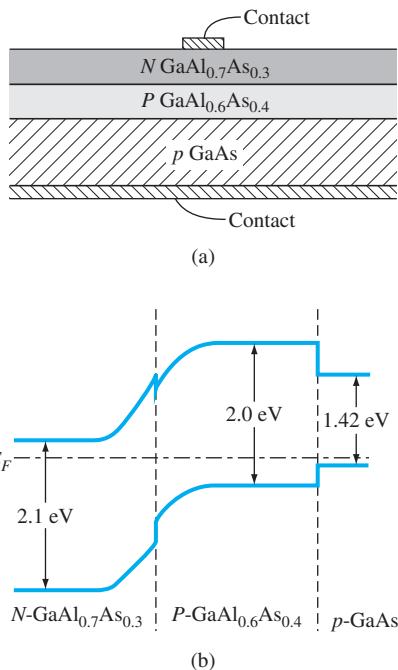


Figure 14.30 | The (a) cross section and (b) thermal equilibrium energy-band diagram of a GaAlAs heterojunction LED. (From Yang [22].)

in the p material can recombine radiatively. Since $E_{gp} < E_{gN}$, the photons are emitted through the wide-bandgap N material with essentially no absorption. The wide bandgap N material acts as an optical window and the external quantum efficiency increases.

14.6 | LASER DIODES

The photon output of the LED is due to an electron giving up energy as it makes a transition from the conduction band to the valence band. The LED photon emission is spontaneous in that each band-to-band transition is an independent event. The spontaneous emission process yields a spectral output of the LED with a fairly wide bandwidth. If the structure and operating condition of the LED are modified, the device can operate in a new mode, producing a coherent spectral output with a bandwidth of wavelengths less than 0.1 nm. This new device is a laser diode, where laser stands for **L**ight **A**mplification by **S**timulated **E**mission of **R**adiation. Although there are many different types of lasers, we are here concerned only with the pn junction laser diode.

14.6.1 Stimulated Emission and Population Inversion

Figure 14.31a shows the case when an incident photon is absorbed and an electron is elevated from an energy state E_1 to an energy state E_2 . This process is known as induced absorption. If the electron spontaneously makes the transition back to the lower energy level with a photon being emitted, we have a spontaneous emission process as indicated in Figure 14.31b. On the other hand, if there is an incident photon at a time when an electron is in the higher energy state as shown in Figure 14.31c, the incident photon can interact with the electron, causing the electron to make a transition downward. The downward transition produces a photon. Since this process was initiated by the incident photon, the process is called *stimulated* or *induced emission*. Note that this stimulated emission process has produced two photons; thus, we can have optical gain or amplification. The two emitted photons are in phase so that the spectral output will be coherent.

In thermal equilibrium, the electron distribution in a semiconductor is determined by the Fermi–Dirac statistics. If the Boltzmann approximation applies, then we can write

$$\frac{N_2}{N_1} = \exp\left[\frac{-(E_2 - E_1)}{kT}\right] \quad (14.62)$$

where N_1 and N_2 are the electron concentrations in the energy levels E_1 and E_2 , respectively, and where $E_2 > E_1$. In thermal equilibrium, $N_2 < N_1$. The probability of an induced absorption event is exactly the same as that of an induced emission event. The number of photons absorbed is proportional to N_1 and the number of additional photons emitted is proportional to N_2 . In order to achieve optical amplification or for lasing action to occur, we must have $N_2 > N_1$; this is called population inversion. We cannot achieve lasing action at thermal equilibrium.

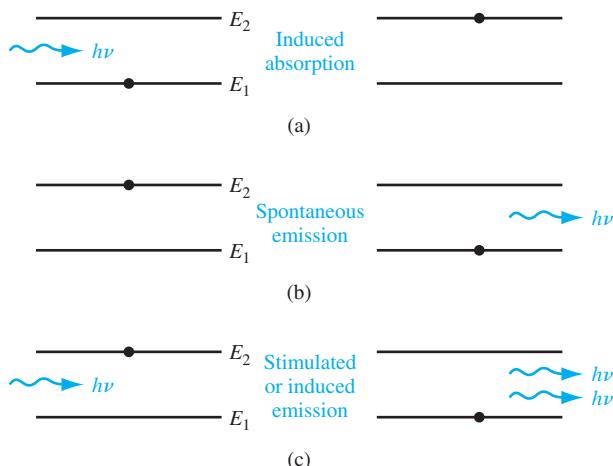


Figure 14.31 | Schematic diagram showing (a) induced absorption, (b) spontaneous emission, and (c) stimulated or induced emission processes.

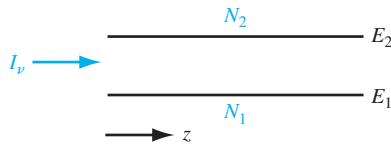


Figure 14.32 | Light propagating in z direction through a material with two energy levels.

Figure 14.32 shows the two energy levels with a light wave at an intensity I_ν propagating in the z direction. The change in intensity as a function of z can be written as

$$\frac{dI_\nu}{dz} \propto \frac{\# \text{ photons emitted}}{\text{cm}^3} - \frac{\# \text{ photons absorbed}}{\text{cm}^3}$$

or

$$\frac{dI_\nu}{dz} = N_2 W_i \cdot h\nu - N_1 W_i \cdot h\nu \quad (14.63)$$

where W_i is the induced transition probability. Equation (14.63) assumes no loss mechanisms and neglects the spontaneous transitions.

Equation (14.63) can be written as

$$\frac{dI_\nu}{dz} = \gamma(\nu) I_\nu \quad (14.64)$$

where $\gamma(\nu) \propto (N_2 - N_1)$ and is the amplification factor. From Equation (14.64), the intensity is

$$I_\nu = I_\nu(0) e^{\gamma(\nu)z} \quad (14.65)$$

Amplification occurs when $\gamma(\nu) > 0$ and absorption occurs when $\gamma(\nu) < 0$.

We can achieve population inversion and lasing in a forward-biased pn homojunction diode, if both sides of the junction are degenerately doped. Figure 14.33a shows the energy-band diagram of a degenerately doped pn junction in thermal equilibrium. The Fermi level is in the conduction band in the n-region and the Fermi level is in the valence band in the p region. Figure 14.33b shows the energy bands of the pn junction when a forward bias is applied. The gain factor in a pn homojunction diode is given by

$$\gamma(\nu) \propto \left\{ 1 - \exp \left[\frac{h\nu - (E_{Fn} - E_{Fp})}{kT} \right] \right\} \quad (14.66)$$

In order for $\gamma(\nu) > 1$, we must have $h\nu < (E_{Fn} - E_{Fp})$, which implies that the junction must be degenerately doped since we also have the requirement that $h\nu \geq E_g$. In the vicinity of the junction, there is a region in which population inversion occurs. There are large numbers of electrons in the conduction band directly above a large number of empty states. If band-to-band recombination occurs, photons will be emitted with energies in the range $E_g < h\nu < (E_{Fn} - E_{Fp})$.

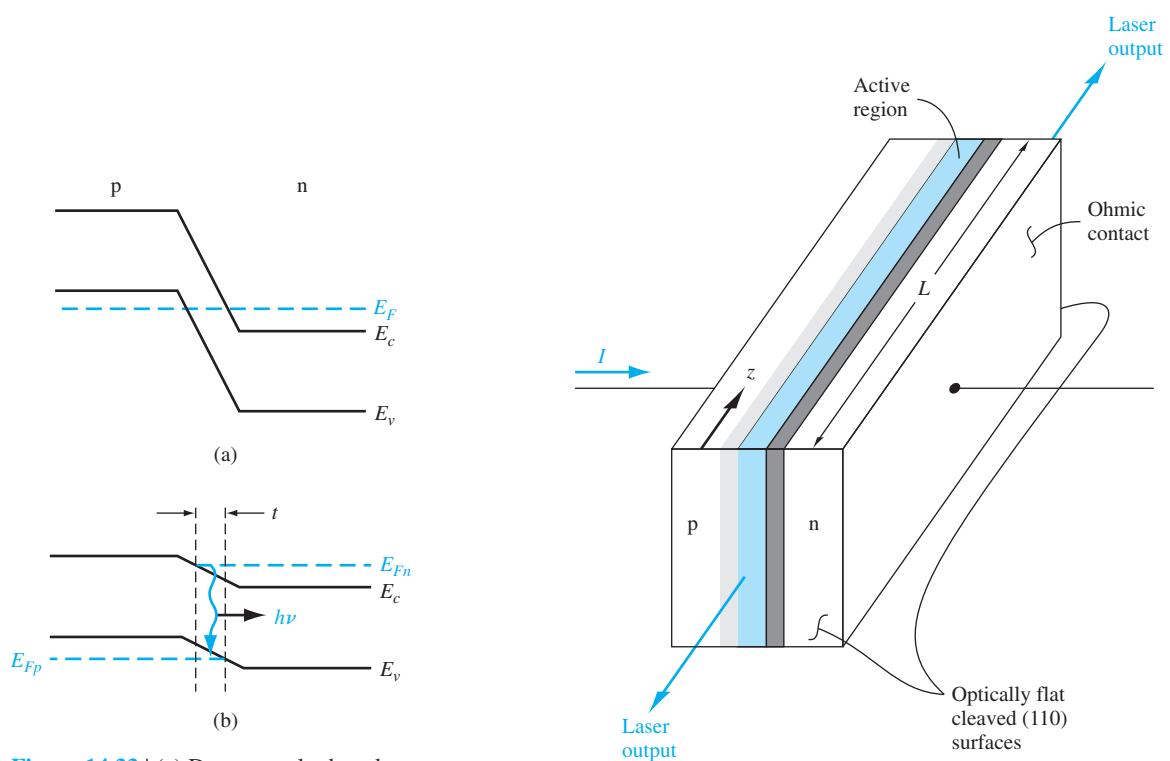


Figure 14.33 | (a) Degenerately doped pn junction at zero bias. (b) Degenerately doped pn junction under forward bias with photon emission.

Figure 14.34 | A pn junction laser diode with cleaved (110) planes forming the Fabry-Perot cavity.
(After Yang [22].)

14.6.2 Optical Cavity

Population inversion is one requirement for lasing action to occur. Coherent emission output is achieved by using an optical cavity. The cavity will cause a buildup of the optical intensity from positive feedback. A resonant cavity consisting of two parallel mirrors is known as a Fabry-Perot resonator. The resonant cavity can be fabricated, for example, by cleaving a gallium arsenide crystal along the (110) planes as shown in Figure 14.34. The optical wave propagates through the junction in the z direction, bouncing back and forth between the end mirrors. The mirrors are actually only partially reflecting so that a portion of the optical wave will be transmitted out of the junction.

For resonance, the length of the cavity L must be an integral number of half wavelengths, or

$$N \left(\frac{\lambda}{2} \right) = L \quad (14.67)$$

where N is an integer. Since λ is small and L is relatively large, there can be many resonant modes in the cavity. Figure 14.35a shows the resonant modes as a function of wavelength.

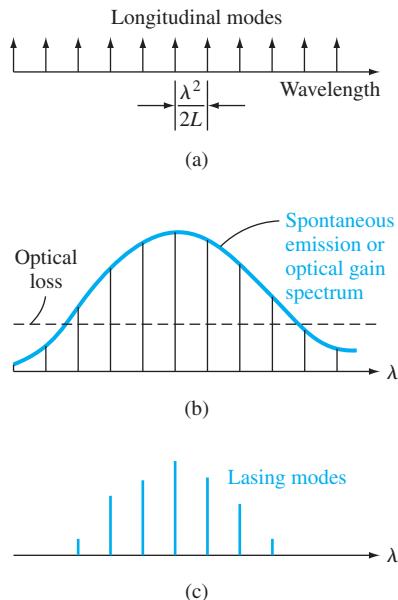


Figure 14.35 | Schematic diagram showing (a) resonant modes of a cavity with length L , (b) spontaneous emission curve, and (c) actual emission modes of a laser diode.
(After Yang [22].)

When a forward-bias current is applied to the pn junction, spontaneous emission will initially occur. The spontaneous emission spectrum is relatively broadband and is superimposed on the possible lasing modes as shown in Figure 14.35b. In order for lasing to be initiated, the spontaneous emission gain must be larger than the optical losses. By positive feedback in the cavity, lasing can occur at several specific wavelengths as indicated in Figure 14.35c.

14.6.3 Threshold Current

The optical intensity in the device can be written from Equation (14.65) as $I_\nu \propto e^{\gamma(\nu)z}$, where $\gamma(\nu)$ is the amplification factor. We have two basic loss mechanisms. The first is the photon absorption in the semiconductor material. We can write

$$I_\nu \propto e^{-\alpha(\nu)z} \quad (14.68)$$

where $\alpha(\nu)$ is the absorption coefficient. The second loss mechanism is due to the partial transmission of the optical signal through the ends, or through the partially reflecting mirrors.

At the onset of lasing, which is known as threshold, the optical loss of one round trip through the cavity is just offset by the optical gain. The threshold condition is then expressed as

$$\Gamma_1 \Gamma_2 \exp [(2\gamma_t(\nu) - 2\alpha(\nu))L] = 1 \quad (14.69)$$

where Γ_1 and Γ_2 are the reflectivity coefficients of the two end mirrors. For the case when the optical mirrors are cleaved (110) surfaces of gallium arsenide, the reflectivity coefficients are given approximately by

$$\Gamma_1 = \Gamma_2 = \left(\frac{\bar{n}_2 - \bar{n}_1}{\bar{n}_2 + \bar{n}_1} \right)^2 \quad (14.70)$$

where \bar{n}_2 and \bar{n}_1 are the index of refraction parameters for the semiconductor and air, respectively. The parameter $\gamma_t(\nu)$ is the optical gain at threshold.

The optical gain at threshold, $\gamma_t(\nu)$, may be determined from Equation (14.69) as

$$\gamma_t(\nu) = \alpha + \frac{1}{2L} \ln \left(\frac{1}{\Gamma_1 \Gamma_2} \right) \quad (14.71)$$

Since the optical gain is a function of the pn junction current, we can define a threshold current density as

$$J_{th} = \frac{1}{\beta} \left[\alpha + \frac{1}{2L} \ln \left(\frac{1}{\Gamma_1 \Gamma_2} \right) \right] \quad (14.72)$$

where β can be determined theoretically or experimentally. Figure 14.36 shows the threshold current density as a function of the mirror losses. We may note the relatively high threshold current density for a pn junction laser diode.

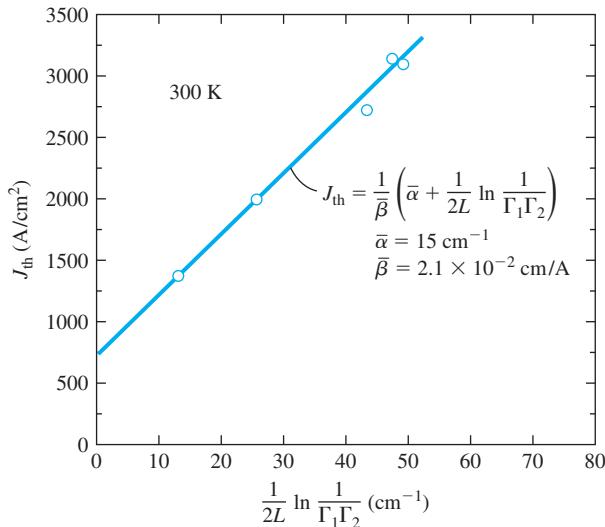


Figure 14.36 | Threshold current density of a laser diode as a function of Fabry-Perot cavity end losses.
(After Yang [22].)

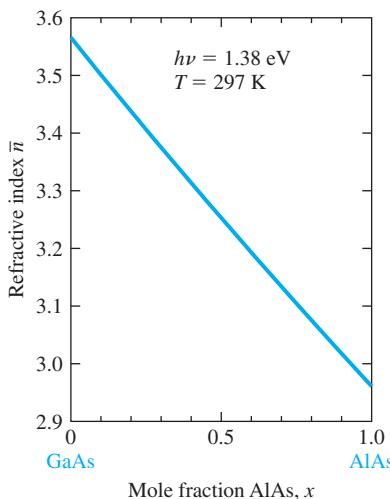


Figure 14.37 | Index of refraction of $\text{Al}_x\text{Ga}_{1-x}\text{As}$ as a function of mole fraction x .
(From Sze [18].)

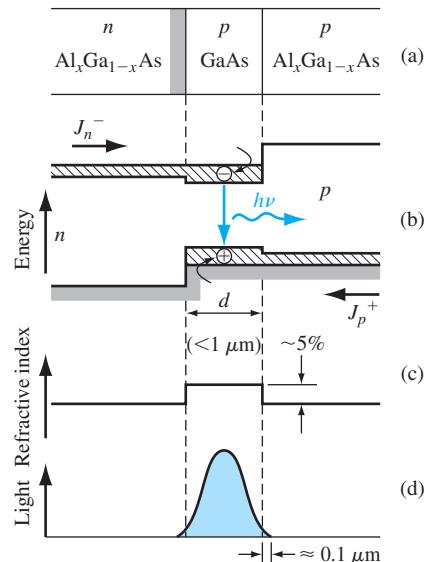


Figure 14.38 | (a) Basic double heterojunction structure. (b) Energy-band diagram under forward bias. (c) Refractive index change through the structure. (d) Confinement of light in the dielectric waveguide.
(From Yang [22].)

14.6.4 Device Structures and Characteristics

We have seen that in a homojunction LED, the photons may be emitted in any direction, which lowers the external quantum efficiency. Significant improvement in device characteristics can be made if the emitted photons are confined to a region near the junction. This confinement can be achieved by using an optical dielectric waveguide. The basic device is a three-layered, double heterojunction structure known as a double heterojunction laser. A requirement for a dielectric waveguide is that the index of refraction of the center material be larger than that of the other two dielectrics. Figure 14.37 shows the index of refraction for the AlGaAs system. We may note that GaAs has the highest index of refraction.

An example of a double heterojunction laser is shown in Figure 14.38a. A thin p-GaAs layer is between P-AlGaAs and N-AlGaAs layers. A simplified energy-band diagram is shown in Figure 14.38b for the forward-biased diode. Electrons are injected from the N-AlGaAs into the p-GaAs. Population inversion is easily obtained since the conduction band potential barrier prevents the electrons from diffusing into the P-AlGaAs region. Radiative recombination is then confined to the p-GaAs region. Since the index of refraction of GaAs is larger than that of AlGaAs, the light wave is also confined to the GaAs region. An optical cavity can be formed by cleaving the semiconductor perpendicular to the N-AlGaAs–p-GaAs junction.

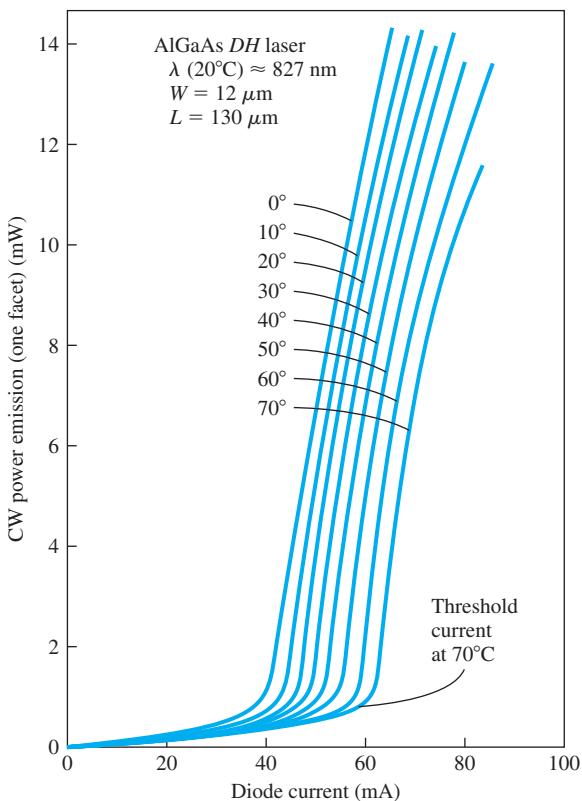


Figure 14.39 | Typical output power versus laser diode current at various temperatures.
(From Yang [22].)

Typical optical output versus diode current characteristics are shown in Figure 14.39. The threshold current is defined to be the current at the breakpoint. At low currents, the output spectrum is very wide and is the result of the spontaneous transitions. When the diode current is slightly above the threshold value, the various resonant frequencies are observed. When the diode current becomes large, a single dominant mode with a narrow bandwidth is produced.

The performance of the laser diode can be further improved if a very narrow recombination region is used with a somewhat wider optical waveguide. Very complex structures using multilayers of compound semiconductor materials have been fabricated in a continuing effort to improve semiconductor laser performance.

14.7 | SUMMARY

- The absorption or emission of light (photons) in semiconductors leads to the study of a general class of devices called optoelectronics. A few of these devices have been discussed and analyzed in this chapter.

- The photon absorption process has been discussed and the absorption coefficient data for semiconductors has been presented.
- Solar cells convert optical power into electrical power. The simple pn junction solar cell was initially considered. The short-circuit current, open-circuit voltage, and maximum power were considered.
- Heterojunction and amorphous silicon solar cells were also considered. Heterojunction cells can be fabricated that tend to increase the conversion efficiency and produce relatively large open-circuit voltages. Amorphous silicon offers the possibility of low-cost, large-area solar cell arrays.
- Photodetectors are semiconductor devices that convert optical signals into electrical signals. The photoconductor is perhaps the simplest type of photodetector. The change in conductivity of the semiconductor due to the creation of excess electrons and holes by the incident photons is the basis of this device.
- Photodiodes are diodes that have reverse-biased voltages applied. Excess carriers that are created by incident photons in the space-charge region are swept out by the electric field creating a photocurrent. The photocurrent is directly proportional to the incident photon intensity. PIN and avalanche photodiodes are variations of the basic photodiode.
- The photocurrent generated in a phototransistor is multiplied by the transistor gain. However, the time response of the phototransistor may be slower than that of a photodiode because of the Miller effect and Miller capacitance.
- The inverse mechanism of photon absorption in a pn junction is injection electroluminescence. The recombination of excess electrons and holes in a direct bandgap semiconductor can result in the emission of photons.
- The light emitting diodes (LEDs) are the class of pn junction diodes whose photon output is a result of spontaneous recombinations of excess electrons and holes. A fairly wide bandwidth in the output signal, on the order of 30 nm, is a result of the spontaneous process.
- The output of a laser diode is the result of stimulated emission. An optical cavity, or Fabry–Perot resonator, is used in conjunction with a diode so that the photon output is in phase, or coherent. Multilayered heterojunction structures can be fabricated to improve the laser diode characteristics.

GLOSSARY OF IMPORTANT TERMS

absorption coefficient The relative number of photons absorbed per unit distance in a semiconductor and denoted by the parameter α .

conversion efficiency The ratio of output electrical power to incident optical power in a solar cell.

delayed photocurrent The component of photocurrent in a semiconductor device due to diffusion currents.

external quantum efficiency The ratio of emitted photons to generated photons in a semiconductor device.

fill factor The ratio $I_m V_m$ to $I_{sc} V_{oc}$, which is a measure of the realizable power from a solar cell. The parameters I_m and V_m are the current and voltage at the maximum power point, respectively, and I_{sc} and V_{oc} are the short-circuit current and open-circuit voltage.

fresnel loss The ratio of reflected to incident photons at an interface due to a change in the index of refraction.

internal quantum efficiency The fraction of diode current that produces luminescence.

LASER diode An acronym for Light Amplification by Stimulated Emission of Radiation; the stimulated emission of photons produced in a forward-biased pn junction in conjunction with an optical cavity.

LED An acronym for Light Emitting Diode; the spontaneous photon emission due to electron-hole recombination in a forward-biased pn junction.

luminescence The general property of light emission.

open-circuit voltage The voltage generated across the open-circuited terminals of a solar cell.

photocurrent The current generated in a semiconductor device due to the flow of excess carriers generated by the absorption of photons.

population inversion The condition whereby the concentration of electrons in one energy state is greater than that in a lower energy state; a nonequilibrium condition.

prompt photocurrent The component of photocurrent generated within the space charge region of a semiconductor device.

radiative recombination The recombination process of electrons and holes that produces a photon, such as the direct band-to-band transition in gallium arsenide.

short-circuit current The current produced in a solar cell when the two terminals are shorted together.

stimulated emission The process whereby an electron is induced by an incident photon to make a transition to a lower energy state, emitting a second photon.

CHECKPOINT

After studying this chapter, the reader should have the ability to:

- Describe the optical absorption process in semiconductors. When is optical absorption essentially zero?
- Describe the basic operation and characteristics of a solar cell, including the short-circuit current and open-circuit voltage.
- Discuss the factors that contribute to the solar cell conversion efficiency.
- Describe the advantages and disadvantages of an amorphous silicon solar cell.
- Describe the characteristics of a photoconductor, including the concept of the photoconductor gain.
- Discuss the operation and characteristics of a simple pn junction photodiode.
- Discuss the advantages of PIN and avalanche photodiodes compared to the simple pn junction photodiode.
- Discuss the operation and characteristics of a phototransistor.
- Describe the operation of an LED.
- Describe the operation of a laser diode.

REVIEW QUESTIONS

1. Sketch the general shape of the optical absorption coefficient in a semiconductor as a function of wavelength. When does the absorption coefficient become zero?
2. Sketch the I - V characteristic of a pn junction solar cell. Define short-circuit current and open-circuit voltage.

3. Discuss how a pn junction solar cell becomes forward biased.
4. Write an expression for the steady-state photocurrent in a simple photoconductor.
5. What is the source of prompt photocurrent in a photodiode? Does the prompt photocurrent depend on the reverse-biased voltage? Why or why not.
6. Sketch the cross section of a phototransistor and show the currents that are created by incident photons. Explain how current gain is achieved.
7. Explain the basic operation of an LED. State two factors that affect the efficiency of the device.
8. How can different colors be obtained in an LED?
9. Discuss the difference between an LED and a laser diode.
10. Discuss the concept of population inversion in a laser diode.

PROBLEMS

Section 14.1 Optical Absorption

- 14.1** Determine the maximum wavelength λ of a light source that can generate electron–hole pairs in (a) Si, (b) Ge, (c) GaAs, and (d) InP.
- 14.2** (a) Two sources generate light at wavelengths of $\lambda = 480$ nm and $\lambda = 725$ nm, respectively. What are the corresponding photon energies? (b) Three sources generate light with photon energies of $E = 0.87$ eV, $E = 1.32$ eV, and $E = 1.90$ eV, respectively. What are the corresponding wavelengths?
- 14.3** (a) A sample of GaAs is $1.2 \mu\text{m}$ thick. The sample is illuminated with a light source that generates photons with energies of $h\nu = 1.65$ eV. Determine the (i) absorption coefficient and (ii) fraction of energy that is absorbed in the material. (b) Repeat part (a) for a sample of GaAs that is $0.80 \mu\text{m}$ thick and is illuminated with photons with energies of $h\nu = 1.90$ eV.
- 14.4** A light source with $h\nu = 1.3$ eV and at a power density of 10^{-2} W/cm^2 is incident on a thin slab of silicon. The excess minority carrier lifetime is 10^{-6} s. Determine the electron–hole generation rate and the steady-state excess carrier concentration. Neglect surface effects.
- 14.5** An n-type GaAs sample has a minority carrier lifetime of $\tau_p = 2 \times 10^{-7}$ s. Incident photons with energies $h\nu = 1.65$ eV generate an excess carrier concentration of $\delta p = 5 \times 10^{15} \text{ cm}^{-3}$ at the surface of the semiconductor. (a) Determine the incident power required. (b) At what distance in the semiconductor does the generation rate drop to 10 percent of that at the surface?
- 14.6** Consider a silicon semiconductor that is illuminated with photons with energies $h\nu = 1.40$ eV. (a) Determine the thickness of the material such that 90 percent of the energy is absorbed. (b) Determine the thickness of the material such that 30 percent of the energy is transmitted through the material.
- 14.7** If the thickness of a GaAs semiconductor is $1 \mu\text{m}$ and 50 percent of the incident monochromatic photon energy is absorbed, determine the incident photon energy and wavelength.
- *14.8** Consider monochromatic light at an intensity $I_{\nu 0}$ incident on the surface at $x = 0$ of an n-type semiconductor that extends to $x = \infty$. Assume the electric field is zero in the semiconductor and assume a surface recombination velocity, s . Taking into

*Asterisks next to problems indicate problems that are more difficult.

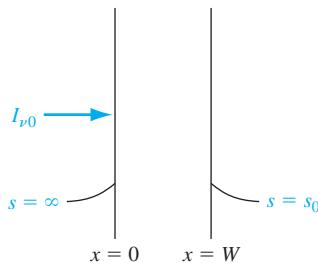


Figure P14.9 | Figure for Problem 14.9.

account the absorption coefficient, determine the steady-state excess hole concentration as a function of x .

- *14.9 Monochromatic light with intensity $I_{\nu 0}$ is incident on a p-type semiconductor as shown in Figure P14.9. Assume the surface recombination velocity at $x = 0$ is $s = \infty$ and assume the surface recombination velocity at $x = W$ is $s = s_0$. Derive the expression for the steady-state excess electron concentration as a function of x .

Section 14.2 Solar Cells

- 14.10 A long silicon pn junction solar cell at $T = 300$ K has the following parameters: $N_a = 10^{16} \text{ cm}^{-3}$, $N_d = 10^{15} \text{ cm}^{-3}$, $D_n = 25 \text{ cm}^2/\text{s}$, $D_p = 10 \text{ cm}^2/\text{s}$, $\tau_{n0} = 10^{-6} \text{ s}$, and $\tau_{p0} = 5 \times 10^{-7} \text{ s}$. The cross-sectional area of the solar cell is 5 cm^2 . The entire junction is uniformly illuminated such that the generation rate of electron–hole pairs is $G_L = 5 \times 10^{21} \text{ cm}^{-3} \text{ s}^{-1}$. (a) Calculate the short circuit photocurrent generated in the space charge region. (b) Using the results of part (a), calculate the open-circuit voltage. (c) Determine the ratio of V_{oc} to V_{bi} .
- 14.11 A long silicon pn junction solar cell has the same parameters as described in Problem 14.10. The generated photocurrent in the cell is $I_L = 120 \text{ mA}$. Determine the (a) open-circuit voltage, (b) the voltage across the junction that will produce a total solar cell current of $I = 100 \text{ mA}$, (c) the maximum power output of the solar cell, and (d) the external load resistance that will produce the maximum power.
- 14.12 Consider the solar cell described in Problem 14.10. (a) The generated photocurrent is $I_L = 10 \text{ mA}$. Determine (i) the open-circuit voltage and (ii) the maximum power output. (b) The solar cell now uses a solar concentrator such that the photocurrent increases by a factor of 10. Determine the new values of (i) open circuit voltage and (ii) maximum power output. (c) Determine the ratio of maximum power from part (b) to that from part (a).
- 14.13 Consider an ideal long n⁺p junction GaAs solar cell at $T = 300$ K in which excess carriers are uniformly generated. The parameters of the diode are as follows:

$$\begin{aligned} N_d &= 10^{19} \text{ cm}^{-3} & D_n &= 225 \text{ cm}^2/\text{s} \\ \tau_{n0} &= \tau_{p0} = 5 \times 10^{-8} \text{ s} & D_p &= 7 \text{ cm}^2/\text{s}. \end{aligned}$$

The generated photocurrent density is $J_L = 30 \text{ mA/cm}^2$. Plot the open-circuit voltage as a function of the acceptor doping concentration for $10^{15} \leq N_a \leq 10^{18} \text{ cm}^{-3}$.

- 14.14** A long silicon pn junction solar cell with an area of 2 cm^2 has the following parameters:

$$\begin{array}{ll} N_d = 10^{19} \text{ cm}^{-3} & N_a = 3 \times 10^{16} \text{ cm}^{-3} \\ D_p = 6 \text{ cm}^2/\text{s} & D_n = 18 \text{ cm}^2/\text{s} \\ \tau_{p0} = 5 \times 10^{-7} \text{ s} & \tau_{n0} = 5 \times 10^{-6} \text{ s} \end{array}$$

Assume that excess carriers are uniformly generated in the solar cell and that $J_L = 25 \text{ mA/cm}^2$. Let $T = 300 \text{ K}$. (a) Plot the I - V characteristics of the diode, (b) determine the maximum power output of the solar cell, and (c) calculate the external load resistance that will produce the maximum power.

- 14.15** A silicon solar cell at $T = 300 \text{ K}$ has a cross-sectional area of 6 cm^2 and a reverse saturation current of $I_s = 2 \times 10^{-9} \text{ A}$. The induced short-circuit photocurrent is $I_L = 180 \text{ mA}$. Determine the (a) open-circuit voltage, (b) maximum power output, and (c) load resistance that will produce the maximum output power. (d) If the load resistance determined in part (c) is increased by 50 percent, what is the new value of the maximum output power?
- 14.16** Consider a silicon solar cell at $T = 300 \text{ K}$ with a reverse saturation current of $I_s = 10^{-10} \text{ A}$ and an induced short-circuit photocurrent of $I_L = 100 \text{ mA}$. (a) Determine V_{oc} . (b) Find V_m , I_m , and P_m . (c) How many cells, operating at the maximum output power, must be connected in series to produce an output voltage of at least 10 V ? (d) How many of the 10 V cells in part (c) must be connected in parallel to produce an output power of at least 5.2 W ? (e) Considering the results of part (d), what must be the load resistance connected across the solar cell system to produce the maximum output power?
- *14.17** Consider the pn junction solar cell with nonuniform absorption. Derive the expression for the excess minority carrier electron concentration for the short-circuit condition and for the case when the p region is very long and the n region is short.
- 14.18** The absorption coefficient in amorphous silicon is approximately 10^4 cm^{-1} at $h\nu = 1.7 \text{ eV}$ and 10^5 cm^{-1} at $h\nu = 2.0 \text{ eV}$. Determine the amorphous silicon thickness for each case so that 90 percent of the photons are absorbed.

Section 14.3 Photodetectors

- 14.19** Consider an n-type silicon photoconductor at $T = 300 \text{ K}$ doped at $N_d = 5 \times 10^{15} \text{ cm}^{-3}$. The cross-sectional area is $A = 5 \times 10^{-4} \text{ cm}^2$ and the length is $L = 120 \mu\text{m}$. The carrier parameters are $\mu_n = 1200 \text{ cm}^2/\text{V}\cdot\text{s}$, $\mu_p = 400 \text{ cm}^2/\text{V}\cdot\text{s}$, $\tau_{n0} = 5 \times 10^{-7} \text{ s}$, and $\tau_{p0} = 10^{-7} \text{ s}$. The photoconductor is uniformly illuminated such that the generation rate of electron-hole pairs is $G_L = 10^{21} \text{ cm}^{-3} \text{ s}^{-1}$. For 3 volts applied to the photoconductor, determine (a) the thermal equilibrium current, (b) the steady-state excess carrier concentration, (c) the photoconductivity, (d) the steady-state photocurrent, and (e) the photocurrent gain.
- 14.20** Excess carriers are uniformly generated in a GaAs photoconductor at a rate of $G_L = 10^{21} \text{ cm}^{-3} \text{ s}^{-1}$. The area is $A = 10^{-4} \text{ cm}^2$ and the length is $L = 100 \mu\text{m}$. The other parameters are:

$$\begin{array}{ll} N_d = 5 \times 10^{16} \text{ cm}^{-3} & N_a = 0 \\ \mu_n = 8000 \text{ cm}^2/\text{V}\cdot\text{s} & \mu_p = 250 \text{ cm}^2/\text{V}\cdot\text{s} \\ \tau_{n0} = 10^{-7} \text{ s} & \tau_{p0} = 10^{-8} \text{ s.} \end{array}$$

If a voltage of 5 volts is applied, calculate (a) the steady-state excess carrier concentration, (b) the photoconductivity, (c) the steady-state photocurrent, and (d) the photoconductor gain.

- *14.21** Consider an n-type silicon photoconductor that is $1 \mu\text{m}$ thick, $50 \mu\text{m}$ wide, and has an applied electric field in the longitudinal dimension of 50 V/cm . If the incident photon flux is $\Phi_0 = 10^{16} \text{ cm}^{-2}\text{s}^{-1}$ and the absorption coefficient is $\alpha = 5 \times 10^4 \text{ cm}^{-1}$, calculate the steady-state photocurrent if $\mu_n = 1200 \text{ cm}^2/\text{V}\cdot\text{s}$, $\mu_p = 450 \text{ cm}^2/\text{V}\cdot\text{s}$, and $\tau_{p0} = 2 \times 10^{-7} \text{ s}$.
- 14.22** A long silicon pn junction photodiode has the following parameters at $T = 300 \text{ K}$: $N_a = 10^{16} \text{ cm}^{-3}$, $N_d = 2 \times 10^{15} \text{ cm}^{-3}$, $D_p = 10 \text{ cm}^2/\text{s}$, $D_n = 25 \text{ cm}^2/\text{s}$, $\tau_{p0} = 10^{-7} \text{ s}$, and $\tau_{n0} = 5 \times 10^{-7} \text{ s}$. The cross-sectional area of the diode is $A = 10^{-3} \text{ cm}^2$. Assume that a reverse-biased voltage of 5 volts is applied and that a uniform generation rate for electron-hole pairs of $G_L = 10^{21} \text{ cm}^{-3} \text{ s}^{-1}$ exists throughout the entire photodiode. (a) Determine the prompt component of photocurrent. (b) Find the steady-state excess carrier concentrations in the p and n regions far from the junction. (c) Determine the total steady-state photocurrent.
- *14.23** Starting with the ambipolar transport equation for minority carrier holes, derive Equation (14.41) using the geometry shown in Figure 14.17.
- 14.24** Three silicon PIN photodiodes A, B, and C, at $T = 300 \text{ K}$ have intrinsic region widths of 2 , 10 , and $80 \mu\text{m}$, respectively. A photon flux of $\Phi_0 = 5 \times 10^{17} \text{ cm}^{-2} \text{ s}^{-1}$ is incident on the surface of each diode as shown in Figure 14.19. (a) For an absorption coefficient of $\alpha = 10^4 \text{ cm}^{-1}$, calculate the prompt photocurrent density in each diode. (b) Repeat part (a) for an absorption coefficient of $\alpha = 5 \times 10^2 \text{ cm}^{-1}$.
- 14.25** Consider a silicon PIN photodiode at $T = 300 \text{ K}$ with the geometry shown in Figure 14.19. The intrinsic region width is $100 \mu\text{m}$. Assume that a reverse-biased voltage is applied such that the intrinsic region is completely depleted. The incident photon power is $I_{\nu 0} = 0.080 \text{ W/cm}^2$, the absorption coefficient is $\alpha = 10^3 \text{ cm}^{-1}$, and the photon energy is 1.5 eV . Neglect any absorption in the p^+ top layer of the photodiode. (a) Determine the steady-state electron-hole generation rate, G_L , versus distance in the intrinsic region. (b) Determine the steady-state photocurrent density.
- 14.26** A silicon PIN photodiode at $T = 300 \text{ K}$ has the geometry shown in Figure 14.19. The intrinsic region width is $20 \mu\text{m}$ and is fully depleted. (a) The electron-hole pair generation rate in the intrinsic region is $G_L = 10^{21} \text{ cm}^{-3} \text{ s}^{-1}$ and is uniform throughout the intrinsic region. Calculate the steady-state photocurrent density for this condition. (b) The generation rate of electron-hole pairs is $G_L = 10^{21} \text{ cm}^{-3} \text{ s}^{-1}$ at $x = 0$ and the absorption coefficient is $\alpha = 10^3 \text{ cm}^{-1}$. Determine the steady-state photocurrent density for this situation.
- 14.27** Consider a silicon PIN photodiode exposed to sunlight. Calculate the intrinsic region width so that at least 90 percent of all photons with wavelengths $\lambda \leq 1 \mu\text{m}$ are absorbed in the intrinsic region. Neglect any absorption in the p^+ or n^+ regions.

Section 14.4 Photoluminescence and Electroluminescence

- 14.28** Consider the $\text{Al}_x\text{Ga}_{1-x}\text{As}$ system. Determine the range of the direct bandgap energies possible and the corresponding range of wavelengths.
- 14.29** Consider the $\text{GaAs}_{1-x}\text{P}_x$ system. (a) For a mole fraction $x = 0.2$, determine the (i) bandgap energy and (ii) corresponding photon wavelength. (b) Repeat part (a) for a mole fraction $x = 0.32$.

- 14.30** Using Figure 14.23, determine the mole fraction x in $\text{Al}_x\text{Ga}_{1-x}\text{As}$ such that the material would emit light at a wavelength of $\lambda = 0.670 \mu\text{m}$. What is the corresponding bandgap energy?
- 14.31** Repeat Problem 14.30 for the $\text{GaAs}_{1-x}\text{P}_x$ system.

Section 14.5 Light Emitting Diodes

- 14.32** Consider a pn junction GaAs LED. Assume that photons are generated uniformly in all directions in a plane perpendicular to the junction at a distance of $0.50 \mu\text{m}$ from the surface. (a) Taking into account total internal reflection, calculate the fraction of photons that have the potential of being emitted from the semiconductor. (b) Using the results of part (a) and including Fresnel loss, determine the fraction of generated photons that will be emitted from the semiconductor into air (neglect absorption losses).
- *14.33** In a pn junction LED, consider a point source in the semiconductor at the junction and assume that photons are emitted uniformly in all directions. Show that (neglecting photon absorption) the external quantum efficiency of the LED is given by

$$\eta_{\text{ext}} = \frac{2\bar{n}_1\bar{n}_2}{(\bar{n}_1 + \bar{n}_2)^2}(1 - \cos \theta_c)$$

where \bar{n}_1 and \bar{n}_2 are the index of refraction parameters for the air and semiconductor, respectively, and θ_c is the critical angle.

Section 14.6 Laser Diodes

- 14.34** Consider an optical cavity. If $N \gg 1$, show that the wavelength separation between two adjacent resonant modes is $\Delta\lambda = \lambda^2/2L$.
- 14.35** If the photon output of a laser diode is equal to the bandgap energy, find the wavelength separation between adjacent resonant modes in a GaAs laser with $L = 75 \mu\text{m}$.

READING LIST

1. Bhattacharya, P. *Semiconductor Optoelectronic Devices*, 2nd ed. Upper Saddle River, NJ: Prentice Hall, 1997.
2. Carlson, D. E. "Amorphous Silicon Solar Cells." *IEEE Transactions on Electron Devices* ED-24 (April 1977), pp. 449–53.
3. Fonash, S. J. *Solar Cell Device Physics*. New York: Academic Press, 1981.
4. Kano, K. *Semiconductor Devices*. Upper Saddle River, NJ: Prentice Hall, 1998.
5. Kressel, H. *Semiconductor Devices for Optical Communications: Topics in Applied Physics*. Vol. 39. New York: Springer-Verlag, 1987.
6. MacMillan, H. F., H. C. Hamaker, G. F. Virshup, and J. G. Werthen. "Multijunction III-V Solar Cells: Recent and Projected Results." *Twentieth IEEE Photovoltaic Specialists Conference* (1988), pp. 48–54.
7. Madan, A. "Amorphous Silicon: From Promise to Practice." *IEEE Spectrum* 23 (September 1986), pp. 38–43.
8. Pankove, J. I. *Optical Processes in Semiconductors*. New York: Dover Publications, 1971.

9. Pierret, R. F. *Semiconductor Device Fundamentals*. Reading, MA: Addison-Wesley, 1996.
10. Roulston, D. J. *An Introduction to the Physics of Semiconductor Devices*. New York: Oxford University Press, 1999.
11. Schroder, D. K. *Semiconductor Material and Devices Characterization*, 3rd ed. Hoboken, NJ: John Wiley and Sons, 2006.
12. Shur, M. *Introduction to Electronic Devices*. New York: John Wiley and Sons, 1996.
- *13. _____. *Physics of Semiconductor Devices*. Englewood Cliffs, NJ: Prentice Hall, 1990.
14. Singh, J. *Optoelectronics: An Introduction to Materials and Devices*. New York: McGraw-Hill, 1996.
15. _____. *Semiconductor Devices: Basic Principles*. New York: John Wiley and Sons, 2001.
16. Streetman, B. G., and S. K. Banerjee. *Solid State Electronic Devices*, 6th ed. Upper Saddle River, NJ: Pearson Prentice-Hall, 2006.
17. Sze, S. M. *Semiconductor Devices: Physics and Technology*. New York: John Wiley and Sons, 1985.
18. Sze, S. M. and K. K. Ng. *Physics of Semiconductor Devices*, 3rd ed. Hoboken, NJ: John Wiley and Sons, 2007.
- *19. Wang, S. *Fundamentals of Semiconductor Theory and Device Physics*. Englewood Cliffs, NJ: Prentice Hall, 1989.
20. Wilson, J., and J. F. B. Hawkes. *Optoelectronics: An Introduction*. Englewood Cliffs, NJ: Prentice Hall, 1983.
- *21. Wolfe, C. M., N. Holonyak, Jr., and G. E. Stillman. *Physical Properties of Semiconductors*. Englewood Cliffs, NJ: Prentice Hall, 1989.
22. Yang, E. S. *Microelectronic Devices*. New York: McGraw-Hill, 1988.

*Indicates reference that is at an advanced level compared to this text.

10

C H A P T E R

It should be page 313 from this page!

Fundamentals of the Metal–Oxide–Semiconductor Field-Effect Transistor

The single-junction semiconductor devices that we have considered, including the pn homojunction diode, can be used to produce rectifying current–voltage characteristics and to form electronic switching circuits. The transistor is a multijunction semiconductor device that, in conjunction with other circuit elements, is capable of current gain, voltage gain, and signal power gain. The basic transistor action is the control of current at one terminal by the voltage applied across the other two terminals of the device.

The Metal–Oxide–Semiconductor Field-Effect Transistor (MOSFET) is one of two major types of transistors. The fundamental physics of the MOSFET is developed in this chapter. The MOSFET is used extensively in digital circuit applications where, because of its small size, millions of devices can be fabricated in a single integrated circuit.

Two complementary configurations of MOS transistors, the n-channel MOSFET and the p-channel MOSFET, can be fabricated. Electronic circuit design becomes very versatile when the two types of devices are used in the same circuit. These circuits are referred to as complementary MOS (CMOS) circuits. ■

10.0 | PREVIEW

In this chapter, we will:

- Study the characteristics of energy bands as a function of applied voltage in the metal–oxide–semiconductor structure known as the MOS capacitor. The MOS capacitor is the heart of the MOSFET.
- Discuss the concept of surface inversion in the semiconductor of the MOS capacitor.

- Define and derive the expression for the threshold voltage, which is a basic parameter of the MOSFET.
- Discuss various physical structures of MOSFETs, including enhancement and depletion mode devices.
- Derive the ideal current–voltage relationship of the MOSFET.
- Develop the small-signal equivalent circuit of the MOSFET. This circuit is used to relate small-signal currents and voltages in analog circuits.
- Derive the frequency limiting factors of the MOSFET.

10.1 | THE TWO-TERMINAL MOS STRUCTURE

The heart of the MOSFET is the MOS capacitor shown in Figure 10.1. The metal may be aluminum or some other type of metal, although in many cases, it is actually a high-conductivity polycrystalline silicon that has been deposited on the oxide; however, the term metal is usually still used. The parameter t_{ox} in the figure is the thickness of the oxide and ϵ_{ox} is the permittivity of the oxide.

10.1.1 Energy-Band Diagrams

The physics of the MOS structure can be more easily explained with the aid of the simple parallel-plate capacitor. Figure 10.2a shows a parallel-plate capacitor with the top plate at a negative voltage with respect to the bottom plate. An insulator material separates the two plates. With this bias, a negative charge exists on the top plate, a positive charge exists on the bottom plate, and an electric field is induced between the two plates as shown. The capacitance per unit area for this geometry is

$$C' = \frac{\epsilon}{d} \quad (10.1)$$

where ϵ is the permittivity of the insulator and d is the distance between the two plates. The magnitude of the charge per unit area on either plate is

$$Q' = C'V \quad (10.2)$$

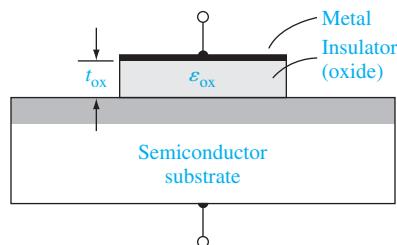


Figure 10.1 | The basic MOS capacitor structure.

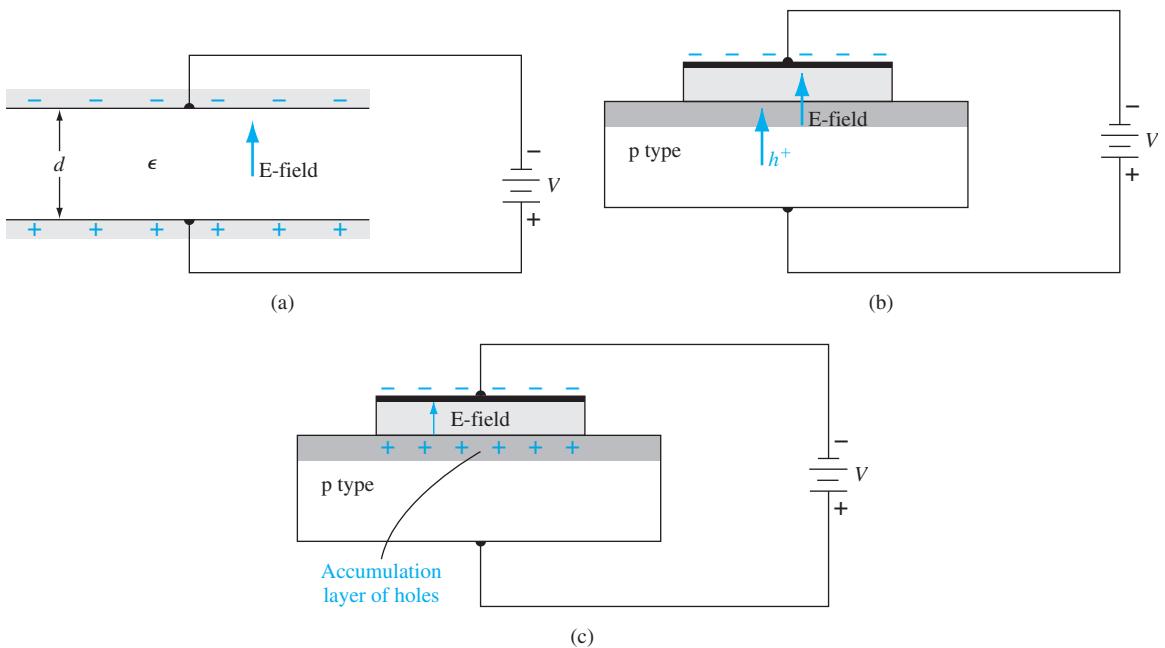


Figure 10.2 | (a) A parallel-plate capacitor showing the electric field and conductor charges. (b) A corresponding MOS capacitor with a negative gate bias showing the electric field and charge flow. (c) The MOS capacitor with an accumulation layer of holes.

where the prime indicates charge or capacitance per unit area. The magnitude of the electric field is

$$E = \frac{V}{d} \quad (10.3)$$

Figure 10.2b shows a MOS capacitor with a p-type semiconductor substrate. The top metal gate is at a negative voltage with respect to the semiconductor substrate. From the example of the parallel-plate capacitor, we can see that a negative charge will exist on the top metal plate and an electric field will be induced with the direction shown in the figure. If the electric field were to penetrate into the semiconductor, the majority carrier holes would experience a force toward the oxide–semiconductor interface. Figure 10.2c shows the equilibrium distribution of charge in the MOS capacitor with this particular applied voltage. An *accumulation layer* of holes at the oxide–semiconductor junction corresponds to the positive charge on the bottom “plate” of the MOS capacitor.

Figure 10.3a shows the same MOS capacitor in which the polarity of the applied voltage is reversed. A positive charge now exists on the top metal plate and the induced electric field is in the opposite direction as shown. If the electric field penetrates the semiconductor in this case, majority carrier holes will experience a force away from the oxide–semiconductor interface. As the holes are pushed away from

the interface, a negative space charge region is created because of the fixed ionized acceptor atoms. The negative charge in the induced depletion region corresponds to the negative charge on the bottom “plate” of the MOS capacitor. Figure 10.3b shows the equilibrium distribution of charge in the MOS capacitor with this applied voltage.

The energy-band diagrams of the MOS capacitor with a p-type substrate for various gate biases are shown in Figure 10.4. Figure 10.4a shows the *ideal* case when zero bias is applied across the MOS device. The energy bands in the semiconductor are flat indicating no net charge exists in the semiconductor. This condition is known as *flat band* and is discussed in more detail later in the chapter.

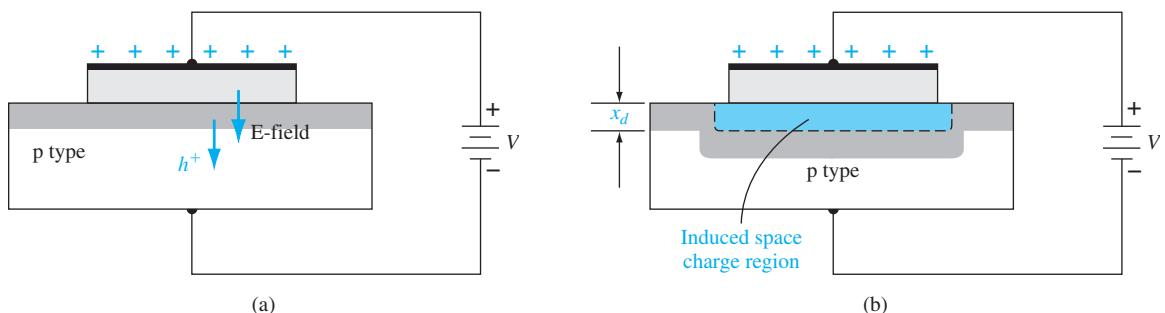


Figure 10.3 | The MOS capacitor with a moderate positive gate bias, showing (a) the electric field and charge flow and (b) the induced space charge region.

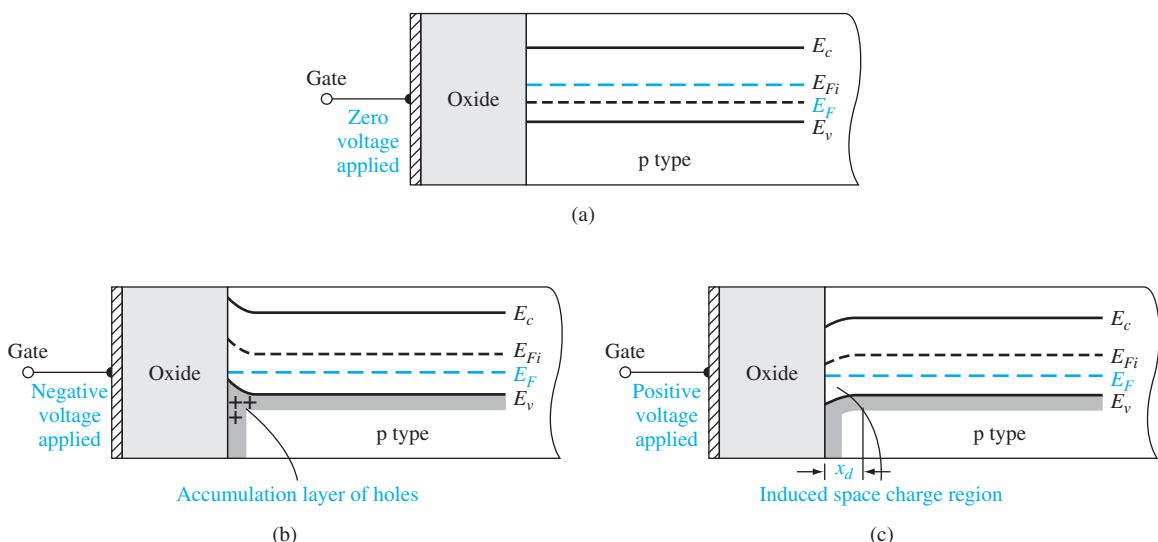


Figure 10.4 | The energy-band diagram of a MOS capacitor with a p-type substrate for (a) a zero applied gate bias showing the *ideal* case, (b) a negative gate bias, and (c) a moderate positive gate bias.

Figure 10.4b shows the energy-band diagram for the case when a negative bias is applied to the gate. (Remember that positive electron energy is plotted “upward” and positive voltage is plotted “downward.”) The valence-band edge is closer to the Fermi level at the oxide–semiconductor interface than in the bulk material, which implies that there is an accumulation of holes. The semiconductor surface appears to be more p-type than the bulk material. The Fermi level is a constant in the semiconductor since the MOS system is in thermal equilibrium and there is no current through the oxide.

Figure 10.4c shows the energy-band diagram of the MOS system when a positive voltage is applied to the gate. The conduction- and valence-band edges bend as shown in the figure, indicating a space charge region similar to that in a pn junction. The conduction band and intrinsic Fermi levels move closer to the Fermi level. The induced space charge width is x_d .

Now consider the case when a still larger positive voltage is applied to the top metal gate of the MOS capacitor. We expect the induced electric field to increase in magnitude and the corresponding positive and negative charges on the MOS capacitor to increase. A larger negative charge in the MOS capacitor implies a larger induced space charge region and more band bending. Figure 10.5 shows such a condition. The intrinsic Fermi level at the surface is now below the Fermi level. The conduction band at the surface is now close to the Fermi level, whereas the valence band is close to the Fermi level in the bulk semiconductor. This result implies that the surface in the semiconductor adjacent to the oxide–semiconductor interface is n type. By applying a sufficiently large positive gate voltage, we have inverted the surface of the semiconductor from a p-type to an n-type semiconductor. We have created an *inversion layer* of electrons at the oxide–semiconductor interface.

In the MOS capacitor structure that we have just considered, we assumed a p-type semiconductor substrate. The same type of energy-band diagrams can be constructed for a MOS capacitor with an n-type semiconductor substrate. Figure 10.6a shows the MOS capacitor structure with a positive voltage applied to the top gate terminal. A positive charge exists on the top gate and an electric field is induced with the direction shown in the figure. An accumulation layer of electrons will be induced in the n-type substrate. The case when a negative voltage is applied to the top gate

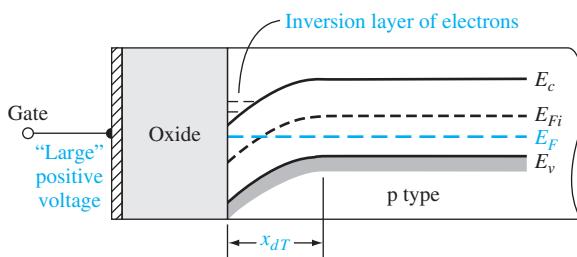


Figure 10.5 | The energy-band diagram of the MOS capacitor with a p-type substrate for a “large” positive gate bias.

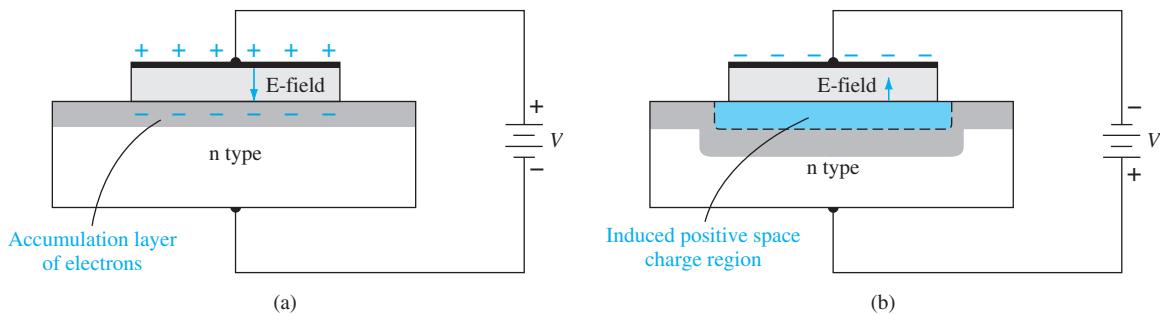


Figure 10.6 | The MOS capacitor with an n-type substrate for (a) a positive gate bias and (b) a moderate negative gate bias.

is shown in Figure 10.6b. A positive space charge region is induced in the n-type semiconductor in this situation.

The energy-band diagrams for this MOS capacitor with the n-type substrate are shown in Figure 10.7. Figure 10.7a shows the case when a positive voltage is applied to the gate and an accumulation layer of electrons is formed. Figure 10.7b shows the energy bands when a negative voltage is applied to the gate. The conduction and valence bands now bend upward indicating that a space charge region has been induced in the n-type substrate. Figure 10.7c shows the energy bands when a larger negative voltage is applied to the gate. The conduction and valence bands are bent even more and the intrinsic Fermi level has moved above the Fermi level. The valence band at the surface is now close to the Fermi level, whereas the conduction band is close to the Fermi level in the bulk semiconductor. This result implies that the semiconductor surface adjacent to the oxide–semiconductor interface is p type. By applying a sufficiently large negative voltage to the gate of the MOS capacitor, the semiconductor surface has been inverted from n type to p type. An inversion layer of holes has been induced at the oxide–semiconductor interface.

10.1.2 Depletion Layer Thickness

We may calculate the width of the induced space charge region adjacent to the oxide–semiconductor interface. Figure 10.8 shows the space charge region in a p-type semiconductor substrate. The potential ϕ_{fp} is the difference (in V) between E_{Fi} and E_F and is given by

$$\phi_{fp} = V_t \ln \left(\frac{N_a}{n_i} \right) \quad (10.4)$$

where N_a is the acceptor doping concentration and n_i is the intrinsic carrier concentration.

The potential ϕ_s is called the surface potential; it is the difference (in V) between E_{Fi} measured in the bulk semiconductor and E_{Fi} measured at the surface. The surface

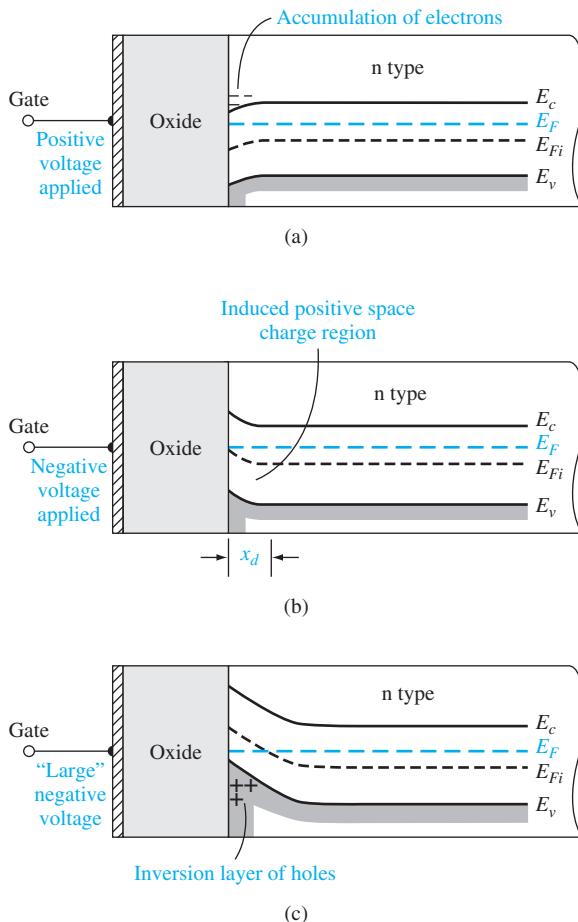


Figure 10.7 | The energy-band diagram of the MOS capacitor with an n-type substrate for (a) a positive gate bias, (b) a moderate negative bias, and (c) a “large” negative gate bias.

potential is the potential difference across the space charge layer. The space charge width can now be written in a form similar to that of a one-sided pn junction. We can write that

$$x_d = \left(\frac{2\epsilon_s \phi_s}{eN_a} \right)^{1/2} \quad (10.5)$$

where ϵ_s is the permittivity of the semiconductor. Equation (10.5) assumes that the abrupt depletion approximation is valid.

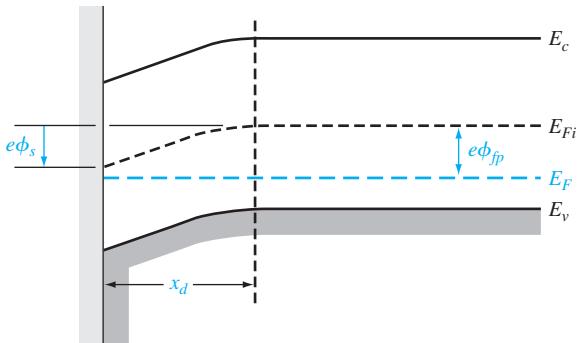


Figure 10.8 | The energy-band diagram in the p-type semiconductor, indicating surface potential.

Figure 10.9 shows the energy bands for the case in which $\phi_s = 2\phi_{fp}$. The Fermi level at the surface is as far above the intrinsic level as the Fermi level is below the intrinsic level in the bulk semiconductor. The electron concentration at the surface is the same as the hole concentration in the bulk material. This condition is known as the *threshold inversion point*. The applied gate voltage creating this condition is known as the *threshold voltage*. If the gate voltage increases above this threshold value, the conduction band will bend slightly closer to the Fermi level, but the change in the conduction band at the surface is now only a slight function of gate voltage. The electron concentration at the surface, however, is an exponential function of the surface potential. The surface potential may increase by a few (kT/e) volts, which will change the electron concentration by orders of magnitude, but the space charge width changes only slightly. In this case, then, the space charge region has essentially reached a maximum width.

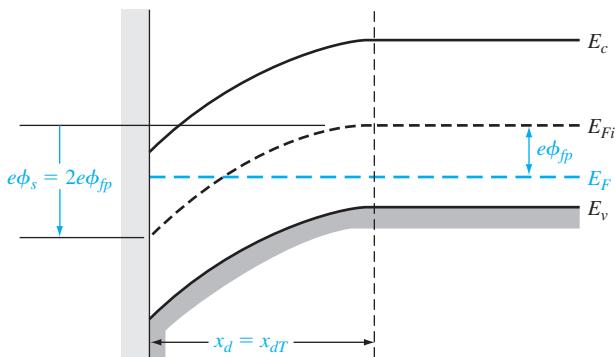


Figure 10.9 | The energy-band diagram in the p-type semiconductor at the threshold inversion point.

The maximum space charge width, x_{dT} , at this inversion transition point can be calculated from Equation (10.5) by setting $\phi_s = 2\phi_{fp}$. Then

$$x_{dT} = \left(\frac{4\epsilon_s \phi_{fp}}{eN_a} \right)^{1/2} \quad (10.6)$$

Objective: Calculate the maximum space charge width for a given semiconductor doping concentration.

EXAMPLE 10.1

Consider silicon at $T = 300$ K doped to $N_a = 10^{16} \text{ cm}^{-3}$. The intrinsic carrier concentration is $n_i = 1.5 \times 10^{10} \text{ cm}^{-3}$.

Solution

From Equation (10.4), we have

$$\phi_{fp} = V_t \ln \left(\frac{N_a}{n_i} \right) = (0.0259) \ln \left(\frac{10^{16}}{1.5 \times 10^{10}} \right) = 0.3473 \text{ V}$$

Then the maximum space charge width is

$$x_{dT} = \left[\frac{4\epsilon_s \phi_{fp}}{eN_a} \right]^{1/2} = \left[\frac{4(11.7)(8.85 \times 10^{-14})(0.3473)}{(1.6 \times 10^{-19})(10^{16})} \right]^{1/2}$$

or

$$x_{dT} \cong 0.30 \times 10^{-4} \text{ cm} = 0.30 \mu\text{m}$$

Comment

The maximum induced space charge width is on the same order of magnitude as pn junction space charge widths.

EXERCISE PROBLEM

Ex 10.1 Consider an oxide-to-p-type silicon junction at $T = 300$ K. The impurity doping concentration in the silicon is $N_a = 2 \times 10^{15} \text{ cm}^{-3}$. Calculate the maximum space charge width. Does the space charge width increase or decrease as the p-type doping concentration decreases?

(Ans. $x_{dT} = 0.629 \mu\text{m}$, increase)

We have been considering a p-type semiconductor substrate. The same maximum induced space charge region width occurs in an n-type substrate. Figure 10.10 is the energy-band diagram at the threshold voltage with an n-type substrate. We can write

$$\phi_{fn} = V_t \ln \left(\frac{N_d}{n_i} \right) \quad (10.7)$$

and

$$x_{dT} = \left(\frac{4\epsilon_s \phi_{fn}}{eN_d} \right)^{1/2} \quad (10.8)$$

Note that we are always assuming the parameters ϕ_{fp} and ϕ_{fn} to be positive quantities.

Figure 10.11 is a plot of x_{dT} at $T = 300$ K as a function of doping concentration in silicon. The semiconductor doping can be either n type or p type.

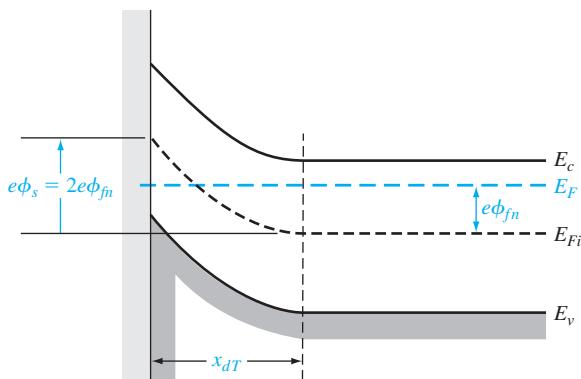


Figure 10.10 | The energy-band diagram in the n-type semiconductor at the threshold inversion point.

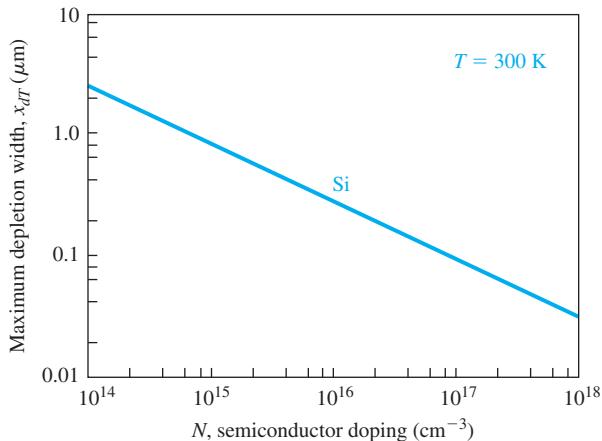


Figure 10.11 | Maximum induced space charge region width versus semiconductor doping.

10.1.3 Surface Charge Density

From the results of Chapter 4, the electron concentration in the conduction band can be written in the form

$$n = n_i \exp \left[\frac{E_F - E_{Fi}}{kT} \right] \quad (10.9)$$

For a p-type semiconductor substrate, the electron inversion charge density can then be written as (see Figure 10.9)

$$n_s = n_i \exp \left[\frac{e(\phi_{fp} + \Delta\phi_s)}{kT} \right] = n_i \exp \left[\frac{\phi_{fp} + \Delta\phi_s}{V_t} \right] \quad (10.10a)$$

or

$$n_s = n_i \exp\left(\frac{\phi_{fp}}{V_t}\right) \cdot \exp\left(\frac{\Delta\phi_s}{V_t}\right) \quad (10.10b)$$

where $\Delta\phi_s$, is the surface potential greater than $2\phi_{fp}$.

We may note that

$$n_{st} = n_i \exp\left(\frac{\phi_{fp}}{V_t}\right) \quad (10.11)$$

where n_{st} is the surface charge density at the threshold inversion point. The electron inversion charge density can then be written as

$$n_s = n_{st} \exp\left(\frac{\Delta\phi_s}{V_t}\right) \quad (10.12)$$

Figure 10.12 shows the electron inversion charge density as a function of surface potential for the case when the threshold inversion charge density is $n_{st} = 10^{16} \text{ cm}^{-3}$. We may note that the inversion charge density increases by a factor of 10 with a 60-mV increase in surface potential. As discussed previously, the electron inversion charge density increases rapidly with small increases in surface potential, which means that the space charge width essentially reaches a maximum value.

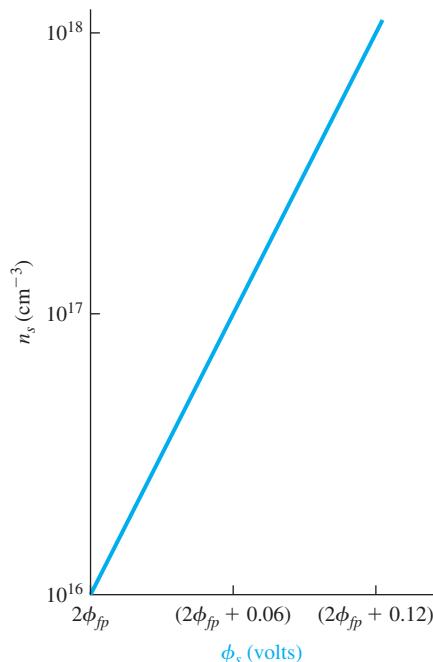


Figure 10.12 | Electron inversion charge density as a function of surface potential.

10.1.4 Work Function Differences

We have been concerned, so far, with the energy-band diagrams of the semiconductor material. Figure 10.13a shows the energy levels in the metal, silicon dioxide (SiO_2), and silicon relative to the vacuum level. The metal work function is ϕ_m and the electron affinity is χ . The parameter χ_i is the oxide electron affinity and, for SiO_2 , $\chi_i = 0.9 \text{ V}$.

Figure 10.13b shows the energy-band diagram of the entire MOS structure with zero gate voltage applied. The Fermi level is a constant through the entire system at thermal equilibrium. We may define ϕ'_m as a modified metal work function—the potential required to inject an electron from the metal into the conduction band of the oxide. Similarly, χ' is defined as a modified electron affinity. The voltage $V_{\text{ox}0}$ is the potential drop across the oxide for zero applied gate voltage and is not necessarily zero because of the difference between ϕ_m and χ . The potential ϕ_{s0} is the surface potential for this case.

If we sum the energies from the Fermi level on the metal side to the Fermi level on the semiconductor side, we have

$$e\phi'_m + eV_{\text{ox}0} = e\chi' + \frac{E_g}{2} - e\phi_{s0} + e\phi_{fp} \quad (10.13)$$

Equation (10.13) can be rewritten as

$$V_{\text{ox}0} + \phi_{s0} = -\left[\phi'_m - \left(\chi' + \frac{E_g}{2e} + \phi_{fp}\right)\right] \quad (10.14)$$

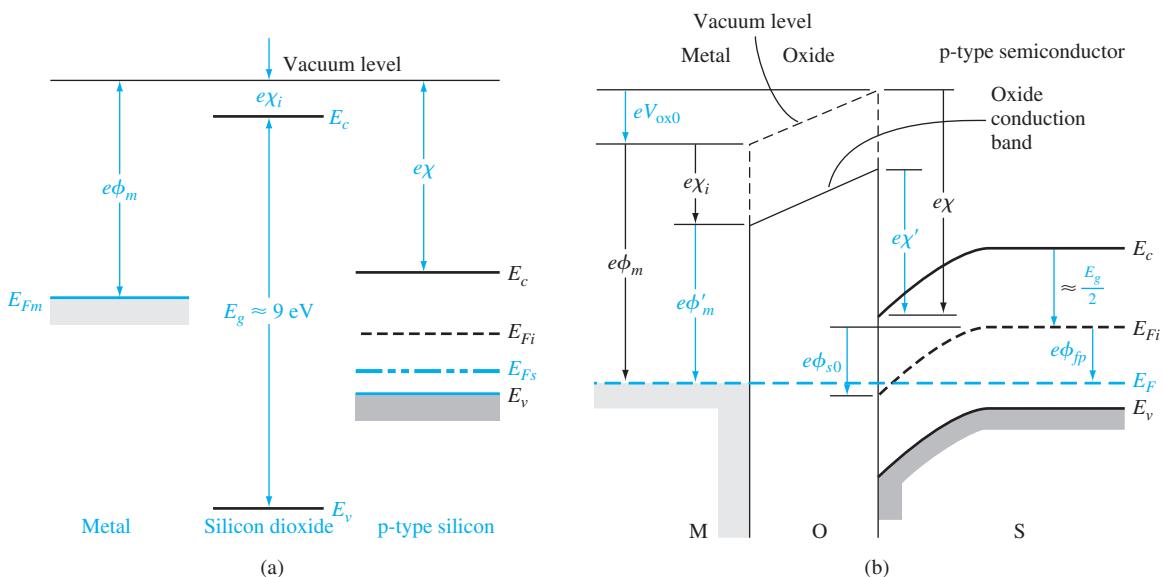


Figure 10.13 | (a) Energy levels in a MOS system prior to contact and (b) energy-band diagram through the MOS structure in thermal equilibrium after contact.

We can define a potential ϕ_{ms} as

$$\phi_{ms} \equiv \left[\phi'_m - \left(\chi' + \frac{E_g}{2e} + \phi_{fp} \right) \right] \quad (10.15)$$

which is known as the metal–semiconductor work function difference.

Objective: Determine the metal–semiconductor work function difference, ϕ_{ms} , for a given MOS system and semiconductor doping.

EXAMPLE 10.2

For an aluminum–silicon dioxide junction, $\phi'_m = 3.20$ V and, for a silicon–silicon dioxide junction, $\chi' = 3.25$ V. We may assume that $E_g = 1.12$ V. Let the p-type doping be $N_a = 10^{15} \text{ cm}^{-3}$.

Solution

For silicon at $T = 300$ K, we may calculate ϕ_{fp} as

$$\phi_{fp} = V_t \ln \left(\frac{N_a}{n_j} \right) = (0.0259) \ln \left(\frac{10^{15}}{1.5 \times 10^{10}} \right) = 0.288 \text{ V}$$

Then the metal–semiconductor work function difference is

$$\phi_{ms} = \phi'_m - \left(\chi' + \frac{E_g}{2e} + \phi_{fp} \right) = 3.20 - (3.25 + 0.560 + 0.288)$$

or

$$\phi_{ms} = -0.898 \text{ V}$$

Comment

The value of ϕ_{ms} will become more negative as the doping of the p-type substrate increases.

EXERCISE PROBLEM

Ex 10.2 Repeat Example 10.2 for a semiconductor doping of $N_a = 10^{16} \text{ cm}^{-3}$.

$$(\Delta \phi_{ms} = 0.056 \text{ V})$$

Degenerately doped polysilicon deposited on the oxide is also often used as the metal gate. Figure 10.14a shows the energy-band diagram of a MOS capacitor with an n⁺ polysilicon gate and a p-type substrate. Figure 10.14b shows the energy-band

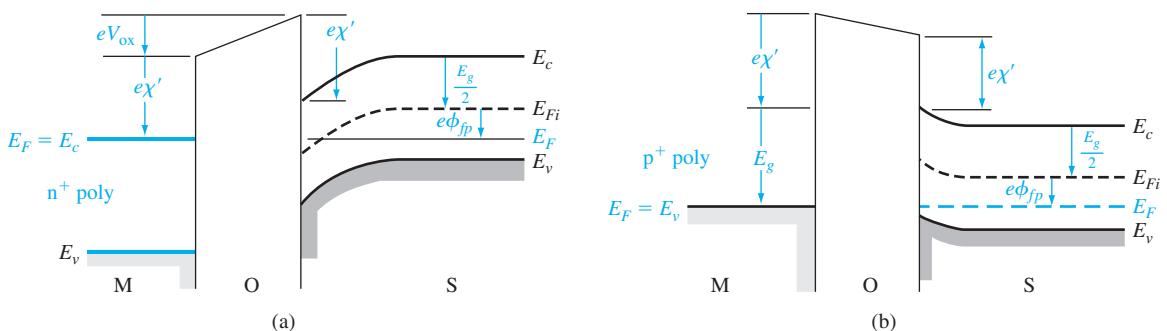


Figure 10.14 | Energy-band diagram through the MOS structure with a p-type substrate at zero gate bias for (a) an n⁺ polysilicon gate and (b) a p⁺ polysilicon gate.

diagram for the case of a p⁺ polysilicon gate and the p-type silicon substrate. In the degenerately doped polysilicon, we will initially assume that $E_F = E_c$ for the n⁺ case and $E_F = E_v$ for the p⁺ case.

For the n⁺ polysilicon gate, the metal–semiconductor work function difference can be written as

$$\phi_{ms} = \left[\chi' - \left(\chi' + \frac{E_g}{2e} + \phi_{fp} \right) \right] = - \left(\frac{E_g}{2e} + \phi_{fp} \right) \quad (10.16)$$

and for the p⁺ polysilicon gate, we have

$$\phi_{ms} = \left[\left(\chi' + \frac{E_g}{e} \right) - \left(\chi' + \frac{E_g}{2e} + \phi_{fp} \right) \right] = \left(\frac{E_g}{2e} - \phi_{fp} \right) \quad (10.17)$$

However, for degenerately doped n⁺ polysilicon and p⁺ polysilicon, the Fermi level can be above E_c and below E_v , respectively, by 0.1 to 0.2 V. The experimental ϕ_{ms} values will then be slightly different from the values calculated by using Equations (10.16) and (10.17).

We have been considering a p-type semiconductor substrate. We may also have an n-type semiconductor substrate in a MOS capacitor. Figure 10.15 shows the energy-band diagram of the MOS capacitor with a metal gate and the n-type semiconductor substrate, for the case when a negative voltage is applied to the gate. The metal–semiconductor work function difference for this case is defined as

$$\phi_{ms} = \phi'_m - \left(\chi' + \frac{E_g}{2e} - \phi_{fn} \right) \quad (10.18)$$

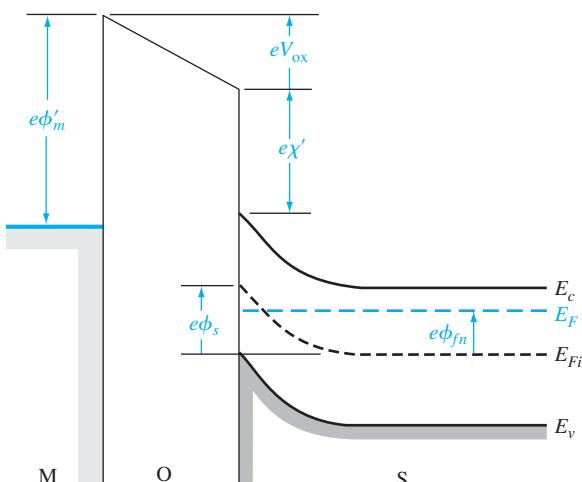


Figure 10.15 | Energy-band diagram through the MOS structure with an n-type substrate for a negative applied gate bias.

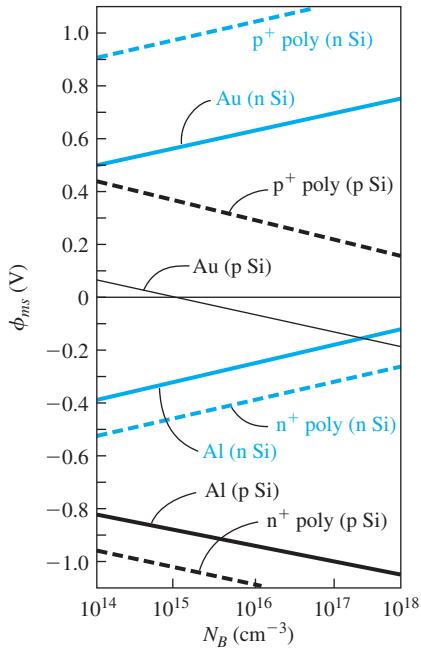


Figure 10.16 | Metal–semiconductor work function difference versus doping for aluminum, gold, and n⁺ and p[−] polysilicon gates.
(From Sze [17] and Werner [20].)

where ϕ_{fm} is assumed to be a positive value. We will have similar expressions for n⁺ and p⁺ polysilicon gates.

Figure 10.16 shows the work function differences as a function of semiconductor doping for the various types of gates. We may note that the magnitudes of ϕ_{ms} for the polysilicon gates are somewhat larger than Equations (10.16) and (10.17) predict. This difference again is because the Fermi level is not equal to the conduction-band energy for the n⁺ gate and is not equal to the valence-band energy for the p⁺ gate. The metal–semiconductor work function difference becomes important in the flat-band and threshold voltage parameters discussed next.

10.1.5 Flat-Band Voltage

The *flat-band voltage* is defined as the applied gate voltage such that there is no band bending in the semiconductor and, as a result, zero net space charge in this region. Figure 10.17 shows this flat-band condition. Because of the work function difference and possible trapped charge in the oxide, the voltage across the oxide for this case is not necessarily zero.

We have implicitly been assuming that there is zero net charge density in the oxide material. This assumption may not be valid—a net fixed charge density,

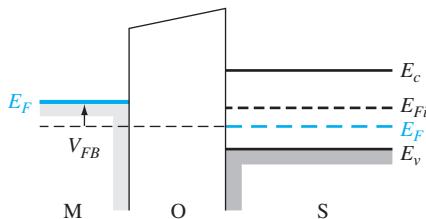


Figure 10.17 | Energy-band diagram of a MOS capacitor at flat band.

usually positive, may exist in the insulator. The positive charge has been identified with broken or dangling covalent bonds near the oxide–semiconductor interface. During the thermal formation of SiO_2 , oxygen diffuses through the oxide and reacts near the Si– SiO_2 interface to form the SiO_2 . Silicon atoms may also break away from the silicon material just prior to reacting to form SiO_2 . When the oxidation process is terminated, excess silicon may exist in the oxide near the interface, resulting in the dangling bonds. The magnitude of this oxide charge seems, in general, to be a strong function of the oxidizing conditions such as oxidizing ambient and temperature. The charge density can be altered to some degree by annealing the oxide in an argon or nitrogen atmosphere. However, the charge is rarely zero.

The net fixed charge in the oxide appears to be located fairly close to the oxide–semiconductor interface. We will assume in the analysis of the MOS structure that an equivalent trapped charge per unit area, Q'_{ss} , is located in the oxide directly adjacent to the oxide–semiconductor interface. For the moment, we will ignore any other oxide-type charges that may exist in the device. The parameter Q'_{ss} is usually given in terms of number of electronic charges per unit area.

Equation (10.14), for zero applied gate voltage, can be written as

$$V_{\text{ox}0} + \phi_{s0} = -\phi_{ms} \quad (10.19)$$

If a gate voltage is applied, the potential drop across the oxide and the surface potential will change. We can then write

$$V_G = \Delta V_{\text{ox}} + \Delta \phi_s = (V_{\text{ox}} - V_{\text{ox}0}) + (\phi_s - \phi_{s0}) \quad (10.20)$$

Using Equation (10.19), we have

$$V_G = V_{\text{ox}} + \phi_s + \phi_{ms} \quad (10.21)$$

Figure 10.18 shows the charge distribution in the MOS structure for the flat-band condition. There is zero net charge in the semiconductor, and we can assume that an equivalent fixed surface charge density exists in the oxide. The charge density on the metal is Q'_m , and from charge neutrality we have

$$Q'_m + Q'_{ss} = 0 \quad (10.22)$$

We can relate Q'_m to the voltage across the oxide by

$$V_{\text{ox}} = \frac{Q'_m}{C_{\text{ox}}} \quad (10.23)$$

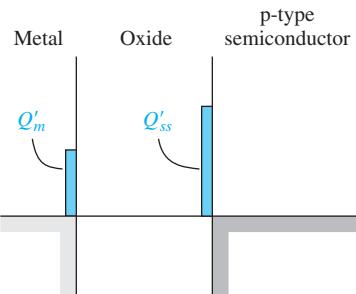


Figure 10.18 | Charge distribution in a MOS capacitor at flat band.

where C_{ox} is the oxide capacitance per unit area.¹ Substituting Equation (10.22) into Equation (10.23), we have

$$V_{\text{ox}} = \frac{-Q'_{ss}}{C_{\text{ox}}} \quad (10.24)$$

In the flat-band condition, the surface potential is zero, or $\phi_s = 0$. Then from Equation (10.21), we have

$$V_G = V_{FB} = \phi_{ms} - \frac{Q'_{ss}}{C_{\text{ox}}} \quad (10.25)$$

Equation (10.25) is the flat-band voltage for this MOS device.

Objective: Calculate the flat-band voltage for a MOS capacitor with a p-type semiconductor substrate.

EXAMPLE 10.3

Consider a MOS capacitor with a p-type silicon substrate doped to $N_a = 10^{16} \text{ cm}^{-3}$, a silicon dioxide insulator with a thickness of $t_{\text{ox}} = 20 \text{ nm} = 200 \text{ \AA}$, and an n⁺ polysilicon gate. Assume that $Q'_{ss} = 5 \times 10^{10}$ electronic charges per cm².

Solution

The work function difference, from Figure 10.16, is $\phi_{ms} \approx -1.1 \text{ V}$. The oxide capacitance is found to be

$$C_{\text{ox}} = \frac{\epsilon_{\text{ox}}}{t_{\text{ox}}} = \frac{(3.9)(8.85 \times 10^{-14})}{200 \times 10^{-8}} = 1.726 \times 10^{-7} \text{ F/cm}^2$$

The equivalent oxide charge density is

$$Q'_{ss} = (5 \times 10^{10})(1.6 \times 10^{-19}) = 8 \times 10^{-9} \text{ C/cm}^2$$

The flat-band voltage is then determined to be

$$V_{FB} = \phi_{ms} - \frac{Q'_{ss}}{C_{\text{ox}}} = -1.1 \frac{-8 \times 10^{-9}}{1.726 \times 10^{-7}} = -1.15 \text{ V}$$

¹Although we will, in general, use the primed notation for capacitance per unit area or charge per unit area, we will omit, for convenience, the prime on the oxide capacitance per unit area parameter.

Comment

The applied gate voltage required to achieve the flat-band condition for this p-type substrate is negative. If the amount of fixed oxide charge increases, the flat-band voltage becomes even more negative.

EXERCISE PROBLEM

- Ex 10.3** Repeat Example 10.3 for a doping concentration of $N_a = 2 \times 10^{15} \text{ cm}^{-3}$, an oxide thickness of $t_{ox} = 4 \text{ nm} = 40 \text{ \AA}$, and $Q'_{ss} = 2 \times 10^{10} \text{ electronic charges per cm}^2$. What is the value of the metal–semiconductor work function difference?

$$\text{Ans. } \phi_{ms} \equiv -1.03 \text{ V}, V_{Fe} = -1.034 \text{ V}$$

10.1.6 Threshold Voltage

The threshold voltage is defined as the applied gate voltage required to achieve the threshold inversion point. The threshold inversion point, in turn, is defined as the condition when the surface potential is $\phi_s = 2\phi_{fp}$ for the p-type semiconductor and $\phi_s = 2\phi_{fn}$ for the n-type semiconductor. These conditions are shown in Figures 10.9 and 10.10. The threshold voltage will be derived in terms of the electrical and geometrical properties of the MOS capacitor.

Figure 10.19 shows the charge distribution through the MOS device at the threshold inversion point for a p-type semiconductor substrate. The space charge width has reached its maximum value. We will assume that there is an equivalent oxide charge Q'_{ss} and the positive charge on the metal gate at threshold is Q'_{mT} . The prime on the charge terms indicates charge per unit area. Even though we are assuming that the surface has been inverted, we will neglect the inversion layer charge at this threshold inversion point. From conservation of charge, we can write

$$Q'_{mT} + Q'_{ss} = |Q'_{SD}(\max)| \quad (10.26)$$

where

$$|Q'_{SD}(\max)| = eN_a x_{dT} \quad (10.27)$$

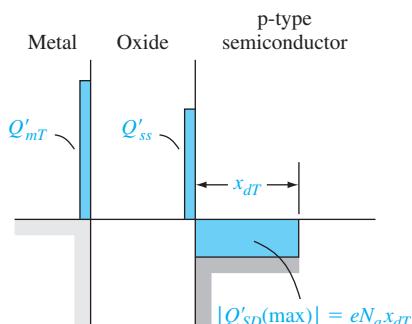


Figure 10.19 | Charge distribution in a MOS capacitor with a p-type substrate at the threshold inversion point.

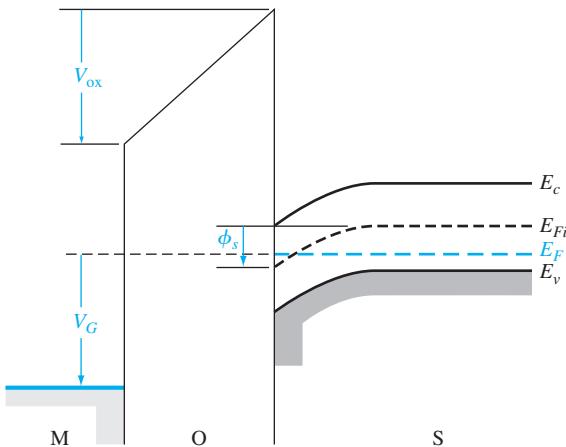


Figure 10.20 | Energy-band diagram through the MOS structure with a positive applied gate bias.

and is the magnitude of the maximum space charge density per unit area of the depletion region.

The energy-band diagram of the MOS system with an applied positive gate voltage is shown in Figure 10.20. As we mentioned, an applied gate voltage will change the voltage across the oxide and will change the surface potential. We had from Equation (10.20) that

$$V_G = \Delta V_{ox} + \Delta \phi_s = V_{ox} + \phi_s + \phi_{ms}$$

At threshold, we can define $V_G = V_{TN}$, where V_{TN} is the threshold voltage that creates the electron inversion layer charge. The surface potential is $\phi_s = 2\phi_{fp}$ at threshold, so Equation (10.20) can be written as

$$V_{TN} = V_{oxT} + 2\phi_{fp} + \phi_{ms} \quad (10.28)$$

where V_{oxT} is the voltage across the oxide at this threshold inversion point.

The voltage V_{oxT} can be related to the charge on the metal and to the oxide capacitance by

$$V_{oxT} = \frac{Q'_{mT}}{C_{ox}} \quad (10.29)$$

where again C_{ox} is the oxide capacitance per unit area. Using Equation (10.26), we can write

$$V_{oxT} = \frac{Q'_{mT}}{C_{ox}} = \frac{1}{C_{ox}} (|Q'_{SD}(\max)| - Q'_{ss}) \quad (10.30)$$

Finally, the threshold voltage can be written as

$$V_{TN} = \frac{|Q'_{SD}(\max)|}{C_{ox}} - \frac{Q'_{ss}}{C_{ox}} + \phi_{ms} + 2\phi_{fp} \quad (10.31a)$$

or

$$V_{TN} = (|Q'_{SD}(\max)| - Q'_{ss}) \left(\frac{t_{ox}}{\epsilon_{ox}} \right) + \phi_{ms} + 2\phi_{fp} \quad (10.31b)$$

Using the definition of flat-band voltage from Equation (10.25), we can also express the threshold voltage as

$$V_{TN} = \frac{|Q'_{SD}(\text{max})|}{C_{ox}} + V_{FB} + 2\phi_{fp} \quad (10.31c)$$

For a given semiconductor material, oxide material, and gate metal, the threshold voltage is a function of semiconductor doping, oxide charge Q'_{ss} , and oxide thickness.

EXAMPLE 10.4

Objective: Calculate the threshold voltage of a MOS system using an aluminum gate.

Consider a p-type silicon substrate at $T = 300$ K doped to $N_a = 10^{15} \text{ cm}^{-3}$. Let $Q'_{ss} = 10^{10} \text{ cm}^{-2}$, $t_{ox} = 12 \text{ nm} = 120 \text{ \AA}$, and assume the oxide is silicon dioxide.

■ Solution

From Figure 10.16, we find $\phi_{ms} \approx -0.88 \text{ V}$. The other parameters are

$$\phi_{fp} = V_t \ln \left(\frac{N_a}{n_i} \right) = (0.0259) \ln \left(\frac{10^{15}}{1.5 \times 10^{10}} \right) = 0.2877 \text{ V}$$

and

$$x_{dT} = \left\{ \frac{4\epsilon_s \phi_{fp}}{eN_a} \right\}^{1/2} = \left\{ \frac{4(11.7)(8.85 \times 10^{-14})(0.2877)}{(1.6 \times 10^{-19})(10^{15})} \right\}^{1/2} = 8.63 \times 10^{-5} \text{ cm}$$

Then

$$|Q'_{SD}(\text{max})| = eN_a x_{dT} = (1.6 \times 10^{-19})(10^{15})(8.63 \times 10^{-5}) = 1.381 \times 10^{-8} \text{ C/cm}^2$$

The threshold voltage is now found to be

$$\begin{aligned} V_{TN} &= (|Q'_{SD}(\text{max})| - Q'_{ss}) \left(\frac{t_{ox}}{\epsilon_{ox}} \right) + \phi_{ms} + 2\phi_{fp} \\ &= [(1.381 \times 10^{-8}) - (10^{10})(1.6 \times 10^{-19})] \cdot \left[\frac{120 \times 10^{-8}}{(3.9)(8.85 \times 10^{-14})} \right] \\ &\quad + (-0.88) + 2(0.2877) \end{aligned}$$

or

$$V_{TN} = -0.262 \text{ V}$$

■ Comment

In this example, the semiconductor is fairly lightly doped, which, in conjunction with the positive charge in the oxide and the work function difference, is sufficient to induce an electron inversion layer charge even with zero applied gate voltage. This condition makes the threshold voltage negative.

■ EXERCISE PROBLEM

Ex 10.4 Determine the metal–semiconductor work function difference and the threshold voltage for a silicon MOS device at $T = 300$ K for the following parameters:

p⁺ polysilicon gate, $N_a = 2 \times 10^{16} \text{ cm}^{-3}$, $t_{ox} = 8 \text{ nm} = 80 \text{ \AA}$, and $Q'_{ss} = 2 \times 10^{10} \text{ cm}^{-2}$.

(Ans. $\phi_{ms} = +0.28 \text{ V}$, $V_{TN} = +1.16 \text{ V}$)

A negative threshold voltage for a p-type substrate implies a depletion mode device. A negative voltage must be applied to the gate in order to make the inversion

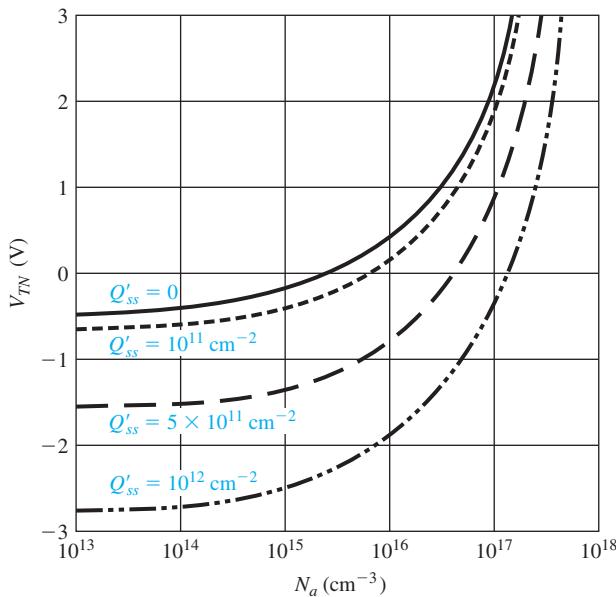


Figure 10.21 | Threshold voltage of an n-channel MOSFET versus the p-type substrate doping concentration for various values of oxide trapped charge ($t_{ox} = 500 \text{ \AA}$, aluminum gate).

layer charge equal to zero, whereas a positive gate voltage will induce a larger inversion layer charge.

Figure 10.21 is a plot of the threshold voltage V_{TN} as a function of the acceptor doping concentration for various positive oxide charge values. We may note that the p-type semiconductor must be somewhat heavily doped in order to obtain an enhancement mode device.

The previous derivation of the threshold voltage assumed a p-type semiconductor substrate. The same type of derivation can be done with an n-type semiconductor substrate, where a negative gate voltage can induce an inversion layer of holes at the oxide–semiconductor interface.

Figure 10.15 shows the energy-band diagram of the MOS structure with an n-type substrate and with an applied negative gate voltage. The threshold voltage for this case can be derived and is given by

$$V_{TP} = (-|Q'_{SD}(\max)| - Q'_{ss}) \left(\frac{t_{ox}}{\epsilon_{ox}} \right) + \phi_{ms} - 2\phi_{fn} \quad (10.32)$$

where

$$\phi_{ms} = \phi'_m - \left(\chi' + \frac{E_g}{2e} - \phi_{fn} \right) \quad (10.33a)$$

$$|Q'_{SD}(\max)| = eN_d x_{dT} \quad (10.33b)$$

$$x_{dT} = \left\{ \frac{4\epsilon_s \phi_{fn}}{eN_d} \right\}^{1/2} \quad (10.33c)$$

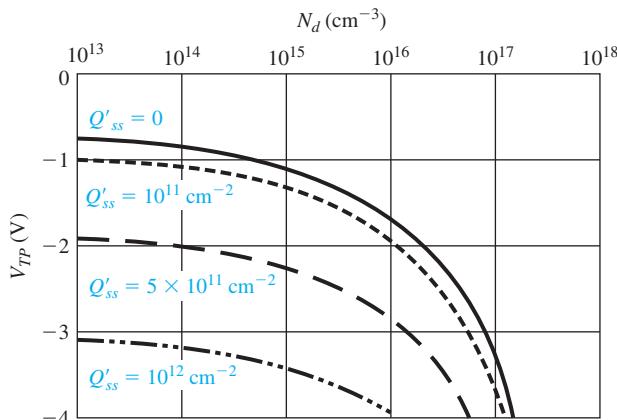


Figure 10.22 | Threshold voltage of a p-channel MOSFET versus the n-type substrate doping concentration for various values of oxide trapped charge ($t_{\text{ox}} = 500 \text{ \AA}$, aluminum gate).

and

$$\phi_{fn} = V_t \ln \left(\frac{N_d}{n_i} \right) \quad (10.33d)$$

We may note that x_{dT} and ϕ_{fn} are defined as positive quantities. We may also note that the notation of V_{TP} is the threshold voltage that will induce an inversion layer of holes. We will later drop the N and P subscript notation on the threshold voltage, but, for the moment, the notation may be useful for clarity.

Figure 10.22 is a plot of V_{TP} versus doping concentration for several values of Q'_{ss} . We may note that, for all values of positive oxide charge, this MOS capacitor is always an enhancement mode device. As the Q'_{ss} charge increases, the threshold voltage becomes more negative, which means that it takes a larger applied gate voltage to create the inversion layer of holes at the oxide–semiconductor interface.

DESIGN

EXAMPLE 10.5

Objective: Determine the gate material and design the semiconductor doping concentration to yield a specified threshold voltage.

Consider a MOS device with silicon dioxide and an n-type silicon substrate. The oxide thickness is $t_{\text{ox}} = 12 \text{ nm} = 120 \text{ \AA}$ and the oxide charge is $Q'_{ss} = 2 \times 10^{10} \text{ cm}^{-2}$. The threshold voltage is to be approximately $V_{TP} = -0.3 \text{ V}$.

■ Solution

The solution to this design problem is not straight forward, since the doping concentration parameter appears in the terms ϕ_{fn} , x_{dT} , Q'_{ss} (max), and ϕ_{ms} . The threshold voltage, then, is a nonlinear function of N_d . We resort to trial and error to obtain a solution.

Figure 10.22 shows the threshold voltage for an aluminum gate system. Since the required threshold voltage in this problem is less negative than the values shown in Figure 10.22,

we require a metal–semiconductor work function difference value that is more positive than for the aluminum gate. We therefore choose a p⁺ polysilicon gate.

Consider a doping concentration of $N_d = 10^{17} \text{ cm}^{-3}$. From Figure 10.16, the metal–semiconductor work function difference is $\phi_{ms} \approx +1.1 \text{ V}$. The remaining parameters are found to be

$$\phi_{fn} = V_t \ln \left(\frac{N_d}{n_i} \right) = (0.0259) \ln \left(\frac{10^{17}}{1.5 \times 10^{10}} \right) = 0.407 \text{ V}$$

$$x_{dt} = \left(\frac{4\epsilon_s \phi_{fn}}{eN_d} \right)^{1/2} = \left\{ \frac{4(11.7)(8.85 \times 10^{-14})(0.407)}{(1.6 \times 10^{-19})(10^{17})} \right\}^{1/2} \\ = 1.026 \times 10^{-5} \text{ cm}$$

$$|Q'_{SD}(\max)| = eN_d x_{dt} = (1.6 \times 10^{-19})(10^{17})(1.026 \times 10^{-5}) \\ = 1.642 \times 10^{-7} \text{ C/cm}^2$$

The threshold voltage is determined to be

$$V_{TP} = [-|Q'_{SD}(\max)| - Q'_{ss}] \cdot \left(\frac{t_{ox}}{\epsilon_{ox}} \right) + \phi_{ms} - 2\phi_{fn}$$

or

$$V_{TP} = \frac{[-(1.642 \times 10^{-7}) - (2 \times 10^{10})(1.6 \times 10^{-19})] \cdot (120 \times 10^{-8})}{(3.9)(8.85 \times 10^{-14})} \\ + 1.1 - 2(0.407)$$

which yields

$$V_{TP} = -0.296 \text{ V} \approx -0.3 \text{ V}$$

Comment

The negative threshold voltage, with the n-type substrate, implies that this MOS capacitor is an enhancement mode device. The inversion layer charge is zero with zero applied gate voltage, and a negative gate voltage must be applied to induce the hole inversion charge.

EXERCISE PROBLEM

Ex 10.5 Consider a MOS capacitor with silicon dioxide and an n-type silicon substrate at $T = 300 \text{ K}$ with the following parameters: p⁺ polysilicon gate, $N_d = 2 \times 10^{16} \text{ cm}^{-3}$, $t_{ox} = 20 \text{ nm} = 200 \text{ \AA}$, and $Q'_{ss} = 5 \times 10^{10} \text{ cm}^{-2}$. Determine the threshold voltage. Is the capacitor an enhancement mode or depletion mode device?

(Ans. $V_{TP} = -0.12 \text{ V}$, enhancement mode)

TEST YOUR UNDERSTANDING

TYU 10.1 (a) Consider an oxide-to-n-type silicon junction at $T = 300 \text{ K}$. The impurity doping concentration in the silicon is $N_d = 8 \times 10^{15} \text{ cm}^{-3}$. Calculate the maximum space charge width in the silicon. (b) Repeat part (a) for a doping concentration of $N_d = 4 \times 10^{16} \text{ cm}^{-3}$.

[Ans. (a) 0.158 μm; (b) 0.332 μm]

- TYU 10.2** Consider an n⁺ polysilicon gate in a MOS structure with a p-type silicon substrate. The doping concentration of the silicon is $N_a = 3 \times 10^{16} \text{ cm}^{-3}$. Using Equation (10.16), find the value of ϕ_{ms} .

$$(\Delta 9\text{E}6.0 - = \text{Ans.} \phi)$$

- TYU 10.3** Repeat TYU 10.2 for a p⁺ polysilicon gate using Equation (10.17).

$$(\Delta 9\text{E}6.0 + = \text{Ans.} \phi)$$

- TYU 10.4** Consider the MOS capacitor described in Exercise TYU 10.3. The oxide thickness is $t_{ox} = 16 \text{ nm} = 160 \text{ \AA}$ and the oxide charge density is $Q'_{ss} = 8 \times 10^{10} \text{ cm}^{-2}$. Determine the flat-band voltage.

$$(\Delta 9\text{E}6.0 + 0.125 \text{ V} = \text{Ans.} \phi)$$

- TYU 10.5** Consider an n⁺ polysilicon gate on silicon dioxide with a p-type silicon substrate doped to $N_a = 3 \times 10^{16} \text{ cm}^{-3}$. Assume $Q'_{ss} = 5 \times 10^{10} \text{ cm}^{-2}$. Determine the required oxide thickness such that the threshold voltage is $V_{TN} = +0.65 \text{ V}$.

$$(\Delta 9\text{E}6.0 = 45.2 \text{ nm} = 452 \text{ \AA} = \text{Ans.} t_{ox})$$

10.2 | CAPACITANCE–VOLTAGE CHARACTERISTICS

As mentioned previously, the MOS capacitor structure is the heart of the MOSFET. A great deal of information about the MOS device and the oxide–semiconductor interface can be obtained from the capacitance versus voltage or C–V characteristics of the device. The capacitance of a device is defined as

$$C = \frac{dQ}{dV} \quad (10.34)$$

where dQ is the magnitude of the differential change in charge on one plate as a function of the differential change in voltage dV across the capacitor. The capacitance is a small-signal or ac parameter and is measured by superimposing a small ac voltage on an applied dc gate voltage. The capacitance, then, is measured as a function of the applied dc gate voltage.

10.2.1 Ideal C–V Characteristics

First we will consider the ideal C–V characteristics of the MOS capacitor and then discuss some of the deviations that occur from these idealized results. We will initially assume that there is zero charge trapped in the oxide and also that there is no charge trapped at the oxide–semiconductor interface.

There are three operating conditions of interest in the MOS capacitor: accumulation, depletion, and inversion. Figure 10.23a shows the energy-band diagram of a MOS capacitor with a p-type substrate for the case when a negative voltage is applied to the gate, inducing an accumulation layer of holes in the semiconductor at the oxide–semiconductor interface. A small differential change in voltage across the MOS structure will cause a differential change in charge on the metal gate and also in the hole accumulation charge, as shown in Figure 10.23b. The differential changes in charge density occur at the edges of the oxide, as in a parallel-plate capacitor. The

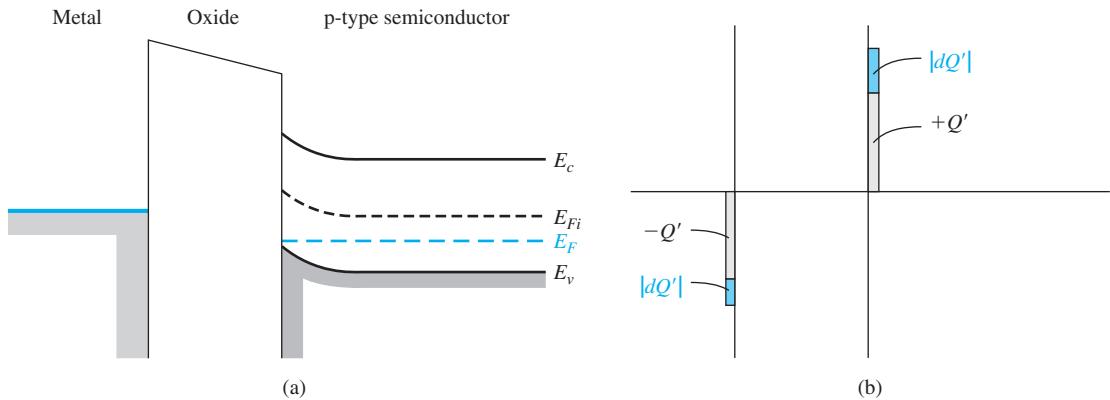


Figure 10.23 | (a) Energy-band diagram through a MOS capacitor for the accumulation mode. (b) Differential charge distribution at accumulation for a differential change in gate voltage.

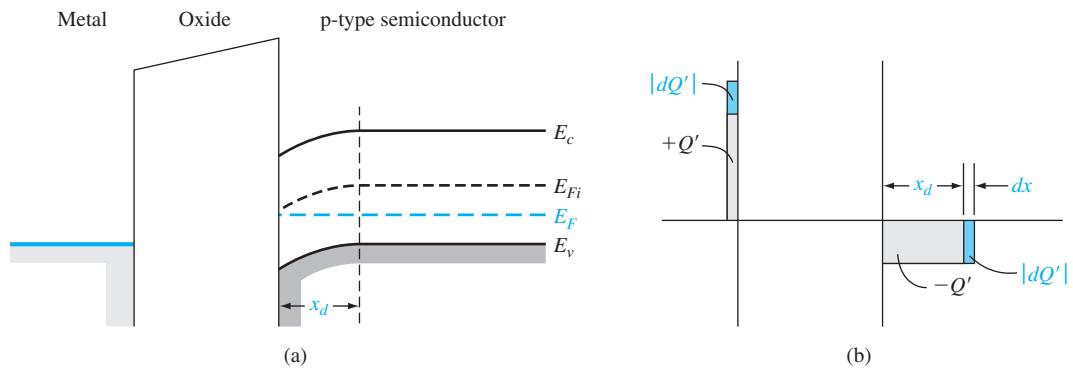


Figure 10.24 | (a) Energy-band diagram through a MOS device for the depletion mode. (b) Differential charge distribution at depletion for a differential change in gate voltage.

capacitance C' per unit area of the MOS capacitor for this accumulation mode is just the oxide capacitance, or

$$C'(\text{acc}) = C_{\text{ox}} = \frac{\epsilon_{\text{ox}}}{t_{\text{ox}}} \quad (10.35)$$

Figure 10.24a shows the energy-band diagram of the MOS device when a small positive voltage is applied to the gate, inducing a space charge region in the semiconductor; Figure 10.24b shows the charge distribution through the device for this condition. The oxide capacitance and the capacitance of the depletion region are in series. A small differential change in voltage across the capacitor will cause a differential change in the space charge width. The corresponding differential changes

in charge densities are shown in the figure. The total capacitance of the series combination is

$$\frac{1}{C'(\text{depl})} = \frac{1}{C_{\text{ox}}} + \frac{1}{C'_{SD}} \quad (10.36a)$$

or

$$C'(\text{depl}) = \frac{C_{\text{ox}} C'_{SD}}{C_{\text{ox}} + C'_{SD}} \quad (10.36b)$$

Since $C_{\text{ox}} = \epsilon_{\text{ox}}/t_{\text{ox}}$ and $C'_{SD} = \epsilon_s/x_d$, Equation (10.36b) can be written as

$$C'(\text{depl}) = \frac{C_{\text{ox}}}{1 + \frac{C_{\text{ox}}}{C'_{SD}}} = \frac{\epsilon_{\text{ox}}}{t_{\text{ox}} + \left(\frac{\epsilon_{\text{ox}}}{\epsilon_s}\right)x_d} \quad (10.37)$$

As the space charge width increases, the total capacitance $C'(\text{depl})$ decreases.

We had defined the threshold inversion point to be the condition when the maximum depletion width is reached, but there is essentially zero inversion charge density. This condition will yield a minimum capacitance C'_{\min} , which is given by

$$C'_{\min} = \frac{\epsilon_{\text{ox}}}{t_{\text{ox}} + \left(\frac{\epsilon_{\text{ox}}}{\epsilon_s}\right)x_{dT}} \quad (10.38)$$

Figure 10.25a shows the energy-band diagram of this MOS device for the inversion condition. In the ideal case, a small incremental change in the voltage across the MOS capacitor will cause a differential change in the inversion layer charge density. The space charge width does not change. If the inversion charge can respond to the change in capacitor voltage as indicated in Figure 10.25b, then the capacitance is again just the oxide capacitance, or

$$C'(\text{inv}) = C_{\text{ox}} = \frac{\epsilon_{\text{ox}}}{t_{\text{ox}}} \quad (10.39)$$

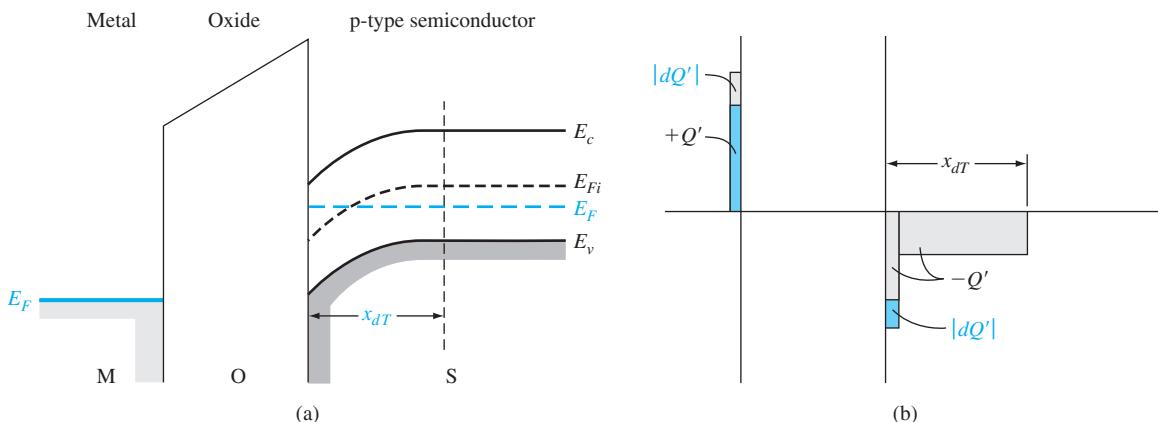


Figure 10.25 | (a) Energy-band diagram through a MOS capacitor for the inversion mode. (b) Differential charge distribution at inversion for a low-frequency differential change in gate voltage.

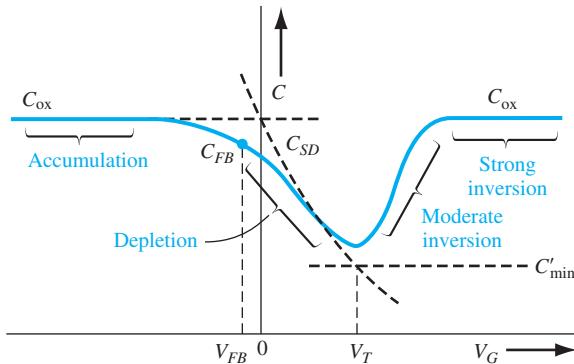


Figure 10.26 | Ideal low-frequency capacitance versus gate voltage of a MOS capacitor with a p-type substrate. Individual capacitance components are also shown.

Figure 10.26 shows the ideal capacitance versus gate voltage, or C - V , characteristics of the MOS capacitor with a p-type substrate. The three dashed segments correspond to the three components C_{ox} , C'_{\min} , and C_{SD} . The solid curve is the ideal net capacitance of the MOS capacitor. Moderate inversion, which is indicated in the figure, is the transition region between the point when only the space charge density changes with gate voltage and when only the inversion charge density changes with gate voltage.

The point on the curve that corresponds to the flat-band condition is of interest. The flat-band condition occurs between the accumulation and depletion conditions. The capacitance at flat band is given by

$$C'_{FB} = \frac{\epsilon_{ox}}{t_{ox} + \left(\frac{\epsilon_{ox}}{\epsilon_s}\right) \sqrt{\left(\frac{kT}{e}\right) \left(\frac{\epsilon_s}{eN_a}\right)}} \quad (10.40)$$

We may note that the flat-band capacitance is a function of oxide thickness as well as semiconductor doping. The general location of this point on the C - V plot is shown in Figure 10.26.

Objective: Calculate C_{ox} , C'_{\min} , and C'_{FB} for a MOS capacitor.

EXAMPLE 10.6

Consider a p-type silicon substrate at $T = 300$ K doped to $N_a = 10^{16}$ cm $^{-3}$.

The oxide is silicon dioxide with a thickness of $t_{ox} = 18$ nm = 180 Å, and the gate is aluminum.

■ Solution

The oxide capacitance is

$$C_{ox} = \frac{\epsilon_{ox}}{t_{ox}} = \frac{(3.9)(8.85 \times 10^{-14})}{180 \times 10^{-8}} = 1.9175 \times 10^{-7} \text{ F/cm}^2$$

To find the minimum capacitance, we need to calculate

$$\phi_{fp} = V_t \ln \left(\frac{N_a}{n_i} \right) = (0.0259) \ln \left(\frac{10^{16}}{1.5 \times 10^{10}} \right) = 0.3473 \text{ V}$$

and

$$x_{dT} = \left\{ \frac{4\epsilon_s \phi_{fp}}{eN_a} \right\}^{1/2} = \left\{ \frac{4(11.7)(8.85 \times 10^{-14})(0.3473)}{(1.6 \times 10^{-19})(10^{16})} \right\}^{1/2} \\ \cong 0.30 \times 10^{-4} \text{ cm}$$

Then

$$C'_{\min} = \frac{\epsilon_{ox}}{t_{ox} + \left(\frac{\epsilon_{ox}}{\epsilon_s}\right)x_{dT}} = \frac{(3.9)(8.85 \times 10^{-14})}{180 \times 10^{-8} + \left(\frac{3.9}{11.7}\right)(0.30 \times 10^{-4})} \\ = 2.925 \times 10^{-8} \text{ F/cm}^2$$

We may note that

$$\frac{C'_{\min}}{C_{ox}} = \frac{2.925 \times 10^{-8}}{1.9175 \times 10^{-7}} = 0.1525$$

The flat-band capacitance is

$$C'_{FB} = \frac{\epsilon_{ox}}{t_{ox} + \left(\frac{\epsilon_{ox}}{\epsilon_s}\right)\sqrt{\frac{V_f \epsilon_s}{eN_a}}} \\ = \frac{(3.9)(8.85 \times 10^{-14})}{180 \times 10^{-8} + \left(\frac{3.9}{11.7}\right)\sqrt{\frac{(0.0259)(11.7)(8.85 \times 10^{-14})}{(1.6 \times 10^{-19})(10^{16})}}} \\ = 1.091 \times 10^{-7} \text{ F/cm}^2$$

We also note that

$$\frac{C'_{FB}}{C_{ox}} = \frac{1.091 \times 10^{-7}}{1.9175 \times 10^{-7}} = 0.569$$

Comment

The ratios of C'_{\min}/C_{ox} and C'_{FB}/C_{ox} are typical values obtained in $C-V$ plots.

EXERCISE PROBLEM

Ex 10.6 Consider a MOS capacitor with the following parameters: n^+ polysilicon gate, $N_a = 3 \times 10^{16} \text{ cm}^{-3}$, $t_{ox} = 8 \text{ nm} = 80 \text{ \AA}$, and $Q_{ss} = 2 \times 10^{10} \text{ cm}^{-2}$. Determine the ratios C'_{\min}/C_{ox} and C'_{FB}/C_{ox} .

(Ans. $C'_{\min}/C_{ox} = 0.118$, $C'_{FB}/C_{ox} = 0.504$)

If we assume the oxide capacitance per unit area is $C_{ox} = 1.9175 \times 10^{-7} \text{ F/cm}^2$ and the channel length and width are $L = 2 \mu\text{m}$ and $W = 20 \mu\text{m}$, respectively, then the total gate oxide capacitance is

$$C_{oxT} = C_{ox} LW = (1.9175 \times 10^{-7})(2 \times 10^{-4})(20 \times 10^{-4}) \\ = 7.67 \times 10^{-14} \text{ F} = 0.0767 \text{ pF} = 76.7 \text{ fF}$$

The total oxide capacitance in a typical MOS device is quite small.

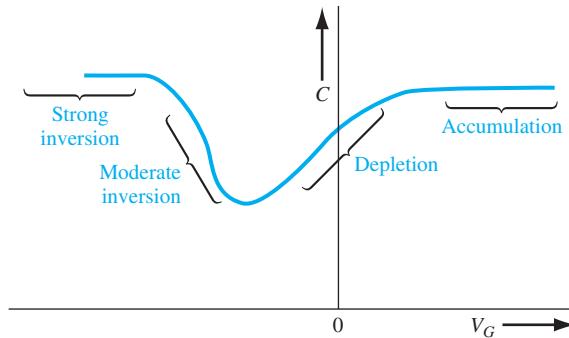


Figure 10.27 | Ideal low-frequency capacitance versus gate voltage of a MOS capacitor with an n-type substrate.

The same type of ideal C – V characteristics is obtained for a MOS capacitor with an n-type substrate by changing the sign of the voltage axis. The accumulation condition is obtained for a positive gate bias and the inversion condition is obtained for a negative gate bias. This ideal curve is shown in Figure 10.27.

10.2.2 Frequency Effects

Figure 10.25a shows the MOS capacitor with a p-type substrate and biased in the inversion condition. We have argued that a differential change in the capacitor voltage in the ideal case causes a differential change in the inversion layer charge density. However, we must consider the source of electrons that produces a change in the inversion charge density.

There are two sources of electrons that can change the charge density of the inversion layer. The first source is by diffusion of minority carrier electrons from the p-type substrate across the space charge region. This diffusion process is the same as that in a reverse-biased pn junction that generates the ideal reverse saturation current. The second source of electrons is by thermal generation of electron–hole pairs within the space charge region. This process is again the same as that in a reverse-biased pn junction generating the reverse-biased generation current. Both of these processes generate electrons at a particular rate. The electron concentration in the inversion layer, then, cannot change instantaneously. If the ac voltage across the MOS capacitor changes rapidly, the change in the inversion layer charge will not be able to respond. The C – V characteristics will then be a function of the frequency of the ac signal used to measure the capacitance.

In the limit of a very high frequency, the inversion layer charge will not respond to a differential change in capacitor voltage. Figure 10.28 shows the charge distribution in the MOS capacitor with a p-type substrate. At a high-signal frequency, the differential change in charge occurs at the metal and in the space charge width in the semiconductor. The capacitance of the MOS capacitor is then C'_{\min} , which we discussed earlier.

The high-frequency and low-frequency limits of the C – V characteristics are shown in Figure 10.29. In general, high frequency corresponds to a value on the order of 1 MHz and low frequency corresponds to values in the range of 5 to 100 Hz. Typically, the high-frequency characteristics of the MOS capacitor are measured.

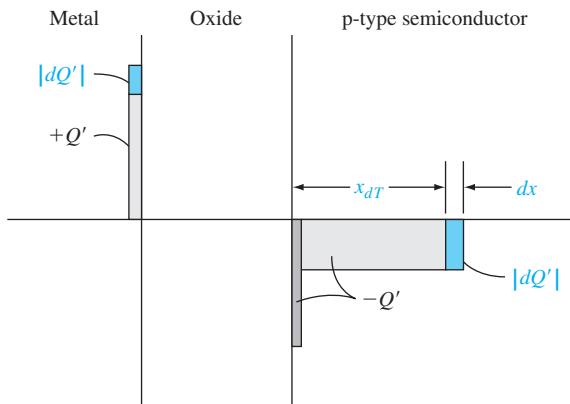


Figure 10.28 | Differential charge distribution at inversion for a high-frequency differential change in gate voltage.

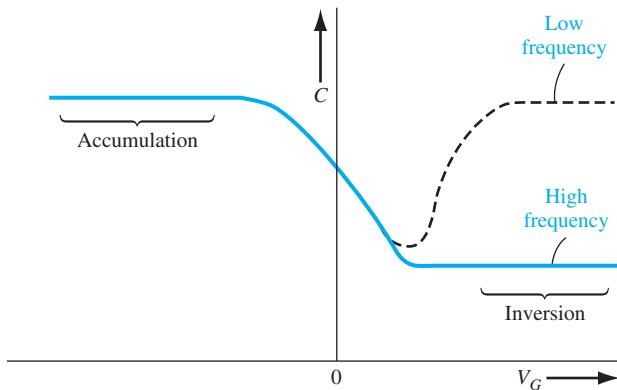


Figure 10.29 | Low-frequency and high-frequency capacitance versus gate voltage of a MOS capacitor with a p-type substrate.

10.2.3 Fixed Oxide and Interface Charge Effects

In all of the discussion concerning C - V characteristics so far, we have assumed an ideal oxide in which there are no fixed oxide or oxide–semiconductor interface charges. These two types of charges will change the C - V characteristics.

We previously discussed how the fixed oxide charge affects the threshold voltage. This charge will also affect the flat-band voltage. The flat-band voltage from Equation (10.25) is given by

$$V_{FB} = \phi_{ms} - \frac{Q'_{ss}}{C_{ox}}$$

where Q'_{ss} is the equivalent fixed oxide charge and ϕ_{ms} is the metal–semiconductor work function difference. The flat-band voltage shifts to more negative voltages for a positive fixed oxide charge. Since the oxide charge is not a function of gate voltage, the curves

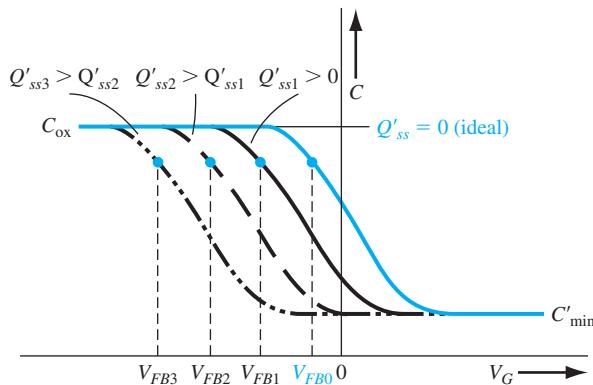


Figure 10.30 | High-frequency capacitance versus gate voltage of a MOS capacitor with a p-type substrate for several values of effective trapped oxide charge.

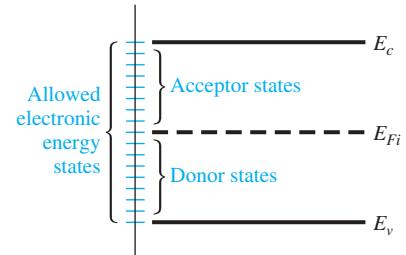


Figure 10.31 | Schematic diagram showing interface states at the oxide–semiconductor interface.

show a parallel shift with oxide charge, and the shape of the C – V curves remains the same as the ideal characteristics. Figure 10.30 shows the high-frequency characteristics of a MOS capacitor with a p-type substrate for several values of fixed positive oxide charge.

The C – V characteristics can be used to determine the equivalent fixed oxide charge. For a given MOS structure, ϕ_{ms} and C_{ox} are known, so the ideal flat-band voltage and flat-band capacitance can be calculated. The experimental value of flat-band voltage can be measured from the C – V curve, and the value of fixed oxide charge can then be determined. The C – V measurements are a valuable diagnostic tool to characterize a MOS device. This characterization is especially useful in the study of radiation effects on MOS devices, for example.

We first encountered oxide–semiconductor interface states in Chapter 9 in the discussion of Schottky barrier diodes. Figure 10.31 shows the energy-band diagram of a semiconductor at the oxide–semiconductor interface. The periodic nature of the semiconductor is abruptly terminated at the interface so that allowed electronic energy levels will exist within the forbidden bandgap. These allowed energy states are referred to as interface states. Charge can flow between the semiconductor and interface states, in contrast to the fixed oxide charge. The net charge in these interface states is a function of the position of the Fermi level in the bandgap.

In general, acceptor states exist in the upper half of the bandgap and donor states exist in the lower half of the bandgap. An acceptor state is neutral if the Fermi level is below the state and becomes negatively charged if the Fermi level is above the state. A donor state is neutral if the Fermi level is above the state and becomes positively charged if the Fermi level is below the state. The charge of the interface states is then a function of the gate voltage applied across the MOS capacitor.

Figure 10.32a shows the energy-band diagram in a p-type semiconductor of a MOS capacitor biased in the accumulation condition. In this case, there is a net positive charge trapped in the donor states. Now let the gate voltage change to produce the energy-band diagram shown in Figure 10.32b. The Fermi level corresponds to

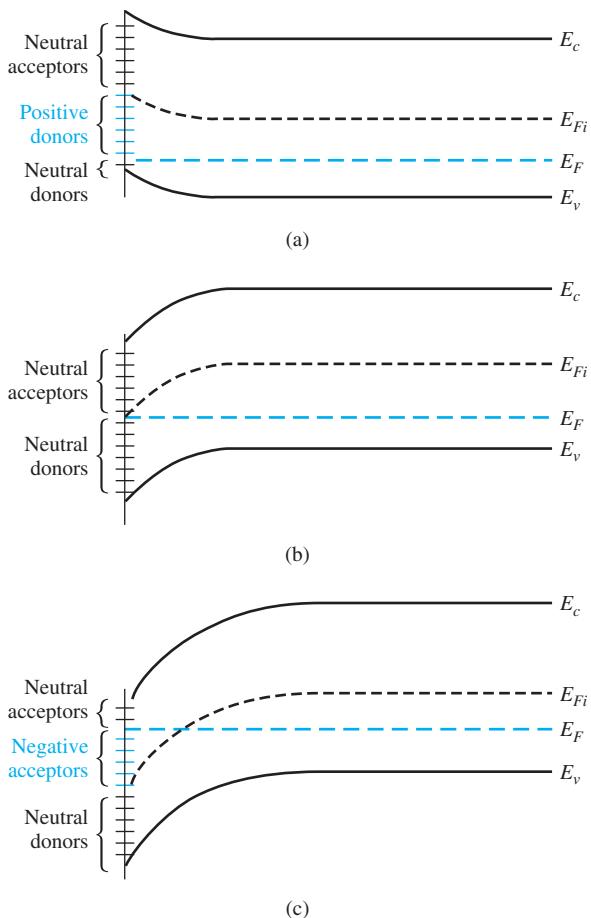


Figure 10.32 | Energy-band diagram in a p-type semiconductor showing the charge trapped in the interface states when the MOS capacitor is biased (a) in accumulation, (b) at midgap, and (c) at inversion.

the intrinsic Fermi level at the surface; thus, all interface states are neutral. This particular bias condition is known as *midgap*. Figure 10.32c shows the condition at inversion in which there is now a net negative charge in the acceptor states.

The net charge in the interface states changes from positive to negative as the gate voltage sweeps from the accumulation, depletion, to the inversion condition. We noted that the $C-V$ curves shifted in the negative gate voltage direction due to positive fixed oxide charge. When interface states are present, the amount and direction of the shift change as we sweep through the gate voltage, since the amount and sign of the interface trapped charge change. The $C-V$ curves now become “smeared out” as shown in Figure 10.33.

Again, the $C-V$ measurements can be used as a diagnostic tool in semiconductor device process control. For a given MOS device, the ideal $C-V$ curve can be determined.

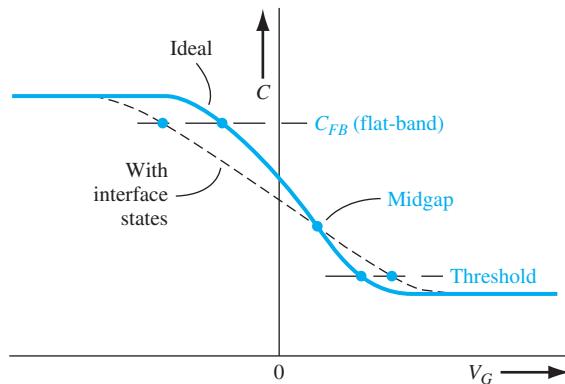


Figure 10.33 | High-frequency C–V characteristics of a MOS capacitor showing effects of interface states.

Any “smearing out” in the experimental curve indicates the presence of interface states and any parallel shift indicates the presence of fixed oxide charge. The amount of smearing out can be used to determine the density of interface states. These types of measurement are extremely useful in the study of radiation effects on MOS devices.

10.3 | THE BASIC MOSFET OPERATION

The current in a MOSFET is due to the flow of charge in the inversion layer or channel region adjacent to the oxide–semiconductor interface. We have discussed the creation of the inversion layer charge in enhancement-type MOS capacitors. We may also have depletion-type devices in which a channel already exists at zero gate voltage.

10.3.1 MOSFET Structures

There are four basic MOSFET device types. Figure 10.34 shows an n-channel enhancement mode MOSFET. Implicit in the enhancement mode notation is the idea that the semiconductor substrate is not inverted directly under the oxide with zero gate voltage. A positive gate voltage induces the electron inversion layer, which then “connects” the n-type source and the n-type drain regions. The source terminal is the source of carriers that flow through the channel to the drain terminal. For this n-channel device, electrons flow from the source to the drain so the conventional current will enter the drain and leave the source. The conventional circuit symbol for this n-channel enhancement mode device is also shown in the figure.

Figure 10.35 shows an n-channel depletion mode MOSFET. An n-channel region exists under the oxide with 0 V applied to the gate. However, we have shown that the threshold voltage of a MOS device with a p-type substrate may be negative; this means that an electron inversion layer already exists with zero gate voltage applied. Such a device is also considered to be a depletion mode device. The n-channel shown in the figure can be an electron inversion layer or an intentionally doped n region. The conventional circuit symbol for the n-channel depletion mode MOSFET is also shown in the figure.

Figures 10.36a, b show a p-channel enhancement mode MOSFET and a p-channel depletion mode MOSFET. In the p-channel enhancement mode device, a negative

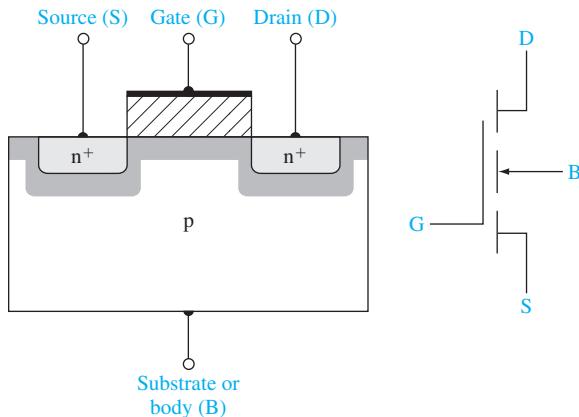


Figure 10.34 | Cross section and circuit symbol for an n-channel enhancement mode MOSFET.

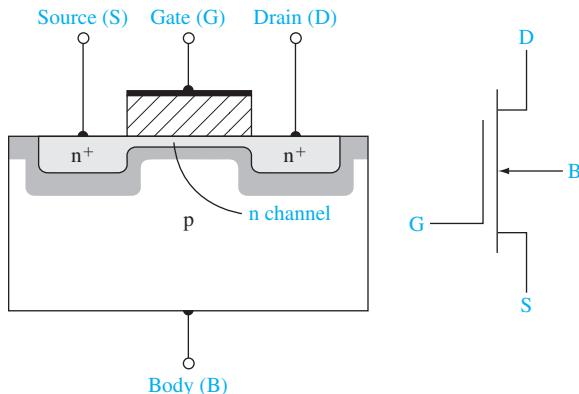


Figure 10.35 | Cross section and circuit symbol for an n-channel depletion mode MOSFET.

gate voltage must be applied to create an inversion layer of holes that will “connect” the p-type source and drain regions. Holes flow from the source to the drain, so the conventional current will enter the source and leave the drain. A p-channel region exists in the depletion mode device even with zero gate voltage. The conventional circuit symbols are shown in the figure.

10.3.2 Current–Voltage Relationship—Concepts

Figure 10.37a shows an n-channel enhancement mode MOSFET with a gate-to-source voltage that is less than the threshold voltage and with only a very small drain-to-source voltage. The source and substrate, or body, terminals are held at ground potential. With this bias configuration, there is no electron inversion layer, the drain-to-substrate pn junction is reverse biased, and the drain current is zero (disregarding pn junction leakage currents).

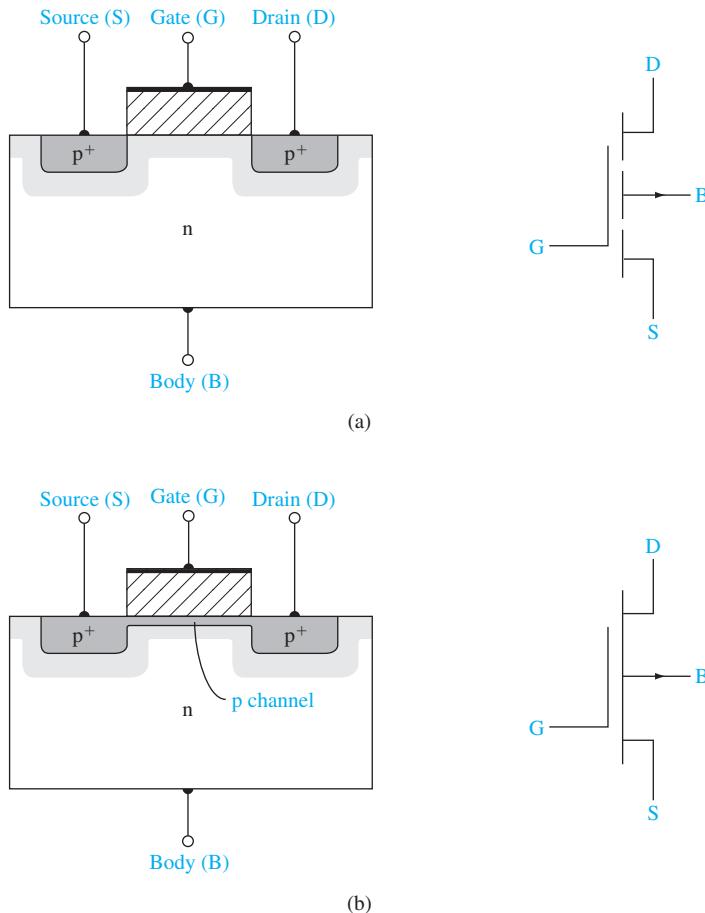


Figure 10.36 | Cross section and circuit symbol for (a) a p-channel enhancement mode MOSFET and (b) a p-channel depletion mode MOSFET.

Figure 10.37b shows the same MOSFET with an applied gate voltage such that $V_{GS} > V_T$. An electron inversion layer has been created so that when a small drain voltage is applied, the electrons in the inversion layer will flow from the source to the positive drain terminal. The conventional current enters the drain terminal and leaves the source terminal. In this ideal case, there is no current through the oxide to the gate terminal.

For small V_{DS} values, the channel region has the characteristics of a resistor, so we can write

$$I_D = g_d V_{DS} \quad (10.41)$$

where g_d is defined as the channel conductance in the limit as $V_{DS} \rightarrow 0$. The channel conductance is given by

$$g_d = \frac{W}{L} \cdot \mu_n |Q'_n| \quad (10.42)$$

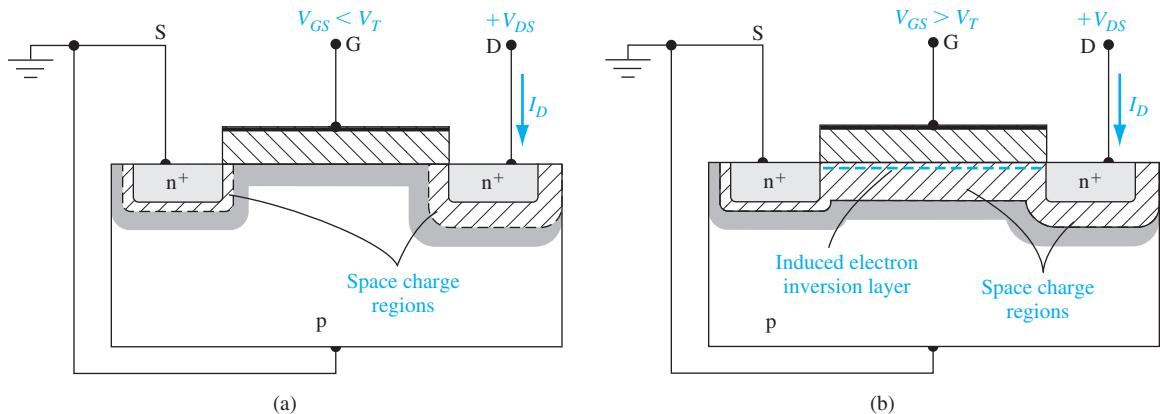


Figure 10.37 | The n-channel enhancement mode MOSFET (a) with an applied gate voltage $V_{GS} < V_T$ and (b) with an applied gate voltage $V_{GS} > V_T$.

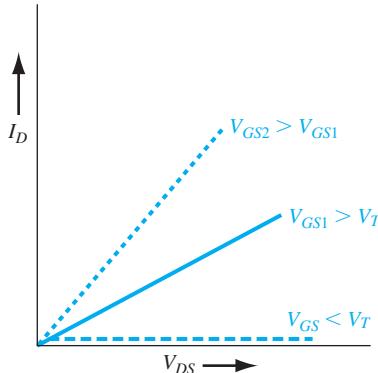


Figure 10.38 | I_D versus V_{DS} characteristics for small values of V_{DS} at three V_{GS} voltages.

where μ_n is the mobility of the electrons in the inversion layer and $|Q'_n|$ is the magnitude of the inversion layer charge per unit area. The inversion layer charge is a function of the gate voltage; thus, the basic MOS transistor action is the modulation of the channel conductance by the gate voltage. The channel conductance, in turn, determines the drain current. We will initially assume that the mobility is a constant; we will discuss mobility effects and variations in the next chapter.

The I_D versus V_{DS} characteristics, for small values of V_{DS} , are shown in Figure 10.38. When $V_{GS} < V_T$, the drain current is zero. As V_{GS} becomes larger than V_T , channel inversion charge density increases, which increases the channel conductance. A larger value of g_d produces a larger initial slope of the I_D versus V_{DS} characteristic as shown in the figure.

Figure 10.39a shows the basic MOS structure for the case when $V_{GS} > V_T$ and the applied V_{DS} voltage is small. The thickness of the inversion channel layer in the

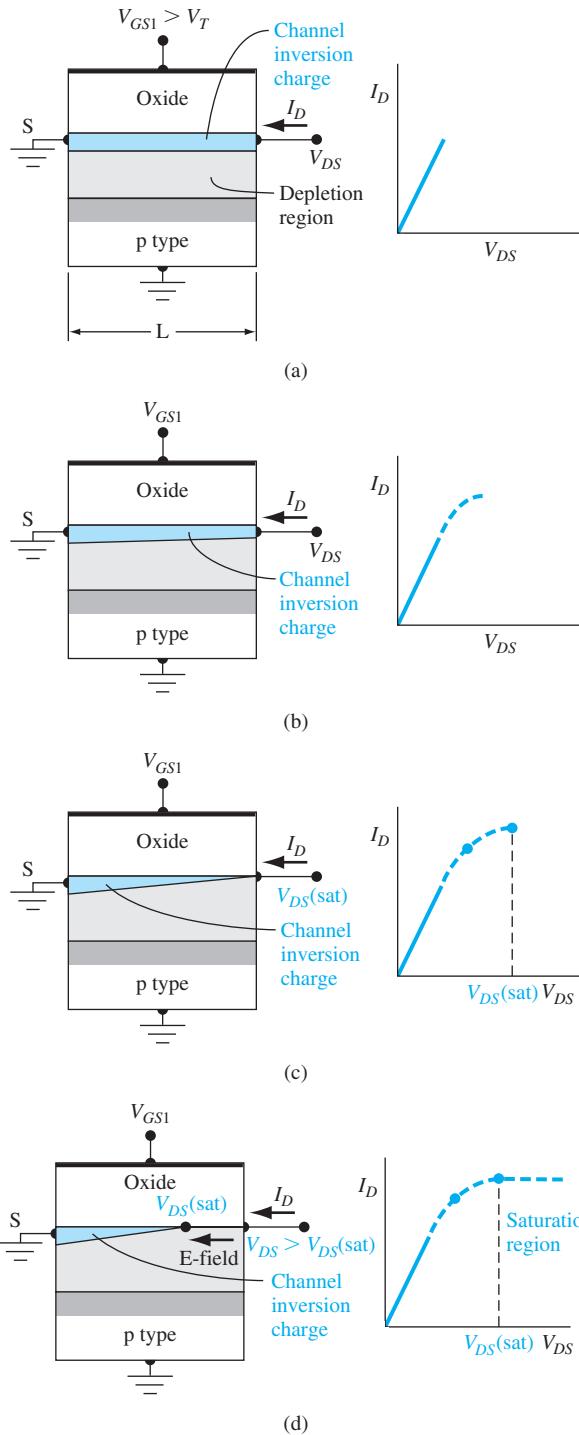


Figure 10.39 | Cross section and I_D versus V_{DS} curve when $V_{GS} < V_T$ for (a) a small V_{DS} value, (b) a larger V_{DS} value, (c) a value of $V_{DS} = V_{DS(\text{sat})}$, and (d) a value of $V_{DS} > V_{DS(\text{sat})}$.

figure qualitatively indicates the relative charge density, which is essentially constant along the entire channel length for this case. The corresponding I_D versus V_{DS} curve is shown in the figure.

Figure 10.39b shows the situation when the V_{DS} value increases. As the drain voltage increases, the voltage drop across the oxide near the drain terminal decreases, which means that the induced inversion charge density near the drain also decreases. The incremental conductance of the channel at the drain decreases, which then means that the slope of the I_D versus V_{DS} curve will decrease. This effect is shown in the I_D versus V_{DS} curve in the figure.

When V_{DS} increases to the point where the potential drop across the oxide at the drain terminal is equal to V_T , the induced inversion charge density is zero at the drain terminal. This effect is schematically shown in Figure 10.39c. At this point, the incremental conductance at the drain is zero, which means that the slope of the I_D versus V_{DS} curve is zero. We can write

$$V_{GS} - V_{DS}(\text{sat}) = V_T \quad (10.43a)$$

or

$$V_{DS}(\text{sat}) = V_{GS} - V_T \quad (10.43b)$$

where $V_{DS}(\text{sat})$ is the drain-to-source voltage producing zero inversion charge density at the drain terminal.

When V_{DS} becomes larger than the $V_{DS}(\text{sat})$ value, the point in the channel at which the inversion charge is just zero moves toward the source terminal. In this case, electrons enter the channel at the source, travel through the channel toward the drain, and then, at the point where the charge goes to zero, the electrons are injected into the space charge region where they are swept by the E-field to the drain contact. If we assume that the change in channel length ΔL is small compared to the original length L , then the drain current will be a constant for $V_{DS} > V_{DS}(\text{sat})$. The region of the I_D versus V_{DS} characteristic is referred to as the *saturation region*. Figure 10.39d shows this region of operation.

When V_{GS} changes, the I_D versus V_{DS} curve will change. We saw that, if V_{GS} increases, the initial slope of I_D versus V_{DS} increases. We can also note from Equation (10.43b) that the value of $V_{DS}(\text{sat})$ is a function of V_{GS} . We can generate the family of curves for this n-channel enhancement mode MOSFET as shown in Figure 10.40.

Figure 10.41 shows an n-channel depletion mode MOSFET. If the n-channel region is actually an induced electron inversion layer created by the metal–semiconductor work function difference and fixed charge in the oxide, the current–voltage characteristics are exactly the same as we have discussed, except that V_T is a negative quantity. We may also consider the case when the n-channel region is actually an n-type semiconductor region. In this type of device, a negative gate voltage will induce a space charge region under the oxide, reducing the thickness of the n-channel region. The reduced thickness decreases the channel conductance, which reduces the drain current. A positive gate voltage will create an electron accumulation layer, which increases the drain current. One basic requirement for this device is that the channel

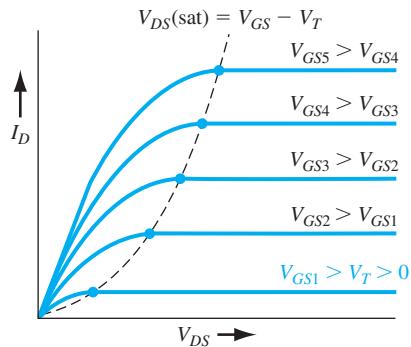


Figure 10.40 | Family of I_D versus V_{DS} curves for an n-channel enhancement mode MOSFET.

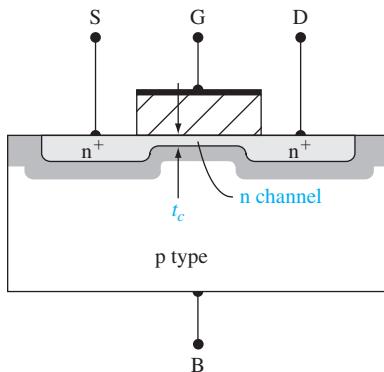


Figure 10.41 | Cross section of an n-channel depletion mode MOSFET.

thickness t_c must be less than the maximum induced space charge width in order to be able to turn the device off. The general I_D versus V_{DS} family of curves for an n-channel depletion mode MOSFET is shown in Figure 10.42.

In the next section, we derive the ideal current–voltage relation for the n-channel MOSFET. In the **nonsaturation region**, we obtain

$$I_D = \frac{W\mu_n C_{ox}}{2L} [2(V_{GS} - V_T)V_{DS} - V_{DS}^2] \quad (10.44a)$$

which can be written as

$$I_D = \frac{k'_n}{2} \cdot \frac{W}{L} \cdot [2(V_{GS} - V_T)V_{DS} - V_{DS}^2] \quad (10.44b)$$

or

$$I_D = K_n [2(V_{GS} - V_T)V_{DS} - V_{DS}^2] \quad (10.44c)$$

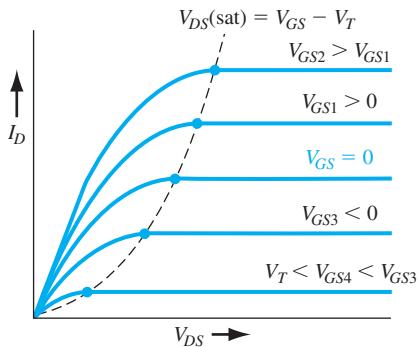


Figure 10.42 | Family of I_D versus V_{DS} curves for an n-channel depletion mode MOSFET.

The parameter $k'_n = \mu_n C_{ox}$ is called the *process conduction parameter* for the n-channel MOSFET and has units of A/V². The parameter $K_n = (W\mu_n C_{ox})/2L = (k'_n/2) \cdot (W/L)$ is called the *conduction parameter* for the n-channel MOSFET and also has units of A/V².

When the transistor is biased in the saturation region, the ideal current–voltage relation is given by

$$I_D = \frac{W\mu_n C_{ox}}{2L} (V_{GS} - V_T)^2 \quad (10.45a)$$

which can be written as

$$I_D = \frac{k'_n}{2} \cdot \frac{W}{L} \cdot (V_{GS} - V_T)^2 \quad (10.45b)$$

or

$$I_D = K_n(V_{GS} - V_T)^2 \quad (10.45c)$$

In general, for a given technology, the process conduction parameter, k'_n , is a constant. From Equations (10.44b) and (10.45b), then, we see that the design of a MOSFET, in terms of current capability, is determined by the width-to-length parameter.

The operation of a p-channel device is the same as that of the n-channel device, except the charge carrier is the hole and the conventional current direction and voltage polarities are reversed.

*10.3.3 Current–Voltage Relationship—Mathematical Derivation

In the previous section, we qualitatively discussed the current–voltage characteristics. In this section, we derive the mathematical relation between the drain current, the gate-to-source voltage, and the drain-to-source voltage. Figure 10.43 shows the geometry of the device that we use in this derivation.

In this analysis, we make the following assumptions:

1. The current in the channel is due to drift rather than diffusion.
2. There is no current through the gate oxide.
3. A gradual channel approximation is used in which $\partial E_y / \partial y \gg \partial E_x / \partial x$. This approximation means that E_x is essentially a constant.
4. Any fixed oxide charge is an equivalent charge density at the oxide–semiconductor interface.
5. The carrier mobility in the channel is constant.

We start the analysis with Ohm's law, which can be written as

$$J_x = \sigma E_x \quad (10.46)$$

where σ is the channel conductivity and E_x is the electric field along the channel created by the drain-to-source voltage. The channel conductivity is given by $\sigma = e\mu_n n(y)$, where μ_n is the electron mobility and $n(y)$ is the electron concentration in the inversion layer.

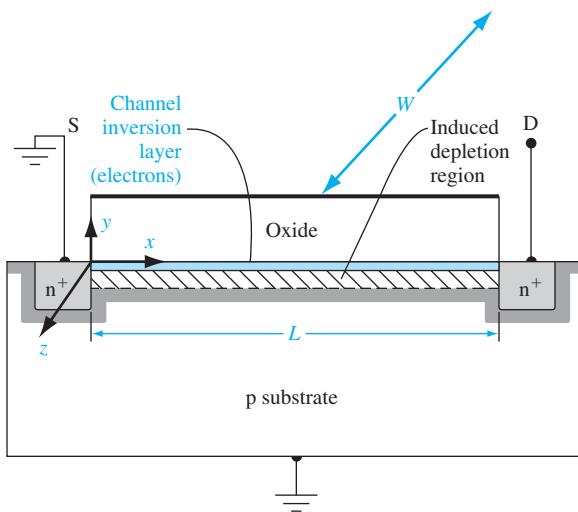


Figure 10.43 | Geometry of a MOSFET for I_D versus V_{DS} derivation.

The total channel current is found by integrating J_x over the cross-sectional area in the y and z directions. Then

$$I_x = \int_y \int_z J_x dy dz \quad (10.47)$$

We may write that

$$Q'_n = - \int en(y) dy \quad (10.48)$$

where Q'_n is the inversion layer charge per unit area and is a negative quantity for this case.

Equation (10.47) then becomes

$$I_x = -W\mu_n Q'_n E_x \quad (10.49)$$

where W is the channel width, the result of integrating over z .

Two concepts we use in the current–voltage derivation are charge neutrality and Gauss's law. Figure 10.44 shows the charge densities through the device for $V_{GS} > V_T$. The charges are all given in terms of charge per unit area. Using the concept of charge neutrality, we can write

$$Q'_m + Q'_{ss} + Q'_n + Q'_{SD(\max)} = 0 \quad (10.50)$$

The inversion layer charge and induced space charge are negative for this n-channel device.

Gauss's law can be written as

$$\oint_s \epsilon E_n dS = Q_T \quad (10.51)$$

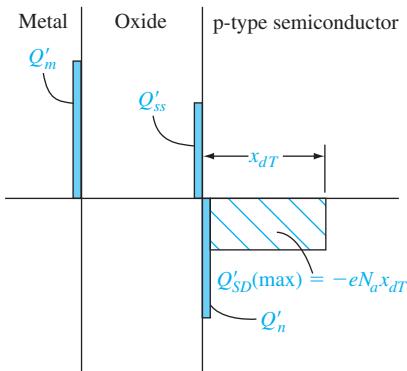


Figure 10.44 | Charge distribution in the n-channel enhancement mode MOSFET for $V_{GS} > V_T$.

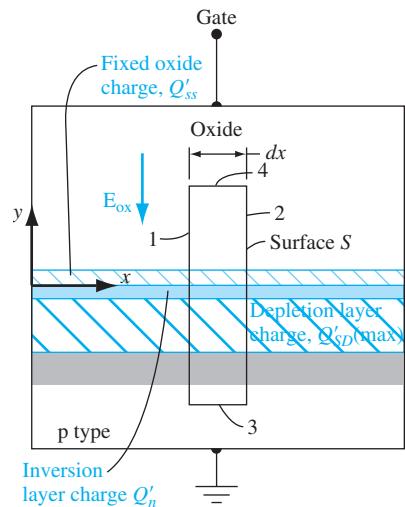


Figure 10.45 | Geometry for applying Gauss's law.

where the integral is over a closed surface. Q_T is the total charge enclosed by the surface, and E_n is the outward directed normal component of the electric field crossing the surface S . Gauss's law is applied to the surface defined in Figure 10.45. Since the surface must be enclosed, we must take into account the two end surfaces in the x - y plane. However, there is no z -component of the electric field so these two end surfaces do not contribute to the integral of Equation (10.51).

Now consider the surfaces labeled 1 and 2 in Figure 10.45. From the gradual channel approximation, we assume that E_x is essentially a constant along the channel length. This assumption means that E_x into surface 2 is the same as E_x out of surface 1. Since the integral in Equation (10.51) involves the outward component of the E -field, the contributions of surfaces 1 and 2 cancel each other. Surface 3 is in the neutral p region, so the electric field is zero at this surface.

Surface 4 is the only surface that contributes to Equation (10.51). Taking into account the direction of the electric field in the oxide, Equation (10.51) becomes

$$\oint_S \epsilon E_n dS = -\epsilon_{ox} E_{ox} W dx = Q_T \quad (10.52)$$

where ϵ_{ox} is the permittivity of the oxide. The total charge enclosed is

$$Q_T = [Q'_{ss} + Q'_n + Q'_{SD(\max)}]W dx \quad (10.53)$$

Combining Equations (10.52) and (10.53), we have

$$-\epsilon_{ox} E_{ox} = Q'_{ss} + Q'_n + Q'_{SD(\max)} \quad (10.54)$$

We now need an expression for E_{ox} . Figure 10.46a shows the oxide and channel. We assume that the source is at ground potential. The voltage V_x is the potential in the channel at a point x along the channel length. The potential difference across the oxide at x is a function of V_{GS} , V_x , and the metal–semiconductor work function difference.

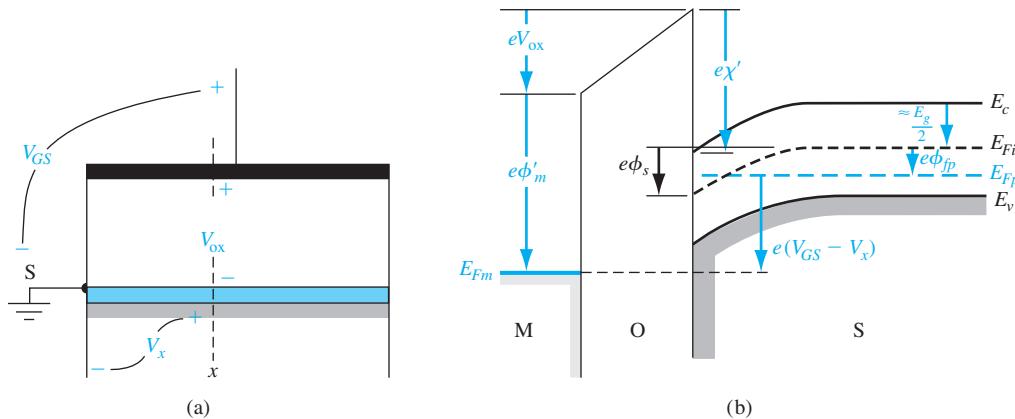


Figure 10.46 | (a) Potentials at a point x along the channel. (b) Energy-band diagram through the MOS structure at the point x .

The energy-band diagram through the MOS structure at point x is shown in Figure 10.46b. The Fermi level in the p-type semiconductor is E_{Fp} and the Fermi level in the metal is E_{Fm} . We have

$$E_{Fp} - E_{Fm} = e(V_{GS} - V_x) \quad (10.55)$$

Considering the potential barriers, we can write

$$V_{GS} - V_x = (\phi'_m + V_{ox}) - \left(\chi' + \frac{E_g}{2e} - \phi_s + \phi_{fp} \right) \quad (10.56)$$

which can also be written as

$$V_{GS} - V_x = V_{ox} + 2\phi_{fp} + \phi_{ms} \quad (10.57)$$

where ϕ_{ms} is the metal–semiconductor work function difference, and $\phi_s = 2\phi_{fp}$ for the inversion condition.

The electric field in the oxide is

$$E_{ox} = \frac{V_{ox}}{t_{ox}} \quad (10.58)$$

Combining Equations (10.54), (10.57), and (10.58), we find that

$$\begin{aligned} -\epsilon_{ox} E_{ox} &= \frac{-\epsilon_{ox}}{t_{ox}} [(V_{GS} - V_x) - (\phi_{ms} + 2\phi_{fp})] \\ &= Q'_{ss} + Q'_n + Q'_{SD}(\max) \end{aligned} \quad (10.59)$$

The inversion charge density, Q'_n , from Equation (10.59) can be substituted into Equation (10.49) and we obtain

$$I_x = -W\mu_n C_{ox} \frac{dV_x}{dx} [(V_{GS} - V_x) - V_T] \quad (10.60)$$

where $E_x = -dV_x/dx$ and V_T is the threshold voltage defined by Equation (10.31b).

We can now integrate Equation (10.60) over the length of the channel. We have

$$\int_0^L I_x dx = -W\mu_n C_{ox} \int_{V_x(0)}^{V_x(L)} [(V_{GS} - V_x) - V_T] dV_x \quad (10.61)$$

We are assuming a constant mobility μ_n . For the n-channel device, the drain current enters the drain terminal and is a constant along the entire channel length. Letting $I_D = -I_x$, Equation (10.61) becomes

$$I_D = \frac{W\mu_n C_{\text{ox}}}{2L} [2(V_{GS} - V_T)V_{DS} - V_{DS}^2] \quad (10.62)$$

Equation (10.62) is valid for $V_{GS} \geq V_T$ and for $0 \leq V_{DS} \leq V_{DS}(\text{sat})$.

Equation (10.62) can also be written as

$$I_D = \frac{k'_n}{2} \cdot \frac{W}{L} \cdot [2(V_{GS} - V_T)V_{DS} - V_{DS}^2] = K_n [2(V_{GS} - V_T)V_{DS} - V_{DS}^2] \quad (10.63)$$

where k'_n is the process conduction parameter and K_n is the conduction parameter. These parameters are described and defined in Equations (10.44b) and (10.44c).

Figure 10.47 shows plots of Equation (10.62) as a function of V_{DS} for several values of V_{GS} . We can find the value of V_{DS} at the peak current value from $\partial I_D / \partial V_{DS} = 0$. Then, using Equation (10.62), the peak current occurs when

$$V_{DS} = V_{GS} - V_T \quad (10.64)$$

This value of V_{DS} is just $V_{DS}(\text{sat})$, the point at which saturation occurs. For $V_{DS} > V_{DS}(\text{sat})$, the ideal drain current is a constant and is equal to

$$I_D(\text{sat}) = \frac{W\mu_n C_{\text{ox}}}{2L} [2(V_{GS} - V_T)V_{DS}(\text{sat}) - V_{DS}^2(\text{sat})] \quad (10.65)$$

Using Equation (10.64) for $V_{DS}(\text{sat})$, Equation (10.65) becomes

$$I_D(\text{sat}) = \frac{W\mu_n C_{\text{ox}}}{2L} (V_{GS} - V_T)^2 \quad (10.66)$$

Equation (10.66) can also be written as

$$I_D = \frac{k'_n}{2} \cdot \frac{W}{L} \cdot (V_{GS} - V_T)^2 = K_n (V_{GS} - V_T)^2 \quad (10.67)$$

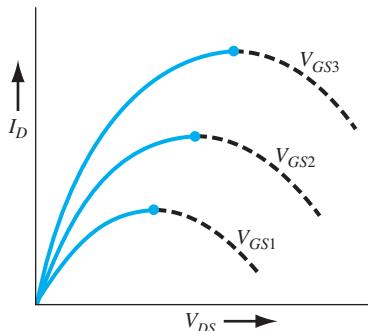


Figure 10.47 | Plots of I_D versus V_{DS} from Equation (10.62).

Equation (10.62) is the ideal current–voltage relationship of the n-channel MOSFET in the nonsaturation region for $0 \leq V_{DS} \leq V_{DS}(\text{sat})$, and Equation (10.66) is the ideal current–voltage relationship of the n-channel MOSFET in the saturation region for $V_{DS} \geq V_{DS}(\text{sat})$. These I – V expressions were explicitly derived for an n-channel enhancement mode device. However, these same equations apply to an n-channel depletion mode MOSFET in which the threshold voltage V_T is a negative quantity.

Objective: Design the width of a MOSFET such that a specified current is induced for a given applied bias.

EXAMPLE 10.7

Consider an ideal n-channel MOSFET with parameters $L = 1.25 \mu\text{m}$, $\mu_n = 650 \text{ cm}^2/\text{V}\cdot\text{s}$, $C_{ox} = 6.9 \times 10^{-8} \text{ F/cm}^2$, and $V_T = 0.65 \text{ V}$. Design the channel width W such that $I_D(\text{sat}) = 4 \text{ mA}$ for $V_{GS} = 5 \text{ V}$.

Solution

For the transition biased in the saturation region, we have, from Equation (10.66),

$$I_D(\text{sat}) = \frac{W\mu_n C_{ox}}{2L} (V_{GS} - V_T)^2$$

or

$$4 \times 10^{-3} = \frac{W(650)(6.9 \times 10^{-8})}{2(1.25 \times 10^{-4})} \cdot (5 - 0.65)^2 = 3.39 \text{ W}$$

Then

$$W = 11.8 \mu\text{m}$$

Comment

The current capability of a MOSFET is directly proportional to the channel width W . The current handling capability can be increased by increasing W .

EXERCISE PROBLEM

- Ex 10.7** The parameters of an n-channel silicon MOSFET are $\mu_n = 650 \text{ cm}^2/\text{V}\cdot\text{s}$, $t_{ox} = 8 \text{ nm} = 80 \text{ \AA}$, $W/L = 12$, and $V_T = 0.40 \text{ V}$. If the transistor is biased in the saturation region, find the drain current for (a) $V_{GS} = 0.8 \text{ V}$, (b) $V_{GS} = 1.2 \text{ V}$, and (c) $V_{GS} = 1.6 \text{ V}$.

[Ans. (a) 0.269 mA; (b) 1.077 mA; (c) 2.423 mA]

We can use the I – V relations to experimentally determine the mobility and threshold voltage parameters. From Equation (10.62), we can write, for very small values of V_{DS} ,

$$I_D = \frac{W\mu_n C_{ox}}{L} (V_{GS} - V_T)V_{DS} \quad (10.68)$$

Figure 10.48a shows a plot of Equation (10.68) as a function of V_{GS} for constant V_{DS} . A straight line is fitted through the points. The deviation from the straight line at low values of V_{GS} is due to subthreshold conduction and the deviation at higher values of V_{GS}

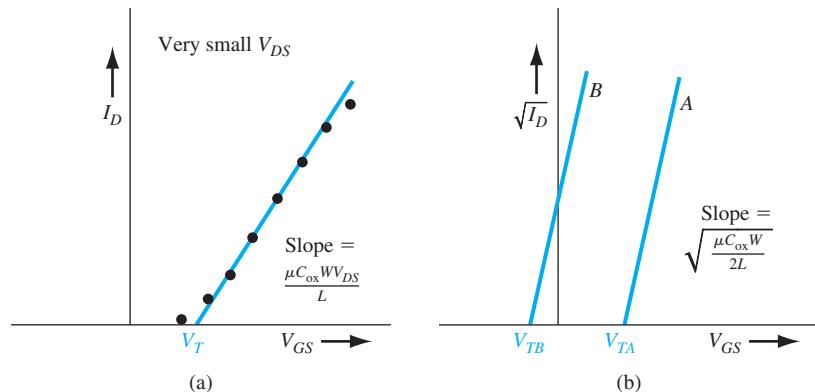


Figure 10.48 | (a) I_D versus V_{GS} (for small V_{DS}) for enhancement mode MOSFET. (b) Ideal $\sqrt{I_D}$ versus V_{GS} in saturation region for enhancement mode (curve A) and depletion mode (curve B) n-channel MOSFETs.

is due to mobility being a function of gate voltage. Both of these effects will be considered in the next chapter. The extrapolation of the straight line to zero current gives the threshold voltage, and the slope is proportional to the inversion carrier mobility.

Now consider the case when the transistor is biased in the saturation region. If we take the square root of Equation (10.66), we obtain

$$\sqrt{I_D(\text{sat})} = \sqrt{\frac{W\mu_n C_{\text{ox}}}{2L}}(V_{GS} - V_T) \quad (10.69)$$

Figure 10.48b is a plot of Equation (10.69). In the ideal case, we can obtain the same information from both curves. However, as we will see in the next chapter, the threshold voltage may be a function of V_{DS} in short-channel devices. Since Equation (10.69) applies to devices biased in the saturation region, the V_T parameter in this equation may differ from the extrapolated value determined in Figure 10.48a. In general, the nonsaturation current–voltage characteristics will produce the more reliable data.

EXAMPLE 10.8

Objective: Determine the inversion carrier mobility from experimental results.

Consider an n-channel MOSFET with $W = 15 \mu\text{m}$, $L = 2 \mu\text{m}$, and $C_{\text{ox}} = 6.9 \times 10^{-8} \text{ F/cm}^2$. Assume that the drain current in the nonsaturation region for $V_{DS} = 0.10 \text{ V}$ is $I_D = 35 \mu\text{A}$ at $V_{GS} = 1.5 \text{ V}$ and $I_D = 75 \mu\text{A}$ at $V_{GS} = 2.5 \text{ V}$.

■ Solution

From Equation (10.68), we can write

$$I_{D2} - I_{D1} = \frac{W\mu_n C_{\text{ox}}}{L} (V_{GS2} - V_{GS1})V_{DS}$$

so that

$$75 \times 10^{-6} - 35 \times 10^{-6} = \left(\frac{15}{2}\right)\mu_n(6.9 \times 10^{-8})(2.5 - 1.5)(0.10)$$

which yields

$$\mu_n = 773 \text{ cm}^2/\text{V}\cdot\text{s}$$

We can then determine

$$V_T = 0.625 \text{ V}$$

Comment

The mobility of carriers in the inversion layer is less than that in the bulk semiconductor due to the surface scattering effect. We will discuss this effect in the next chapter.

EXERCISE PROBLEM

Ex 10.8 An n-channel silicon MOSFET has the following parameters: $W = 6 \mu\text{m}$, $L = 1.5 \mu\text{m}$, and $t_{ox} = 8 \text{ nm} = 80 \text{ \AA}$. When the transistor is biased in the saturation region, the drain current is $I_D(\text{sat}) = 0.132 \text{ mA}$ at $V_{GS} = 1.0 \text{ V}$ and $I_D(\text{sat}) = 0.295 \text{ mA}$ at $V_{GS} = 1.25 \text{ V}$. Determine the electron mobility and the threshold voltage.

$$(A) 564.0 = \frac{I_D}{V_{GS}} \cdot s - A \cdot \frac{I_D}{V_{GS}} = 0.09 \approx 0.09 \text{ mA/V}$$

The current–voltage relationship of a p-channel device can be obtained by the same type of analysis. Figure 10.49 shows a p-channel enhancement mode MOSFET. The voltage polarities and current direction are the reverse of those in the n-channel device. We may note the change in the subscript notation for this device. For the current direction shown in the figure, the I – V relation for the p-channel MOSFET biased in the nonsaturation region is

$$I_D = \frac{W\mu_p C_{ox}}{2L} [2(V_{SG} + V_T)V_{SD} - V_{SD}^2] \quad (10.70)$$

Equation (10.70) is valid for $0 \leq V_{SD} \leq V_{SD}(\text{sat})$.

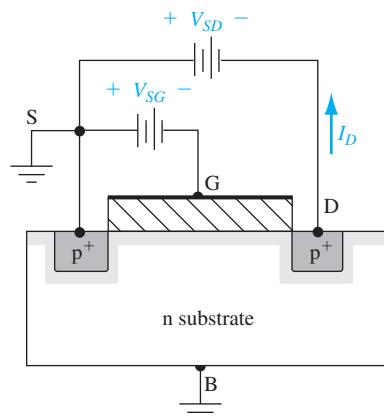


Figure 10.49 | Cross section and bias configuration for a p-channel enhancement mode MOSFET.

Equation (10.70) can also be written as

$$I_D = \frac{k'_p}{2} \cdot \frac{W}{L} \cdot [2(V_{SG} + V_T)V_{SD} - V_{SD}^2] = K_p [2(V_{SG} + V_T)V_{SD} - V_{SD}^2] \quad (10.71)$$

where $k'_p = \mu_p C_{ox}$ is the *process conduction parameter* for the p-channel MOSFET and $K_p = (W\mu_p C_{ox})/(2L) = (k'_p/2) \cdot (W/L)$ is the *conduction parameter*.

When the transistor is biased in the saturation region, the I – V relation is given by

$$I_D(\text{sat}) = \frac{W\mu_p C_{ox}}{2L} (V_{SG} + V_T)^2 \quad (10.72)$$

Equation (10.72) is valid for $V_{SD} \geq V_{SD}(\text{sat})$.

Equation (10.72) can also be written as

$$I_D = \frac{k'_p}{2} \cdot \frac{W}{L} \cdot (V_{SG} + V_T)^2 = K_p (V_{SG} + V_T)^2 \quad (10.73)$$

The source-to-drain saturation voltage is given by

$$V_{SD}(\text{sat}) = V_{SG} + V_T \quad (10.74)$$

Note the change in the sign in front of V_T and note that the mobility is now the mobility of the holes in the hole inversion layer charge. Keep in mind that V_T is negative for a p-channel enhancement mode MOSFET and positive for a depletion mode p-channel device.

One assumption we made in the derivation of the current–voltage relationship was that the charge neutrality condition given by Equation (10.50) was valid over the entire length of the channel. We implicitly assumed that $Q'_{SD}(\text{max})$ was constant along the length of the channel. The space charge width, however, varies between source and drain due to the drain-to-source voltage; it is widest at the drain when $V_{DS} > 0$. A change in the space charge density along the channel length must be balanced by a corresponding change in the inversion layer charge. An increase in the space charge width means that the inversion layer charge is reduced, implying that the drain current and drain-to-source saturation voltage are less than the ideal values. The actual saturation drain current may be as much as 20 percent less than the predicted value due to this bulk charge effect.

10.3.4 Transconductance

The MOSFET transconductance is defined as the change in drain current with respect to the corresponding change in gate voltage, or

$$g_m = \frac{\partial I_D}{\partial V_{GS}} \quad (10.75)$$

The transconductance is sometimes referred to as the transistor gain.

If we consider an n-channel MOSFET operating in the nonsaturation region, then using Equation (10.62), we have

$$g_{mL} = \frac{\partial I_D}{\partial V_{GS}} = \frac{W\mu_n C_{ox}}{L} \cdot V_{DS} \quad (10.76)$$

The transconductance increases linearly with V_{DS} but is independent of V_{GS} in the nonsaturation region.

The I - V characteristics of an n-channel MOSFET in the saturation region are given by Equation (10.66). The transconductance in this region of operation is given by

$$g_{ms} = \frac{\partial I_D(\text{sat})}{\partial V_{GS}} = \frac{W\mu_n C_{ox}}{L}(V_{GS} - V_T) \quad (10.77)$$

In the saturation region, the transconductance is a linear function of V_{GS} and is independent of V_{DS} .

The transconductance is a function of the geometry of the device as well as of carrier mobility and threshold voltage. The transconductance increases as the width of the device increases, and it also increases as the channel length and oxide thickness decrease. In the design of MOSFET circuits, the size of the transistor, in particular the channel width W , is an important engineering design parameter.

10.3.5 Substrate Bias Effects

In all of our analyses so far, the substrate, or body, has been connected to the source and held at ground potential. In MOSFET circuits, the source and body may not be at the same potential. Figure 10.50a shows an n-channel MOSFET and the associated double-subscripted voltage variables. The source-to-substrate pn junction must always be zero or reverse biased, so V_{SB} must always be greater than or equal to zero.

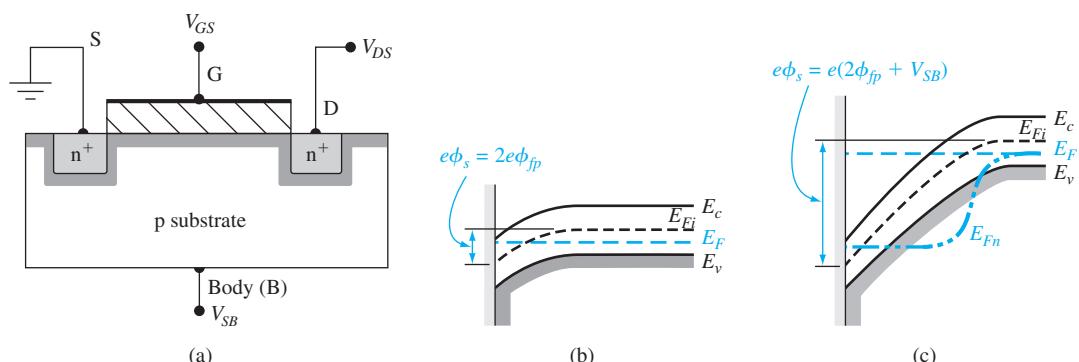


Figure 10.50 | (a) Applied voltages on an n-channel MOSFET. (b) Energy-band diagram at inversion point when $V_{SB} = 0$. (c) Energy-band diagram at inversion point when $V_{SB} > 0$ is applied.

If $V_{SB} = 0$, threshold is defined as the condition when $\phi_s = 2\phi_{fp}$ as we discussed previously and as shown in Figure 10.50b. When $V_{SB} > 0$ the surface will still try to invert when $\phi_s = 2\phi_{fp}$. However, these electrons are at a higher potential energy than are the electrons in the source. The newly created electrons will move laterally and flow out of the source terminal. When $\phi_s = 2\phi_{fp} + V_{SB}$, the surface reaches an equilibrium inversion condition. The energy-band diagram for this condition is shown in Figure 10.50c. The curve represented as E_{Fn} is the Fermi level from the p substrate through the reverse-biased source–substrate junction to the source contact.

The space charge region width under the oxide increases from the original x_{dT} value when a reverse-biased source–substrate junction voltage is applied. With an applied $V_{SB} > 0$, there is more charge associated with this region. Considering the charge neutrality condition through the MOS structure, the positive charge on the top metal gate must increase to compensate for the increased negative space charge in order to reach the threshold inversion point. So when $V_{SB} > 0$, the threshold voltage of the n-channel MOSFET increases.

When $V_{SB} = 0$, we had

$$Q'_{SD}(\text{max}) = -eN_a x_{dT} = -\sqrt{2e\epsilon_s N_a(2\phi_{fp})} \quad (10.78)$$

When $V_{SB} > 0$, the space charge width increases and we now have

$$Q'_{SD} = -eN_a x_d = -\sqrt{2e\epsilon_s N_a(2\phi_{fp} + V_{SB})} \quad (10.79)$$

The change in the space charge density is then

$$\Delta Q'_{SD} = -\sqrt{2e\epsilon_s N_a} [\sqrt{2\phi_{fp} + V_{SB}} - \sqrt{2\phi_{fp}}] \quad (10.80)$$

To reach the threshold condition, the applied gate voltage must be increased. The change in threshold voltage can be written as

$$\Delta V_T = -\frac{\Delta Q'_{SD}}{C_{ox}} = \frac{\sqrt{2e\epsilon_s N_a}}{C_{ox}} [\sqrt{2\phi_{fp} + V_{SB}} - \sqrt{2\phi_{fp}}] \quad (10.81)$$

where $\Delta V_T = V_T(V_{SB} > 0) - V_T(V_{SB} = 0)$. We may note that V_{SB} must always be positive so that, for the n-channel device, ΔV_T is always positive. The threshold voltage of the n-channel MOSFET will increase as a function of the source–substrate junction voltage.

From Equation (10.81), we may define

$$\gamma = \frac{\sqrt{2e\epsilon_s N_a}}{C_{ox}} \quad (10.82)$$

where γ is defined as the *body-effect coefficient*. Equation (10.81) may then be written as

$$\Delta V_T = \gamma [\sqrt{2\phi_{fp} + V_{SB}} - \sqrt{2\phi_{fp}}] \quad (10.83)$$

Objective: Calculate the body-effect coefficient and the change in the threshold voltage due to an applied source-to-body voltage.

EXAMPLE 10.9

Consider an n-channel silicon MOSFET at $T = 300$ K. Assume the substrate is doped to $N_a = 3 \times 10^{16}$ cm $^{-3}$ and assume the oxide is silicon dioxide with a thickness of $t_{ox} = 20$ nm = 200 Å. Let $V_{SB} = 1$ V.

Solution

We can calculate that

$$\phi_{fp} = V_t \ln \left(\frac{N_a}{n_i} \right) = (0.0259) \ln \left(\frac{3 \times 10^{16}}{1.5 \times 10^{10}} \right) = 0.3758 \text{ V}$$

and

$$C_{ox} = \frac{\epsilon_{ox}}{t_{ox}} = \frac{(3.9)(8.85 \times 10^{-14})}{200 \times 10^{-8}} = 1.726 \times 10^{-7} \text{ F/cm}^2$$

From Equation (10.82), we find the body-effect coefficient to be

$$\gamma = \frac{\sqrt{2e\epsilon_s N_a}}{C_{ox}} = \frac{[2(1.6 \times 10^{-19})(11.7)(8.85 \times 10^{-14})(3 \times 10^{16})]^{1/2}}{1.726 \times 10^{-7}}$$

or

$$\gamma = 0.5776 \text{ V}^{1/2}$$

The change in threshold voltage for $V_{SB} = 1$ V is found to be

$$\begin{aligned} \Delta V_T &= \gamma [\sqrt{2\phi_{fp} + V_{SB}} - \sqrt{2\phi_{fp}}] \\ &= (0.5776) [\sqrt{2(0.3758)} + 1 - \sqrt{2(0.3758)}] \\ &= (0.5776)[1.3235 - 0.8669] = 0.264 \text{ V} \end{aligned}$$

Comment

Figure 10.51 shows plots of $\sqrt{I_D(sat)}$ versus V_{GS} for various applied values of V_{SB} . The original threshold voltage is assumed to be $V_{TO} = 0.64$ V.

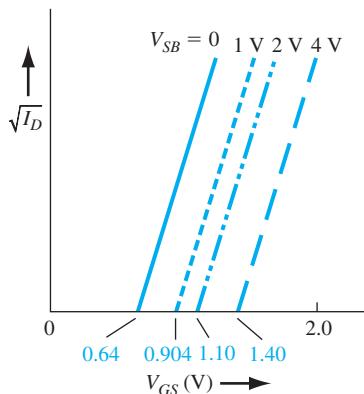


Figure 10.51 | Plots of $\sqrt{I_D}$ versus V_{GS} at several values of V_{SB} for an n-channel MOSFET.

■ EXERCISE PROBLEM

- Ex 10.9** A silicon MOSFET has the following parameters: $N_a = 10^{16} \text{ cm}^{-3}$ and $t_{ox} = 12 \text{ nm} = 120 \text{ \AA}$. Calculate (a) the body-effect coefficient and (b) the change in threshold voltage for (i) $V_{SB} = 1$ and (ii) $V_{SB} = 2 \text{ V}$.

[Ans. (a) $\gamma = 0.200 \text{ V/z}$; (ii) $\gamma = 0.162 \text{ V/z}$; (i) $(q) V_{th} = 0.069 \text{ V}$, (ii) $V_{th} = 0.0314 \text{ V}$]

If a body or substrate bias is applied to a p-channel device, the threshold voltage is shifted to more negative values. Because the threshold voltage of a p-channel enhancement mode MOSFET is negative, a body voltage will increase the applied negative gate voltage required to create inversion. The same general observation was made for the n-channel MOSFET.

TEST YOUR UNDERSTANDING

- TYU 10.6** The silicon n-channel MOSFET described in Exercise Problem Ex 10.7 is to be redesigned by changing the W/L ratio such that $I_D = 100 \mu\text{A}$ when the transistor is biased in the saturation region with $V_{GS} = 1.0 \text{ V}$.

(Ans. $W/L = 1.98$)

- TYU 10.7** The parameters of a p-channel MOSFET are $\mu_p = 310 \text{ cm}^2/\text{V-s}$, $t_{ox} = 220 \text{ \AA}$, $W/L = 60$, and $V_T = -0.40 \text{ V}$. If the transistor is biased in the saturation region, find the drain current for $V_{SG} = 1$, 1.5, and 2 V.

(Ans. $I_D = 0.526, 1.77 \text{ and } 3.74 \text{ mA}$)

- TYU 10.8** The p-channel MOSFET in TYU 10.7 is to be redesigned by changing the (W/L) ratio such that $I_D = 200 \mu\text{A}$ when the transistor is biased in the saturation region with $V_{SG} = 1.25 \text{ V}$.

(Ans. $W/L = 11.4$)

- TYU 10.9** Repeat Exercise Problem Ex 10.9 for a substrate impurity doping concentration of $N_a = 10^{15} \text{ cm}^{-3}$.

[Ans. (a) $\gamma = 0.0536 \text{ V/z}$; (i) $V_{th} = 0.0314 \text{ V}$, (ii) $V_{th} = 0.0633 \text{ V}$]

10.4 | FREQUENCY LIMITATIONS

In many applications, the MOSFET is used in a linear amplifier circuit. A small-signal equivalent circuit for the MOSFET is needed in order to mathematically analyze the electronic circuit. The equivalent circuit contains capacitances and resistances that introduce frequency effects. We initially develop a small-signal equivalent circuit and then discuss the physical factors that limit the frequency response of the MOSFET. A transistor cutoff frequency, which is a figure of merit, is then defined and an expression derived for this factor.

10.4.1 Small-Signal Equivalent Circuit

The small-signal equivalent circuit of the MOSFET is constructed from the basic MOSFET geometry. A model based on the inherent capacitances and resistances

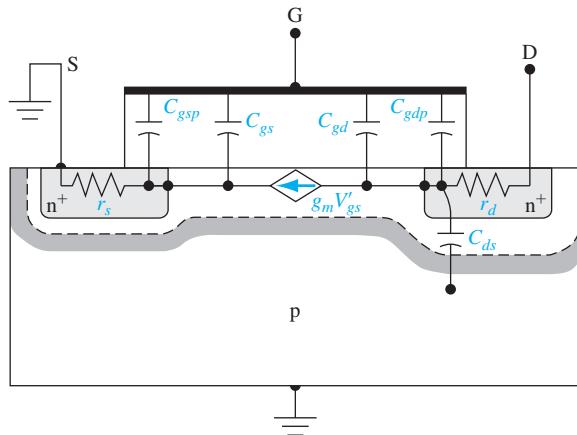


Figure 10.52 | Inherent resistances and capacitances in the n-channel MOSFET structure.

within the transistor structure, along with elements that represent the basic device equations, is shown in Figure 10.52. One simplifying assumption we will make in the equivalent circuit is that the source and substrate are both tied to ground potential.

Two of the capacitances connected to the gate are inherent in the device. These capacitances are C_{gs} and C_{gd} , which represent the interaction between the gate and the channel charge near the source and drain terminals, respectively. The remaining two gate capacitances, C_{gsp} and C_{gdp} , are parasitic or overlap capacitances. In real devices, the gate oxide will overlap the source and drain contacts because of tolerance or fabrication factors. As we will see, the drain overlap capacitance— C_{gdp} , in particular—will lower the frequency response of the device. The parameter C_{ds} is the drain-to-substrate pn junction capacitance, and r_s and r_d are the series resistances associated with the source and drain terminals. The small-signal channel current is controlled by the internal gate-to-source voltage through the transconductance.

The small-signal equivalent circuit for the n-channel common-source MOSFET is shown in Figure 10.53. The voltage V'_{gs} is the internal gate-to-source voltage that

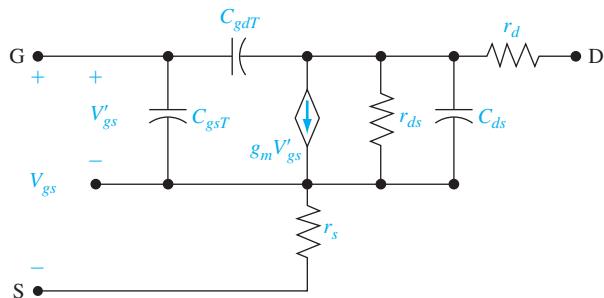


Figure 10.53 | Small-signal equivalent circuit of a common-source n-channel MOSFET.

controls the channel current. The parameters C_{gST} and C_{gdT} are the total gate-to-source and total gate-to-drain capacitances. One parameter, r_{ds} , shown in Figure 10.53, is not shown in Figure 10.52. This resistance is associated with the slope I_D versus V_{DS} . In the ideal MOSFET biased in the saturation region, I_D is independent of V_{DS} so that r_{ds} would be infinite. In short-channel-length devices, in particular, r_{ds} is finite because of channel length modulation, which we will consider in the next chapter.

A simplified small-signal equivalent circuit valid at low frequency is shown in Figure 10.54. The series resistances, r_s and r_d , have been neglected, so the drain current is essentially only a function of the gate-to-source voltage through the transconductance. The input gate impedance is infinite in this simplified model.

The source resistance r_s can have a significant effect on the transistor characteristics. Figure 10.55 shows a simplified, low-frequency equivalent circuit including r_s but neglecting r_{ds} . The drain current is given by

$$I_d = g_m V'_{gs} \quad (10.84)$$

and the relation between V_{gs} and V'_{gs} can be found from

$$V_{gs} = V'_{gs} + (g_m V'_{gs}) r_s = (1 + g_m r_s) V'_{gs} \quad (10.85)$$

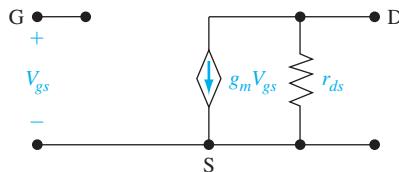


Figure 10.54 | Simplified, low-frequency small-signal equivalent circuit of a common-source n-channel MOSFET.

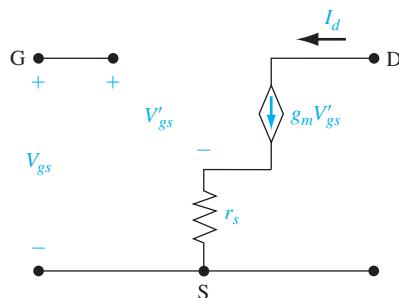


Figure 10.55 | Simplified, low-frequency small-signal equivalent circuit of common-source n-channel MOSFET including source resistance r_s .

The drain current from Equation (10.84) can now be written as

$$I_d = \left(\frac{g_m}{1 + g_m r_s} \right) V_{gs} = g'_m V_{gs} \quad (10.86)$$

The source resistance reduces the effective transconductance or transistor gain.

The equivalent circuit of the p-channel MOSFET is exactly the same as that of the n-channel except that all voltage polarities and current directions are reversed. The same capacitances and resistances that are in the n-channel model apply to the p-channel model.

10.4.2 Frequency Limitation Factors and Cutoff Frequency

There are two basic frequency limitation factors in the MOSFET. The first factor is the **channel transit time**. If we assume that carriers are traveling at their saturation drift velocity v_{sat} , then the transit time is $\tau_t = L/v_{sat}$ where L is the channel length. If $v_{sat} = 10^7$ cm/s and $L = 1\ \mu\text{m}$, then $\tau_t = 10\ \text{ps}$, which translates into a maximum frequency of 100 GHz. This frequency is much larger than the typical maximum frequency response of a MOSFET. The transit time of carriers through the channel is usually not the limiting factor in the frequency responses of MOSFETs.

The second limiting factor is the **gate or capacitance charging time**. If we neglect r_s , r_d , r_{ds} , and C_{ds} , the resulting equivalent small-signal circuit is shown in Figure 10.56 where R_L is a load resistance.

The input gate impedance in this equivalent circuit is no longer infinite. Summing currents at the input gate node, we have

$$I_i = j\omega C_{gsT} V_{gs} + j\omega C_{gdT} (V_{gs} - V_d) \quad (10.87)$$

where I_i is the input current. Likewise, summing currents at the output drain node, we have

$$\frac{V_d}{R_L} + g_m V_{gs} + j\omega C_{gdT} (V_d - V_{gs}) = 0 \quad (10.88)$$

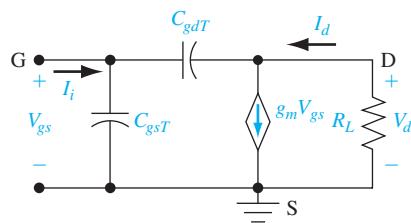


Figure 10.56 | High-frequency small-signal equivalent circuit of common-source n-channel MOSFET.

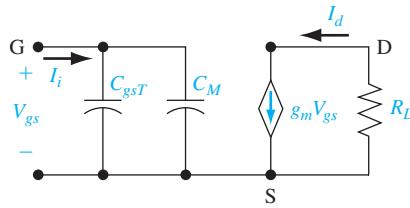


Figure 10.57 | Small-signal equivalent circuit including Miller capacitance.

Combining Equations (10.87) and (10.88) to eliminate the voltage variable V_d , we can determine the input current as

$$I_i = j\omega \left[C_{gsT} + C_{gdT} \left(\frac{1 + g_m R_L}{1 + j\omega R_L C_{gdT}} \right) \right] V_{gs} \quad (10.89)$$

Normally, $\omega R_L C_{gdT}$ is much less than unity; therefore, we may neglect the $(j\omega R_L C_{gdT})$ term in the denominator. Equation (10.89) then simplifies to

$$I_i = j\omega [C_{gsT} + C_{gdT}(1 + g_m R_L)] V_{gs} \quad (10.90)$$

Figure 10.57 shows the equivalent circuit with the equivalent input impedance described by Equation (10.90). The parameter C_M is the Miller capacitance and is given by

$$C_M = C_{gdT}(1 + g_m R_L) \quad (10.91)$$

The serious effect of the drain overlap capacitance now becomes apparent. When the transistor is operating in the saturation region, C_{gd} essentially becomes zero, but C_{gdT} is a constant. This parasitic capacitance is multiplied by the gain of the transistor and can become a significant factor in the input impedance.

The cutoff frequency f_T is defined to be the frequency at which the magnitude of the current gain of the device is unity, or when the magnitude of the input current I_i is equal to the ideal load current I_d . From Figure 10.57, we can see that

$$I_i = j\omega(C_{gsT} + C_M)V_{gs} \quad (10.92)$$

and the ideal load current is

$$I_d = g_m V_{gs} \quad (10.93)$$

The magnitude of the current gain is then

$$\left| \frac{I_d}{I_i} \right| = \frac{g_m}{2\pi f(C_{gsT} + C_M)} \quad (10.94)$$

Setting the magnitude of the current gain equal to unity at the **cutoff frequency**, we find

$$f_T = \frac{g_m}{2\pi(C_{gsT} + C_M)} = \frac{g_m}{2\pi C_G} \quad (10.95)$$

where **C_G** is the equivalent input gate capacitance.

In the ideal MOSFET, the overlap or parasitic capacitances, C_{gsp} and C_{gdsp} , are zero. Also, when the transistor is biased in the saturation region, C_{gd} approaches zero and C_{gs} is approximately $C_{ox}WL$. The transconductance of the ideal MOSFET biased in the saturation region and assuming a constant mobility is given by Equation (10.77) as

$$g_{ms} = \frac{W\mu_n C_{ox}}{L} (V_{GS} - V_T)$$

Then, for this ideal case, the cutoff frequency is

$$f_T = \frac{g_m}{2\pi C_G} = \frac{\frac{W\mu_n C_{ox}}{L} (V_{GS} - V_T)}{2\pi(C_{ox}WL)} = \frac{\mu_n(V_{GS} - V_T)}{2\pi L^2} \quad (10.96)$$

Objective: Calculate the cutoff frequency of an ideal MOSFET with a constant mobility.

Assume that the electron mobility in an n-channel device is $\mu_n = 400 \text{ cm}^2/\text{V}\cdot\text{s}$ and that the channel length is $L = 4 \mu\text{m}$. Also assume that $V_T = 1 \text{ V}$ and $V_{GS} = 3 \text{ V}$.

EXAMPLE 10.10

From Equation (10.96), the cutoff frequency is

$$f_T = \frac{\mu_n(V_{GS} - V_T)}{2\pi L^2} = \frac{400(3-1)}{2\pi(4 \times 10^{-4})^2} = 796 \text{ MHz}$$

Comment

In an actual MOSFET, the effect of the parasitic capacitance will substantially reduce the cutoff frequency from that calculated in this example.

EXERCISE PROBLEM

Ex 10.10 An n-channel silicon MOSFET has the following parameters: $\mu_n = 420 \text{ cm}^2/\text{V}\cdot\text{s}$, $t_{ox} = 18 \text{ nm} = 180 \text{ \AA}$, $L = 1.2 \mu\text{m}$, $W = 24 \mu\text{m}$, and $V_T = 0.4 \text{ V}$. The transistor is biased in the saturation region at $V_{GS} = 1.5 \text{ V}$. Determine the cutoff frequency.
(Ans. 1.1 GHz)

TEST YOUR UNDERSTANDING

TYU 10.10 Consider the n-channel MOSFET described in Exercise Problem Ex 10.10.

The transistor is connected to an effective load resistance of $R_L = 100 \text{ k}\Omega$.

Calculate the ratio of Miller capacitance C_M to gate-to-drain capacitance C_{gdT} .

(Ans. 178)

*10.5 | THE CMOS TECHNOLOGY

The primary objective of this book is to present the basic physics of semiconductor materials and devices without considering in detail the various fabrication processes; this important subject is left to other books. However, there is one MOS technology

that is used extensively, for which the basic fabrication techniques must be considered in order to understand essential characteristics of these devices and circuits. The MOS technology we consider briefly is the complementary MOS, or CMOS, process.

We have considered the physics of both n-channel and p-channel enhancement mode MOSFETs. Both devices are used in a CMOS inverter, which is the basis of CMOS digital logic circuits. The dc power dissipation in a digital circuit can be reduced to very low levels by using a complementary p-channel and n-channel pair.

It is necessary to form electrically isolated p- and n-substrate regions in an integrated circuit to accommodate the n- and p-channel transistors. The p-well process has been a commonly used technique for CMOS circuits. The process starts with a fairly low doped n-type silicon substrate in which the p-channel MOSFET will be fabricated. A diffused p region, called a p well, is formed in which the n-channel MOSFET will be fabricated. In most cases, the p-type substrate doping level must be larger than the n-type substrate doping level to obtain the desired threshold voltages. The larger p doping can easily compensate the initial n doping to form the p well. A simplified cross section of the p-well CMOS structure is shown in Figure 10.58a. The notation FOX stands for field oxide, which is a relatively thick oxide separating the devices. The FOX prevents either the n or p substrate from becoming inverted and helps maintain isolation between the two devices. In practice, additional processing steps must be included; for example, providing connections so that the p well and n substrate can be electrically connected to the appropriate voltages. The n substrate

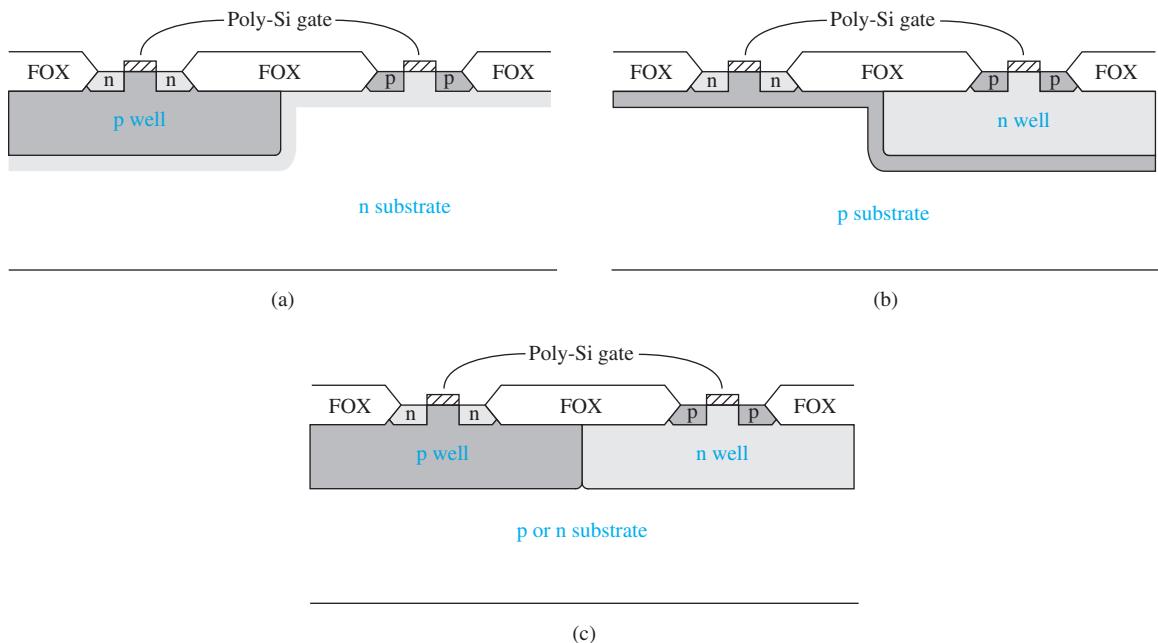


Figure 10.58 | CMOS structures: (a) p well, (b) n well, and (c) twin well.
(From Yang [22].)

must always be at a higher potential than the p well; therefore, this pn junction will always be reverse biased.

With ion implantation now being extensively used for threshold voltage control, both the n-well CMOS process and twin-well CMOS process can be used. The n-well CMOS process, shown in Figure 10.58b, starts with an optimized p-type substrate that is used to form the n-channel MOSFETs. (The n-channel MOSFETs, in general, have superior characteristics, so this starting point should yield excellent n-channel devices.) The n well is then added, in which the p-channel devices are fabricated. The n-well doping can be controlled by ion implantation.

The twin-well CMOS process, shown in Figure 10.58c, allows both the p-well and n-well regions to be optimally doped to control the threshold voltage and transconductance of each transistor. The twin-well process allows a higher packing density because of self-aligned channel stops.

One major problem in CMOS circuits has been latch-up. *Latch-up* refers to a high-current, low-voltage condition that may occur in a four-layer pnpn structure. Figure 10.59a shows the circuit of a CMOS inverter and Figure 10.59b shows a simplified integrated circuit layout of the inverter circuit. In the CMOS layout, p⁺ source to n substrate to p well to p⁺ source forms such a four-layer structure.

The equivalent circuit of this four-layer structure is shown in Figure 10.60. The silicon-controlled rectifier action involves the interaction of the parasitic pnp and npn bipolar transistors. Bipolar transistors are discussed in Chapter 12. The npn transistor corresponds to the vertical n⁺-source to p-well to n-substrate structure and the pnp transistor corresponds to the lateral p-well to n-substrate to p⁺-source structure. Under normal CMOS operation, both parasitic bipolar transistors are cut off. However, under certain conditions, avalanche breakdown may occur in the p-well to n-substrate junction, driving both bipolar transistors into saturation. This high-current, low-voltage condition—latch-up—can sustain itself by positive feedback. The condition can prevent the CMOS circuit from operating and can also cause permanent damage and burnout of the circuit.

Latch-up can be prevented if the product $\beta_n \beta_p$ is less than unity at all times, where β_n and β_p are the common-emitter current gains of the npn and pnp parasitic

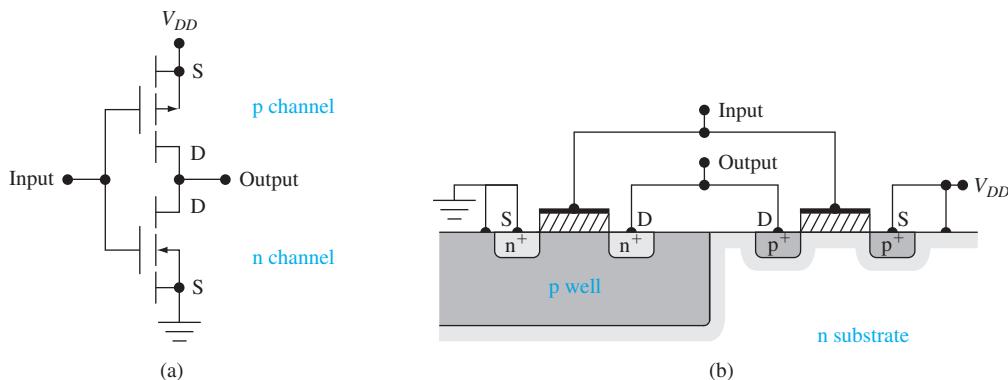


Figure 10.59 | (a) CMOS inverter circuit. (b) Simplified integrated circuit cross section of CMOS inverter.

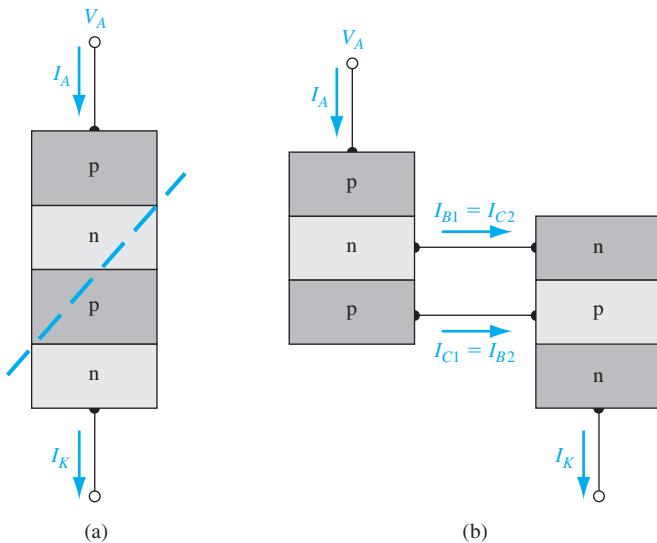


Figure 10.60 | (a) The splitting of the basic pn-pn structure. (b) The two-transistor equivalent circuit of the four-layered pn-pn device.

bipolar transistors, respectively. One method of preventing latch-up is to “kill” the minority carrier lifetime. Minority carrier lifetime degradation can be accomplished by gold doping or neutron irradiation, either of which introduces deep traps within the semiconductor. The deep traps increase the excess minority carrier recombination rate and reduce current gain. A second method of preventing latch-up is by using proper circuit layout techniques. If the two bipolar transistors can be effectively decoupled, then latch-up can be minimized or prevented. The two parasitic bipolar transistors can also be decoupled by using a different fabrication technology. The silicon-on-insulator technology, for example, allows the n-channel and the p-channel MOSFETs to be isolated from each other by an insulator. This isolation decouples the parasitic bipolar transistors.

10.6 | SUMMARY

- The fundamental physics and characteristics of the metal-oxide–semiconductor field-effect transistor (MOSFET) have been considered in this chapter.
- The heart of the MOSFET is the MOS capacitor. The energy bands in the semiconductor adjacent to the oxide–semiconductor interface bend, depending upon the voltage applied to the gate.
- An inversion layer of electrons can be created at the oxide–semiconductor surface in a p-type semiconductor by applying a sufficiently positive gate voltage, and an inversion layer of holes can be created at the oxide–semiconductor surface in an n-type semiconductor by applying a sufficiently negative gate voltage.
- The threshold voltage is the applied gate voltage required to reach the threshold inversion point. The flat-band voltage was defined and discussed.

- The n-channel MOSFET, both enhancement mode and depletion mode, and the p-channel MOSFET, both enhancement mode and depletion mode, were described.
- The basic transistor action is the modulation of the current at the drain terminal by the gate-to-source voltage.
- The ideal MOSFET current–voltage relations were derived.
- The body-effect coefficient was defined and discussed. The expression for the shift in threshold voltage due to the body effect was derived.
- A small-signal equivalent circuit of the MOSFET was developed.
- Various physical factors in the MOSFET that affect the frequency limitations were discussed. An expression for the cutoff frequency was developed.
- The CMOS technology was briefly considered.

GLOSSARY OF IMPORTANT TERMS

accumulation layer charge The induced charge directly under an oxide that is in excess of the thermal-equilibrium majority carrier concentration.

channel conductance The ratio of drain current to drain-to-source voltage in the limit as $V_{DS} \rightarrow 0$.

channel conductance modulation The process whereby the channel conductance varies with gate-to-source voltage.

CMOS Complementary MOS; the technology that uses both p- and n-channel devices in an electronic circuit fabricated in a single semiconductor chip.

conduction parameter The multiplying coefficient of the voltage terms to obtain the MOSFET drain current.

cutoff frequency The signal frequency at which the input ac gate current is equal to the output ac drain current.

depletion mode MOSFET The type of MOSFET in which a gate voltage must be applied to turn the device off.

enhancement mode MOSFET The type of MOSFET in which a gate voltage must be applied to turn the device on.

equivalent fixed oxide charge The effective fixed charge in the oxide, Q'_{ss} , directly adjacent to the oxide–semiconductor interface.

field-effect The phenomenon by which an electric field perpendicular to the surface of a semiconductor can modulate the conductance.

flat-band voltage The gate voltage that must be applied to create the flat-band condition in which there is no space charge region in the semiconductor under the oxide.

interface states The allowed electronic energy states within the bandgap energy at the oxide–semiconductor interface.

inversion layer charge The induced charge directly under the oxide, which is the opposite type compared with the semiconductor doping.

inversion layer mobility The mobility of carriers in the inversion layer.

metal-semiconductor work function difference The parameter ϕ_{ms} , a function of the difference between the metal work function and semiconductor electron affinity.

oxide capacitance The ratio of oxide permittivity to oxide thickness, which is the capacitance per unit area, C_{ox} .

process conduction parameter The product of carrier mobility and oxide capacitance.

saturation The condition in which the inversion charge density is zero at the drain and the drain current is no longer a function of the drain-to-source voltage.

strong inversion The condition in which the inversion charge density is larger than the magnitude of the semiconductor doping concentration.

threshold inversion point The condition in which the inversion charge density is equal in magnitude to the semiconductor doping concentration.

threshold voltage The gate voltage that must be applied to achieve the threshold inversion point.

transconductance The ratio of an incremental change in drain current to the corresponding incremental change in gate voltage.

weak inversion The condition in which the inversion charge density is less than the magnitude of the semiconductor doping concentration.

CHECKPOINT

After studying this chapter, the reader should have the ability to:

- Sketch the energy-band diagrams in the semiconductor of the MOS capacitor under various bias conditions.
- Describe the process by which an inversion layer of charge is created in a MOS capacitor.
- Discuss the reason the space charge width reaches a maximum value once the inversion layer is formed.
- Discuss what is meant by the metal–semiconductor work function difference and why this value is different between aluminum, n⁺ polysilicon, and p⁺ polysilicon gates.
- Describe what is meant by flat-band voltage.
- Define threshold voltage.
- Sketch the C–V characteristics of a MOS capacitor with p-type and n-type semiconductor substrates under high-frequency and low-frequency conditions.
- Discuss the effects of fixed trapped oxide charge and interface states on the C–V characteristics.
- Sketch the cross sections of n-channel and p-channel MOSFET structures.
- Explain the basic operation of the MOSFET.
- Discuss the I–V characteristics of the MOSFET when biased in the nonsaturation and saturation regions.
- Describe the substrate bias effects on the threshold voltage.
- Sketch the small-signal equivalent circuit, including capacitances, of the MOSFET, and explain the physical origin of each capacitance.
- Discuss the condition that defines the cutoff frequency of a MOSFET.
- Sketch the cross section of a CMOS structure.
- Discuss what is meant by latch-up in a CMOS structure.

REVIEW QUESTIONS

1. Sketch the energy-band diagrams in a MOS capacitor with an n-type substrate in accumulation, depletion, and inversion modes.
2. Describe what is meant by an inversion layer of charge. Describe how an inversion layer of charge can be formed in a MOS capacitor with a p-type substrate.

3. Why does the space charge region in the semiconductor of a MOS capacitor reach a maximum width once the inversion layer is formed?
4. Define surface potential. Does the surface potential change significantly with gate voltage once threshold is reached?
5. Sketch the energy-band diagram through a MOS structure with a p-type substrate and an n⁺ polysilicon gate under zero bias.
6. Define the flat-band voltage. Sketch the energy-band diagram in a MOS capacitor at flat band.
7. Define the threshold voltage. What is the surface potential at the threshold voltage?
8. Sketch the C–V characteristics of a MOS capacitor with an n-type substrate under the low-frequency condition. How do the characteristics change for the high-frequency condition?
9. Indicate the approximate capacitance at flat band on the C–V characteristic of a MOS capacitor with a p-type substrate under the high-frequency condition.
10. What is the effect on the C–V characteristics of a MOS capacitor with a p-type substrate if the amount of positive trapped oxide charge increases?
11. Qualitatively sketch the inversion charge density in the channel region when the transistor is biased in the nonsaturation region. Repeat for the case when the transistor is biased in the saturation region.
12. Define $V_{DS}(\text{sat})$.
13. Define enhancement mode and depletion mode for both n-channel and p-channel devices.
14. Sketch the charge distribution through a MOS capacitor with a p-type substrate when biased in the inversion mode. Write the charge neutrality equation.
15. Discuss why the threshold voltage changes when a reverse-biased source-to-substrate voltage is applied to a MOSFET.

PROBLEMS

(Note: In the following problems, assume the semiconductor and oxide in the MOS system are silicon and silicon dioxide, respectively, and assume the temperature is $T = 300$ K unless otherwise stated. Use Figure 10.16 to determine the metal–semiconductor work function difference.)

Section 10.1 The Two-Terminal MOS Structure

- 10.1** The dc charge distributions of four ideal MOS capacitors are shown in Figure P10.1. For each case: (a) Is the semiconductor n or p type? (b) Is the device biased in the accumulation, depletion, or inversion mode? (c) Draw the energy-band diagram in the semiconductor region.
- 10.2** (a) Calculate the maximum space charge width x_{dt} and the maximum space charge density $|Q'_{SD}(\text{max})|$ in a MOS capacitor with a p-type silicon substrate at $T = 300$ K for doping concentrations of (i) $N_a = 7 \times 10^{15} \text{ cm}^{-3}$ and (ii) $N_a = 3 \times 10^{16} \text{ cm}^{-3}$.
 (b) Repeat part (a) for $T = 350$ K.
- 10.3** (a) Consider a MOS capacitor at $T = 300$ K with an n-type silicon substrate. Determine the silicon doping concentration such that $|Q'_{SD}(\text{max})| = 1.25 \times 10^{-8} \text{ C/cm}^{-2}$.
 (b) What is the surface potential that results in the maximum space charge width?

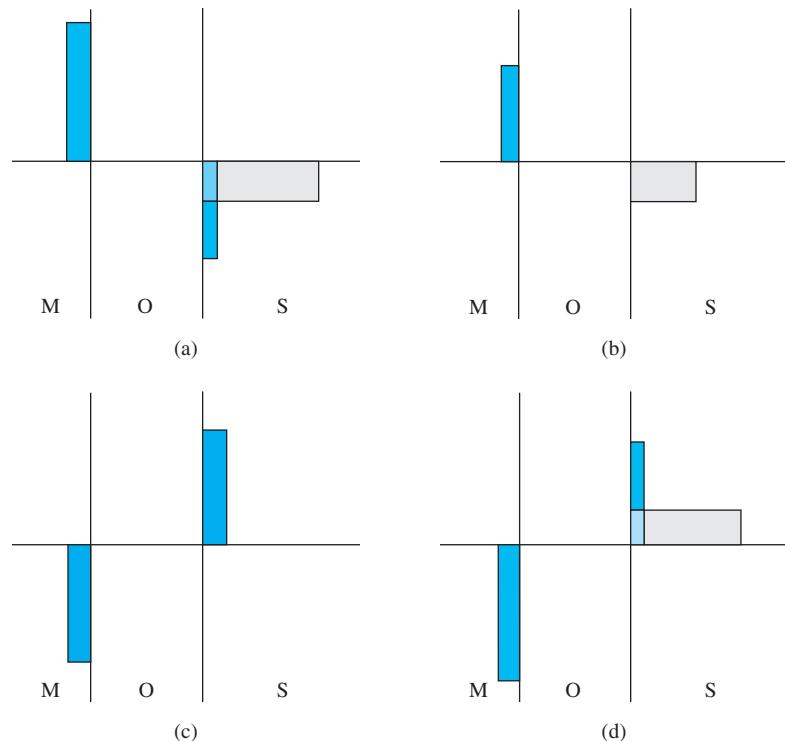


Figure P10.1 | Figure for Problem 10.1.

- 10.4** Determine the metal–semiconductor work function difference ϕ_{ms} in a MOS structure with p-type silicon for the case when the gate is (a) aluminum, (b) n[−] polysilicon, and (c) p⁺ polysilicon. Let $N_a = 6 \times 10^{15} \text{ cm}^{-3}$.
- 10.5** The silicon impurity doping concentration in an aluminum–silicon dioxide–silicon MOS device is $N_a = 4 \times 10^{16} \text{ cm}^{-3}$. Using the parameters in Example 10.2, determine the metal–semiconductor work function difference ϕ_{ms} .
- 10.6** Consider a MOS capacitor with an n-type silicon substrate. A metal–semiconductor work function difference of $\phi_{ms} = -0.30 \text{ V}$ is required. Determine the silicon doping concentration required to meet this specification when the gate is (a) n⁺ polysilicon, (b) p⁺ polysilicon, and (c) aluminum. If a particular gate cannot meet this specification, explain why.
- 10.7** (a) Consider the MOS capacitor described in Problem 10.5. For an oxide thickness of $t_{ox} = 20 \text{ nm} = 200 \text{ \AA}$ and an oxide charge of $Q'_{ss} = 5 \times 10^{10} \text{ cm}^{-2}$, calculate the flat-band voltage. (b) Repeat part (a) for an oxide thickness of $t_{ox} = 8 \text{ nm} = 80 \text{ \AA}$.
- 10.8** (a) Consider an n⁺ polysilicon–silicon dioxide–n-type silicon MOS structure. Let $N_d = 4 \times 10^{15} \text{ cm}^{-3}$. Calculate the ideal flat-band voltage for $t_{ox} = 20 \text{ nm} = 200 \text{ \AA}$. (b) Considering the results of part (a), determine the shift in flat-band voltage for (i) $Q'_{ss} = 4 \times 10^{10} \text{ cm}^{-2}$ and (ii) $Q'_{ss} = 10^{11} \text{ cm}^{-2}$. (c) Repeat parts (a) and (b) for an oxide thickness of $t_{ox} = 12 \text{ nm} = 120 \text{ \AA}$.

- 10.9** Consider an aluminum gate–silicon dioxide–p-type silicon MOS structure with $t_{ox} = 450 \text{ \AA}$. The silicon doping is $N_a = 2 \times 10^{16} \text{ cm}^{-3}$ and the flat-band voltage is $V_{FB} = -1.0 \text{ V}$. Determine the fixed oxide charge Q'_{ss} .
- 10.10** Consider a MOS device with a p-type silicon substrate with $N_a = 2 \times 10^{16} \text{ cm}^{-3}$. The oxide thickness is $t_{ox} = 15 \text{ nm} = 150 \text{ \AA}$ and the equivalent oxide charge is $Q'_{ss} = 7 \times 10^{10} \text{ cm}^{-2}$. Calculate the threshold voltage for (a) an n⁺ polysilicon gate, (b) a p⁺ polysilicon gate, and (c) an aluminum gate.
- 10.11** Repeat Problem 10.10 for an n-type silicon substrate with a doping of $N_d = 3 \times 10^{15} \text{ cm}^{-3}$.
- 10.12** A 400-Å oxide is grown on p-type silicon with $N_a = 5 \times 10^{15} \text{ cm}^{-3}$. The flat-band voltage is -0.9 V . Calculate the surface potential at the threshold inversion point as well as the threshold voltage assuming negligible oxide charge. Also find the maximum space charge width for this device.
- 10.13** A MOS device with an aluminum gate is fabricated on a p-type silicon substrate. The oxide thickness is $t_{ox} = 22 \text{ nm} = 220 \text{ \AA}$ and the trapped oxide charge is $Q'_{ss} = 4 \times 10^{10} \text{ cm}^{-2}$. The measured threshold voltage is $V_T = +0.45 \text{ V}$. Determine the p-type doping concentration.
- 10.14** Consider a MOS device with the following parameters: p⁺ polysilicon gate, n-type silicon substrate, $t_{ox} = 18 \text{ nm} = 180 \text{ \AA}$, and $Q'_{ss} = 4 \times 10^{10} \text{ cm}^{-2}$. Determine the silicon doping concentration such that the threshold voltage is in the range $-0.35 \leq V_{TP} \leq -0.25 \text{ V}$.
- 10.15** Repeat Problem 10.13 for an n-type silicon substrate if the measured threshold voltage is $V_T = -0.975 \text{ V}$. Determine the n-type doping concentration.
- 10.16** An n⁺ polysilicon gate–silicon dioxide–silicon MOS capacitor has an oxide thickness of $t_{ox} = 18 \text{ nm} = 180 \text{ \AA}$ and a doping of $N_a = 10^{15} \text{ cm}^{-3}$. The oxide charge density is $Q'_{ss} = 6 \times 10^{10} \text{ cm}^{-2}$. Calculate the (a) flat-band voltage and (b) threshold voltage.
- 10.17** An n-channel depletion mode MOSFET with an n⁺ polysilicon gate is shown in Figure 10.41. The n-channel doping is $N_d = 10^{15} \text{ cm}^{-3}$ and the oxide thickness is $t_{ox} = 500 \text{ \AA}$. The equivalent fixed oxide charge is $Q'_{ss} = 10^{10} \text{ cm}^{-2}$. The n-channel thickness t_c is equal to the maximum induced space charge width. (Disregard the space charge region at the n-channel–p-substrate junction.) (a) Determine the channel thickness t_c and (b) calculate the threshold voltage.
- 10.18** Consider a MOS capacitor with an n⁺ polysilicon gate and n-type silicon substrate. Assume $N_a = 10^{16} \text{ cm}^{-3}$ and let $E_F - E_c = 0.2 \text{ eV}$ in the n⁺ polysilicon. Assume the oxide has a thickness of $t_{ox} = 300 \text{ \AA}$. Also assume that χ' (polysilicon) = χ' (single-crystal silicon). (a) Sketch the energy-band diagrams (i) for $V_G = 0$ and (ii) at flat band. (b) Calculate the metal–semiconductor work function difference. (c) Calculate the threshold voltage for the ideal case of zero fixed oxide charge and zero interface states.
- *10.19** The threshold voltage of an n-channel MOSFET is given by Equation (10.31a). Plot V_T versus temperature over the range $200 \leq T \leq 450 \text{ K}$. Consider both an aluminum gate and an n⁺ polysilicon gate. Assume the work functions are independent of temperature and use device parameters similar to those in Example 10.4.
- *10.20** Plot the threshold voltage of an n-channel MOSFET versus p-type substrate doping concentration similar to Figure 10.21. Consider both n⁺ and p⁺ polysilicon gates. Use reasonable device parameters.

*Asterisks next to problems indicate problems that are more difficult.

- *10.21 Plot the threshold voltage of a p-channel MOSFET versus n-type substrate doping concentration similar to Figure 10.22. Consider both n⁺ and p⁺ polysilicon gates. Use reasonable device parameters.
- 10.22 Consider an NMOS device with the parameters given in Problem 10.12. Plot V_T versus t_{ox} over the range $20 \leq t_{ox} \leq 500 \text{ \AA}$.

Section 10.2 Capacitance–Voltage Characteristics

- 10.23 An ideal MOS capacitor with an n⁺ polysilicon gate has a silicon dioxide thickness of $t_{ox} = 12 \text{ nm} = 120 \text{ \AA}$ on a p-type silicon substrate doped at $N_a = 10^{16} \text{ cm}^{-3}$. Determine the capacitance C_{ox} , C'_{FB} , C'_{min} , and $C'(\text{inv})$ at (a) $f = 1 \text{ Hz}$ and (b) $f = 1 \text{ MHz}$. (c) Determine V_{FB} and V_T . (d) Sketch C'/C_{ox} versus V_G for parts (a) and (b).
- 10.24 Repeat Problem 10.23 for an ideal MOS capacitor with a p⁺ polysilicon gate and an n-type silicon substrate doped at $N_d = 5 \times 10^{14} \text{ cm}^{-3}$.
- *10.25 Using superposition, show that the shift in the flat-band voltage due to a fixed charge distribution $\rho(x)$ in the oxide is given by

$$\Delta V_{FB} = -\frac{1}{C_{ox}} \int_0^{t_{ox}} \frac{x\rho(x)}{t_{ox}} dx$$

- 10.26 Using the results of Problem 10.25, calculate the shift in flat-band voltage for $t_{ox} = 20 \text{ nm} = 200 \text{ \AA}$ for the following oxide charge distributions:
 (a) $Q'_{ss} = 8 \times 10^{10} \text{ cm}^{-2}$ is entirely located at the oxide–semiconductor interface,
 (b) $Q'_{ss} = 8 \times 10^{10} \text{ cm}^{-2}$ is uniformly distributed throughout the oxide, and
 (c) $Q'_{ss} = 8 \times 10^{10} \text{ cm}^{-2}$ forms a triangular distribution with the peak at the oxide–semiconductor interface and is zero at the metal–oxide interface.
- 10.27 An ideal MOS capacitor is fabricated by using intrinsic silicon and an n⁺ polysilicon gate. (a) Sketch the energy-band diagram through the MOS structure under flat-band conditions. (b) Sketch the low-frequency C–V characteristics from negative to positive gate voltage.
- 10.28 Consider a MOS capacitor with a p-type substrate. Assume that donor-type interface traps exist only at midgap (i.e., at E_{Fi}). Sketch the high-frequency C–V curve from accumulation to inversion. Compare this sketch to the ideal C–V plot.
- 10.29 Consider an SOS capacitor as shown in Figure P10.29. Assume the SiO₂ is ideal (no trapped charge) and has a thickness of $t_{ox} = 500 \text{ \AA}$. The doping concentrations are $N_d = 10^{16} \text{ cm}^{-3}$ and $N_a = 10^{16} \text{ cm}^{-3}$. (a) Sketch the energy-band diagram through the device for (i) flat band, (ii) $V_G = +3 \text{ V}$, and (iii) $V_G = -3 \text{ V}$. (b) Calculate the flat-band voltage. (c) Estimate the voltage across the oxide for (i) $V_G = +3 \text{ V}$ and (ii) $V_G = -3 \text{ V}$. (d) Sketch the high-frequency C–V characteristic curve.

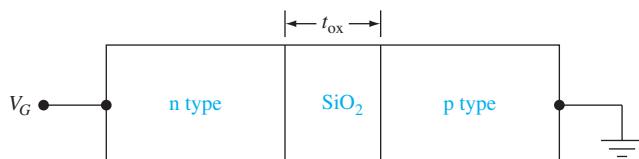


Figure P10.29 | Figure for Problem 10.29.

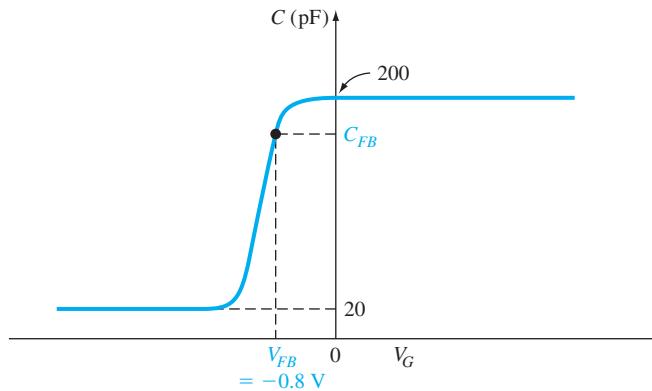


Figure P10.30 | Figure for Problem 10.30.

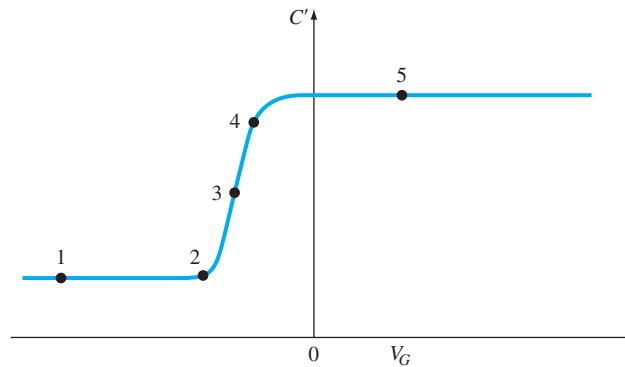


Figure P10.31 | Figure for Problem 10.31.

- 10.30** The high-frequency C - V characteristic curve of a MOS capacitor is shown in Figure P10.30. The area of the device is 2×10^{-3} cm 2 . The metal–semiconductor work function difference is $\phi_{ms} = -0.50$ V, the oxide is SiO $_2$, the semiconductor is silicon, and the semiconductor doping concentration is 2×10^{16} cm $^{-3}$. (a) Is the semiconductor n or p type? (b) What is the oxide thickness? (c) What is the equivalent trapped oxide charge density? (d) Determine the flat-band capacitance.

- 10.31** Consider the high-frequency C - V plot shown in Figure P10.31. (a) Indicate which points correspond to flat-band, inversion, accumulation, threshold, and depletion modes. (b) Sketch the energy-band diagram in the semiconductor for each condition.

Section 10.3 The Basic MOSFET Operation

- 10.32** An expression that includes the inversion charge density is given by Equation (10.59). Consider the definition of threshold voltage and show that the inversion charge density goes to zero at the drain terminal at saturation. (Hint: Let $V_x = V_{DS} = V_{DS}(\text{sat})$.)

- 10.33** Consider an n-channel MOSFET with the following parameters: $k'_n = 0.18 \text{ mA/V}^2$, $W/L = 8$, and $V_T = 0.4 \text{ V}$. Determine the drain current I_D for (a) $V_{GS} = 0.8 \text{ V}$, $V_{DS} = 0.2 \text{ V}$; (b) $V_{GS} = 0.8 \text{ V}$, $V_{DS} = 1.2 \text{ V}$; (c) $V_{GS} = 0.8 \text{ V}$, $V_{DS} = 2.5 \text{ V}$; and (d) $V_{GS} = 1.2 \text{ V}$, $V_{DS} = 2.5 \text{ V}$.
- 10.34** A p-channel MOSFET has the following parameters: $k'_p = 0.10 \text{ mA/V}^2$, $W/L = 15$, and $V_T = -0.4 \text{ V}$. Calculate the drain current I_D for (a) $V_{SG} = 0.8 \text{ V}$, $V_{SD} = 0.25 \text{ V}$; (b) $V_{SG} = 0.8 \text{ V}$, $V_{SD} = 1.0 \text{ V}$; (c) $V_{SG} = 1.2 \text{ V}$, $V_{SD} = 1.0 \text{ V}$; and (d) $V_{SG} = 1.2 \text{ V}$, $V_{SD} = 2.0 \text{ V}$.
- 10.35** The parameters of an n-channel MOSFET are $k'_n = 0.6 \text{ mA/V}^2$ and $V_T = 0.8 \text{ V}$. The drain current is 1 mA with applied voltages of $V_{GS} = 1.4 \text{ V}$, $V_{SB} = 0$, and $V_{DS} = 4 \text{ V}$. (a) What is the W/L value? (b) What is the value of I_D for $V_{GS} = 1.85 \text{ V}$, $V_{SB} = 0$, and $V_{DS} = 6 \text{ V}$? (c) Determine the value of I_D for $V_{GS} = 1.2 \text{ V}$, $V_{SB} = 0$, and $V_{DS} = 0.15 \text{ V}$.
- 10.36** Consider a p-channel MOSFET with the following parameters: $k'_p = 0.12 \text{ mA/V}^2$ and $W/L = 20$. The drain current is 100 μA with applied voltages of $V_{SG} = 0$, $V_{BS} = 0$, and $V_{SD} = 1.0 \text{ V}$. (a) Determine the V_T value. (b) Determine the drain current I_D for $V_{SG} = 0.4 \text{ V}$, $V_{SB} = 0$, and $V_{SD} = 1.5 \text{ V}$. (c) What is the value of I_D for $V_{SG} = 0.6 \text{ V}$, $V_{SB} = 0$, and $V_{SD} = 0.15 \text{ V}$?
- 10.37** An ideal n-channel MOSFET has the following parameters: $V_T = 0.45 \text{ V}$, $\mu_n = 425 \text{ cm}^2/\text{V-s}$, $t_{ox} = 11 \text{ nm} = 110 \text{ \AA}$, $W = 20 \text{ } \mu\text{m}$, and $L = 1.2 \text{ } \mu\text{m}$. (a) Plot I_D versus V_{DS} for $0 \leq V_{DS} \leq 3 \text{ V}$ and for $V_{GS} = 0, 0.6, 1.2, 1.8$, and 2.4 V . Indicate on each curve the $V_{DS}(\text{sat})$ point. (b) Plot $\sqrt{I_D(\text{sat})}$ versus V_{GS} for $0 \leq V_{GS} \leq 2.4 \text{ V}$. (c) Plot I_D versus V_{GS} for $0 \leq V_{GS} \leq 2.4 \text{ V}$ and for $V_{DS} = 0.1 \text{ V}$.
- 10.38** Consider an ideal p-channel MOSFET with the following parameters: $V_T = -0.35 \text{ V}$, $\mu_p = 210 \text{ cm}^2/\text{V-s}$, $t_{ox} = 11 \text{ nm} = 110 \text{ \AA}$, $W = 35 \text{ } \mu\text{m}$, and $L = 1.2 \text{ } \mu\text{m}$. (a) Plot I_D versus V_{SD} for $0 \leq V_{SD} \leq 3 \text{ V}$ and for $V_{SG} = 0, 0.6, 1.2, 1.8$, and 2.4 V . Indicate on each curve the $V_{SD}(\text{sat})$ point. (b) Plot $\sqrt{I_D(\text{sat})}$ versus V_{SG} for $0 \leq V_{SG} \leq 2.4 \text{ V}$. (c) Plot I_D versus V_{SG} for $0 \leq V_{SG} \leq 2.4 \text{ V}$ and for $V_{SD} = 0.1 \text{ V}$.
- 10.39** Consider an n-channel MOSFET with the same parameters as described in Problem 10.37 except that $V_T = -0.8 \text{ V}$. (a) Plot I_D versus V_{DS} for $0 \leq V_{DS} \leq 3 \text{ V}$ and for $V_{GS} = -0.8, 0, +0.8$, and $+1.6 \text{ V}$. (b) Plot $\sqrt{I_D(\text{sat})}$ versus V_{GS} for $-0.8 \leq V_{GS} \leq 1.6 \text{ V}$.
- 10.40** Consider an n-channel enhancement mode MOSFET biased as shown in Figure P10.40. Sketch the current–voltage characteristics, I_D versus V_{DS} , for (a) $V_{GD} = 0$, (b) $V_{GD} = V_T/2$, and (c) $V_{GD} = 2V_T$.
- 10.41** Figure P10.41 shows the cross section of an NMOS device that includes source and drain resistances. These resistances take into account the bulk n^+ semiconductor resistance and the ohmic contact resistance. The current–voltage relations can be generated by replacing V_{GS} by $V_G - I_D R_S$ and V_{DS} by $V_D - I_D (R_S + R_D)$ in the ideal equations. Assume transistor parameters of $V_T = 1 \text{ V}$ and $K_n = 1 \text{ mA/V}^2$. (a) Plot the following curves on the same graph: I_D versus V_D for $V_G = 2 \text{ V}$ and $V_G = 3 \text{ V}$ over the range $0 \leq V_D \leq 5 \text{ V}$ for (i) $R_S = R_D = 0$ and (ii) $R_S = R_D = 1 \text{ k}\Omega$. (b) Plot the following curves on the same graph: $\sqrt{I_D}$ versus V_G for $V_D = 0.1 \text{ V}$ and $V_D = 5 \text{ V}$ over the range $0 \leq I_D \leq 1 \text{ mA}$ for (i) $R_S = R_D = 0$ and (ii) $R_S = R_D = 1 \text{ k}\Omega$.
- 10.42** An n-channel MOSFET has the same parameters as given in Problem 10.37. The gate terminal is connected to the drain terminal. Plot I_D versus V_{DS} for $0 \leq V_{DS} \leq 5 \text{ V}$. Determine the range of V_{DS} over which the transistor is biased in the nonsaturation and saturation regions.
- 10.43** The channel conductance for a p-channel MOSFET is defined as

$$g_d = \frac{\partial I_D}{\partial V_{SD}} \Big|_{V_{SD=0}}$$

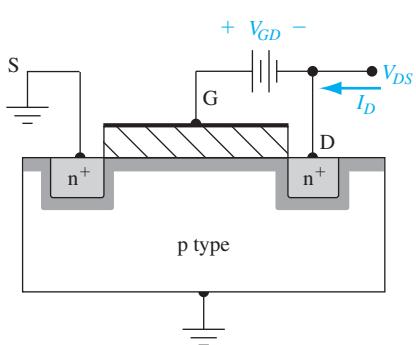


Figure P10.40 | Figure for Problem 10.40.

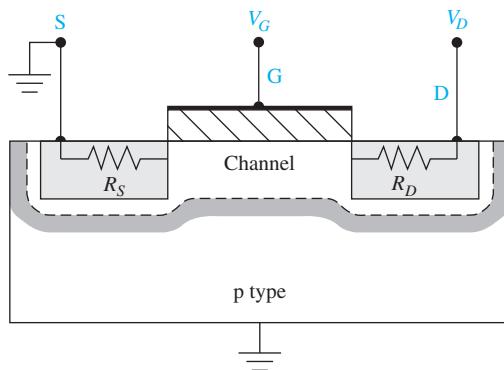


Figure P10.41 | Figure for Problem 10.41.

Plot the channel conductance for the p-channel MOSFET described in Problem 10.38 for $0 \leq V_{GS} \leq 2.4$.

10.44

The transconductance of an n-channel MOSFET is found to be $g_m = \partial I_D / \partial V_{GS} = 1.25 \text{ mA/V}$ when measured at $V_{DS} = 50 \text{ mV}$. The threshold voltage is $V_T = 0.3 \text{ V}$.

(a) Determine the conductance parameter K_n . (b) What is the current at $V_{GS} = 0.8 \text{ V}$ and $V_{DS} = 50 \text{ mV}$? (c) Determine the current at $V_{GS} = 0.8 \text{ V}$ and $V_{DS} = 1.5 \text{ V}$.

10.45

The experimental characteristics of an ideal n-channel MOSFET biased in the saturation region are shown in Figure P10.45. If $W/L = 10$ and $t_{ox} = 425 \text{ \AA}$, determine V_T and μ_n .

10.46

One curve of an n-channel MOSFET is characterized by the following parameters: $I_D(\text{sat}) = 2 \times 10^{-4} \text{ A}$, $V_{DS}(\text{sat}) = 4 \text{ V}$, and $V_T = 0.8 \text{ V}$.

- (a) What is the gate voltage?
- (b) What is the value of the conduction parameter?
- (c) If $V_G = 2 \text{ V}$ and $V_{DS} = 2 \text{ V}$, determine I_D .
- (d) If $V_G = 3 \text{ V}$ and $V_{DS} = 1 \text{ V}$, determine I_D .
- (e) For each of the conditions given in (c) and (d), sketch the inversion charge density and depletion region through the channel.

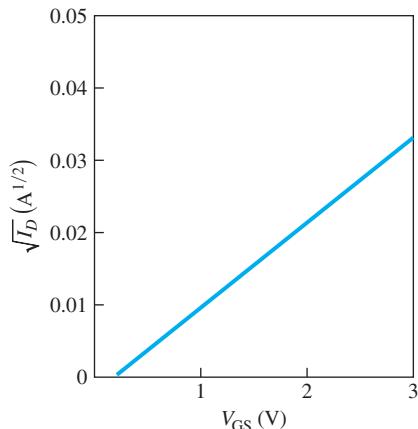


Figure P10.45 | Figure for Problem 10.45.

- 10.47** (a) An ideal n-channel MOSFET has parameters $t_{ox} = 18 \text{ nm} = 180 \text{ \AA}$, $\mu_n = 450 \text{ cm}^2/\text{V-s}$, and $V_T = 0.4 \text{ V}$. The measured current in the saturation region is $I_D(\text{sat}) = 0.8 \text{ mA}$ when biased at $V_{GS} = 2.0 \text{ V}$. Determine the (i) process conduction parameter and (ii) width-to-length ratio. (b) An ideal p-channel MOSFET has the same oxide thickness as given in part (a), a mobility of $\mu_p = 210 \text{ cm}^2/\text{V-s}$, and a threshold voltage of $V_T = -0.4 \text{ V}$. The measured current in the saturation region is also $I_D(\text{sat}) = 0.8 \text{ mA}$ when biased at $V_{SG} = 2.0 \text{ V}$. Determine the (i) process conduction parameter and (ii) width-to-length ratio.
- 10.48** Consider the n-channel MOSFET described in Problem 10.37. (a) Calculate g_{mL} for $V_{DS} = 0.10 \text{ V}$. (b) Find g_{ms} for $V_{GS} = 1.5 \text{ V}$.
- 10.49** Consider the p-channel MOSFET described in Problem 10.38. (a) Calculate g_{mL} for $V_{SD} = 0.10 \text{ V}$. (b) Find g_{ms} for $V_{SG} = 1.5 \text{ V}$.
- 10.50** An n-channel MOSFET has the following parameters: $N_a = 5 \times 10^{16} \text{ cm}^{-3}$, $t_{ox} = 15 \text{ nm} = 150 \text{ \AA}$, $\mu_n = 450 \text{ cm}^2/\text{V-s}$, $V_{FB} = -0.5 \text{ V}$, $L = 1.2 \mu\text{m}$, and $W = 8 \mu\text{m}$. (a) Determine the body-effect coefficient. (b) Plot $\sqrt{I_D(\text{sat})}$ versus V_{GS} over the range $0 \leq I_D \leq 0.5 \text{ mA}$ for source-to-body voltages of (i) $V_{SB} = 0$, (ii) $V_{SB} = 1 \text{ V}$, (iii) $V_{SB} = 2 \text{ V}$, and (iv) $V_{SB} = 4 \text{ V}$. (c) What are the threshold voltages for the conditions given in part (b)?
- 10.51** The substrate doping and body-effect coefficient of an n-channel MOSFET are $N_a = 10^{16} \text{ cm}^{-3}$ and $\gamma = 0.12 \text{ V}^{1/2}$, respectively. The threshold voltage is found to be $V_T = 0.5 \text{ V}$ when biased at $V_{SB} = 2.5 \text{ V}$. What is the threshold voltage at $V_{SB} = 0$?
- 10.52** A p-channel MOSFET has an oxide thickness of $t_{ox} = 20 \text{ nm} = 200 \text{ \AA}$ and a substrate doping of $N_d = 5 \times 10^{15} \text{ cm}^{-3}$. (a) Find the body-effect coefficient. (b) Determine the body-to-source voltage, V_{BS} , such that the shift in threshold voltage, ΔV_T , from the $V_{BS} = 0$ curve is $\Delta V_T = -0.22 \text{ V}$.
- 10.53** An NMOS device has the following parameters: n^+ poly gate, $t_{ox} = 400 \text{ \AA}$, $N_a = 10^{15} \text{ cm}^{-3}$, and $Q'_{ss} = 5 \times 10^{10} \text{ cm}^{-2}$. (a) Determine V_T . (b) Is it possible to apply a V_{SB} voltage such that $V_T = 0$? If so, what is the value of V_{SB} ?
- 10.54** Investigate the threshold voltage shift due to substrate bias. The threshold shift is given by Equation (10.81). Plot ΔV_T versus V_{SB} over the range $0 \leq V_{SB} \leq 5 \text{ V}$ for several values of N_a and t_{ox} . Determine the conditions for which ΔV_T is limited to a maximum value of 0.7 V over the range of V_{SB} .

Section 10.4 Frequency Limitations

- 10.55** Consider an ideal n-channel MOSFET with a width-to-length ratio of $(W/L) = 10$, an electron mobility of $\mu_n = 400 \text{ cm}^2/\text{V-s}$, an oxide thickness of $t_{ox} = 475 \text{ \AA}$, and a threshold voltage of $V_T = +0.65 \text{ V}$. (a) Determine the maximum value of source resistance so that the saturation transconductance g_{ms} is reduced by no more than 20 percent from its ideal value when $V_{GS} = 5 \text{ V}$. (b) Using the value of r_s calculated in part (a), how much is g_{ms} reduced from its ideal value when $V_{GS} = 3 \text{ V}$?
- 10.56** An n-channel MOSFET has the following parameters:

$$\begin{aligned}\mu_n &= 400 \text{ cm}^2/\text{V-s} & t_{ox} &= 500 \text{ \AA} \\ L &= 2 \mu\text{m} & W &= 20 \mu\text{m} \\ V_T &= +0.75 \text{ V} & &\end{aligned}$$

Assume the transistor is biased in the saturation region at $V_{GS} = 4$ V. (a) Calculate the ideal cutoff frequency. (b) Assume that the gate oxide overlaps both the source and drain contacts by $0.75 \mu\text{m}$. If a load resistance of $R_L = 10 \text{ k}\Omega$ is connected to the output, calculate the cutoff frequency.

- 10.57** Repeat Problem 10.56 for the case when the electrons are traveling at a saturation velocity of $v_{\text{sat}} = 4 \times 10^6 \text{ cm/s}$.

Summary and Review

- *10.58** Design an ideal silicon n-channel MOSFET with a polysilicon gate to have a threshold voltage of $V_T = 0.65$ V. Assume an oxide thickness of $t_{\text{ox}} = 300 \text{ \AA}$, a channel length of $L = 1.25 \mu\text{m}$, and a nominal value of $Q'_{ss} = 1.5 \times 10^{11} \text{ cm}^{-2}$. It is desired to have a drain current of $I_D = 50 \mu\text{A}$ at $V_{GS} = 2.5$ V and $V_{DS} = 0.1$ V. Determine the substrate doping concentration, channel width, and type of gate required.
- *10.59** Design an ideal silicon n-channel depletion mode MOSFET with a polysilicon gate to have a threshold voltage of $V_T = -0.65$ V. Assume an oxide thickness of $t_{\text{ox}} = 300 \text{ \AA}$, a channel length of $L = 1.25 \mu\text{m}$, and a nominal value of $Q'_{ss} = 1.5 \times 10^{11} \text{ cm}^{-2}$. It is desired to have a drain current of $I_D(\text{sat}) = 50 \mu\text{A}$ at $V_{GS} = 0$. Determine the type of gate, substrate doping concentration, and channel width required.
- *10.60** Consider the CMOS inverter circuit shown in Figure 10.59a. Ideal n- and p-channel devices are to be designed with channel lengths of $L = 2.5 \mu\text{m}$ and oxide thicknesses of $t_{\text{ox}} = 450 \text{ \AA}$. Assume the inversion channel mobilities are one-half the bulk values. The threshold voltages of the n- and p-channel transistors are to be $+0.5$ V and -0.5 V, respectively. The drain current is to be $I_D = 0.256 \text{ mA}$ when the input voltage to the inverter is 1.5 V and 3.5 V with $V_{DD} = 5$ V. The gate material is to be the same in each device. Determine the type of gate, substrate doping concentrations, and channel widths.
- *10.61** A complementary pair of ideal n-channel and p-channel MOSFETs is to be designed to produce the same I -V characteristics when they are equivalently biased. The devices are to have the same oxide thickness of 250 \AA and the same channel length of $L = 2 \mu\text{m}$. Assume the SiO_2 layer is ideal. The n-channel device is to have a channel width of $W = 20 \mu\text{m}$. Assume constant inversion layer mobilities of $\mu_n = 600 \text{ cm}^2/\text{V}\cdot\text{s}$ and $\mu_p = 220 \text{ cm}^2/\text{V}\cdot\text{s}$. (a) Determine p-type and n-type substrate doping concentrations. (b) What are the threshold voltages? (c) What is the width of the p-channel device?

READING LIST

1. Dimitrijev, S. *Principles of Semiconductor Devices*. New York: Oxford University, 2006.
2. Hu, C. C. *Modern Semiconductor Devices for Integrated Circuits*. Upper Saddle River, NJ: Pearson Prentice Hall, 2010.
3. Kano, K. *Semiconductor Devices*. Upper Saddle River, NJ: Prentice Hall, 1998.
4. Muller, R. S., and T. I. Kamins. *Device Electronics for Integrated Circuits*. 2nd ed. New York: Wiley, 1986.

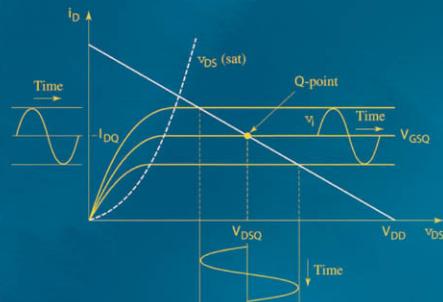
5. Ng, K. K. *Complete Guide to Semiconductor Devices*. New York: McGraw-Hill, 1995.
6. Nicollian, E. H., and J. R. Brews. *MOS Physics and Technology*. New York: Wiley, 1982.
7. Ong, D. G. *Modern MOS Technology: Processes, Devices, and Design*. New York: McGraw-Hill, 1984.
8. Pierret, R. F. *Semiconductor Device Fundamentals*. Reading, MA: Addison-Wesley, 1996.
9. Roulston, D. J. *An Introduction to the Physics of Semiconductor Devices*. New York: Oxford University Press, 1999.
10. Schröder, D. K. *Advanced MOS Devices, Modular Series on Solid State Devices*. Reading, MA: Addison-Wesley, 1987.
11. Shur, M. *Introduction to Electronic Devices*. New York: John Wiley & Sons, Inc., 1996.
- *12. _____ *Physics of Semiconductor Devices*. Englewood Cliffs, NJ: Prentice Hall, 1990.
13. Singh, J. *Semiconductor Devices: An Introduction*. New York: McGraw-Hill, 1994.
14. _____ *Semiconductor Devices: Basic Principles*. New York: Wiley, 2001.
15. Streetman, B. G., and S. K. Banerjee. *Solid State Electronic Devices*. 6th ed. Upper Saddle River, NJ: Pearson Prentice Hall, 2006.
16. Sze, S. M. *High-Speed Semiconductor Devices*. New York: Wiley, 1990.
17. Sze, S. M. and K. K. Ng. *Physics of Semiconductor Devices*, 3rd ed. Hoboken, NJ: John Wiley & Sons, Inc., 2007.
18. Taur, Y. and T. H. Ning. *Fundamentals of Modern VLSI Devices*, 2nd ed. Cambridge University Press, 2009.
- *19. Tsividis, Y. *Operation and Modeling of the MOS Transistor*. 2nd ed. Burr Ridge, IL: McGraw-Hill, 1999.
20. Werner, W. M. “The Work Function Difference of the MOS System with Aluminum Field Plates and Polycrystalline Silicon Field Plates.” *Solid State Electronics* 17, (1974), pp. 769–775.
21. Yamaguchi, T., S. Morimoto, G. H. Kawamoto, and J. C. DeLacy. “Process and Device Performance of 1 μm -Channel n-Well CMOS Technology.” *IEEE Transactions on Electron Devices* ED-31 (February 1984), pp. 205–214.
22. Yang, E. S. *Microelectronic Devices*. New York: McGraw-Hill, 1988.

*Indicates references that are at an advanced level compared to this text.

FOURTH EDITION

Microelectronics

CIRCUIT ANALYSIS AND DESIGN



DONALD A. NEAMEN

Microelectronics: Circuit Analysis and Design

Microelectronics: Circuit Analysis and Design

Fourth Edition

Donald A. Neamen

University of New Mexico



Semiconductor Materials and Diodes

This text deals with the analysis and design of circuits containing electronic devices, such as diodes and transistors. These electronic devices are fabricated using semiconductor materials, so we begin Chapter 1 with a brief discussion of the properties and characteristics of semiconductors. The intent of this brief discussion is to become familiar with some of the semiconductor material terminology, and to gain an understanding of the mechanisms that generate currents in a semiconductor.

A basic electronic device is the pn junction diode. The diode is a two-terminal device, but the i - v relationship is nonlinear. Since the diode is a nonlinear element, the analysis of circuits containing diodes is not as straightforward as the analysis of simple linear resistor circuits. A goal of the chapter is to become familiar with the analysis of diode circuits.

PREVIEW

In this chapter, we will:

- Gain a basic understanding of a few semiconductor material properties including the two types of charged carriers that exist in a semiconductor and the two mechanisms that generate currents in a semiconductor.
- Determine the properties of a pn junction including the ideal current–voltage characteristics of the pn junction diode.
- Examine dc analysis techniques for diode circuits using various models to describe the nonlinear diode characteristics.
- Develop an equivalent circuit for a diode that is used when a small, time-varying signal is applied to a diode circuit.
- Gain an understanding of the properties and characteristics of a few specialized diodes.
- As an application, design a simple electronic thermometer using the temperature characteristics of a diode.


1.1

SEMICONDUCTOR MATERIALS AND PROPERTIES

Objective: • Gain a basic understanding of a few semiconductor material properties including the two types of charged carriers that exist in a semiconductor and the two mechanisms that generate currents in a semiconductor.

Most electronic devices are fabricated by using semiconductor materials along with conductors and insulators. To gain a better understanding of the behavior of the electronic devices in circuits, we must first understand a few of the characteristics of the semiconductor material. Silicon is by far the most common semiconductor material used for semiconductor devices and integrated circuits. Other semiconductor materials are used for specialized applications. For example, gallium arsenide and related compounds are used for very high speed devices and optical devices. A list of some semiconductor materials is given in Table 1.1.

1.1.1 Intrinsic Semiconductors

An atom is composed of a nucleus, which contains positively charged protons and neutral neutrons, and negatively charged electrons that, in the classical sense, orbit the nucleus. The electrons are distributed in various “shells” at different distances from the nucleus, and electron energy increases as shell radius increases. Electrons in the outermost shell are called **valence electrons**, and the chemical activity of a material is determined primarily by the number of such electrons.

Elements in the periodic table can be grouped according to the number of valence electrons. Table 1.2 shows a portion of the periodic table in which the more common semiconductors are found. Silicon (Si) and germanium (Ge) are in group IV and are **elemental semiconductors**. In contrast, gallium arsenide is a group III–V **compound semiconductor**. We will show that the elements in group III and group V are also important in semiconductors.

Figure 1.1(a) shows five noninteracting silicon atoms, with the four valence electrons of each atom shown as dashed lines emanating from the atom. As silicon

Table 1.1 A list of some semiconductor materials

Elemental semiconductors		Compound semiconductors	
Si	Silicon	GaAs	Gallium arsenide
Ge	Germanium	GaP	Gallium phosphide
		AlP	Aluminum phosphide
		AlAs	Aluminum arsenide
		InP	Indium phosphide

Table 1.2 A portion of the periodic table

III	IV	V
5 B Boron	6 C Carbon	
13 Al Aluminum	14 Si Silicon	15 P Phosphorus
31 Ga Gallium	32 Ge Germanium	33 As Arsenic
49 In Indium		51 Sb Antimony

atoms come into close proximity to each other, the valence electrons interact to form a crystal. The final crystal structure is a tetrahedral configuration in which each silicon atom has four nearest neighbors, as shown in Figure 1.1(b). The valence electrons are shared between atoms, forming what are called **covalent bonds**. Germanium, gallium arsenide, and many other semiconductor materials have the same tetrahedral configuration.

Figure 1.1(c) is a two-dimensional representation of the lattice formed by the five silicon atoms in Figure 1.1(a). An important property of such a lattice is that valence electrons are always available on the outer edge of the silicon crystal so that additional atoms can be added to form very large single-crystal structures.

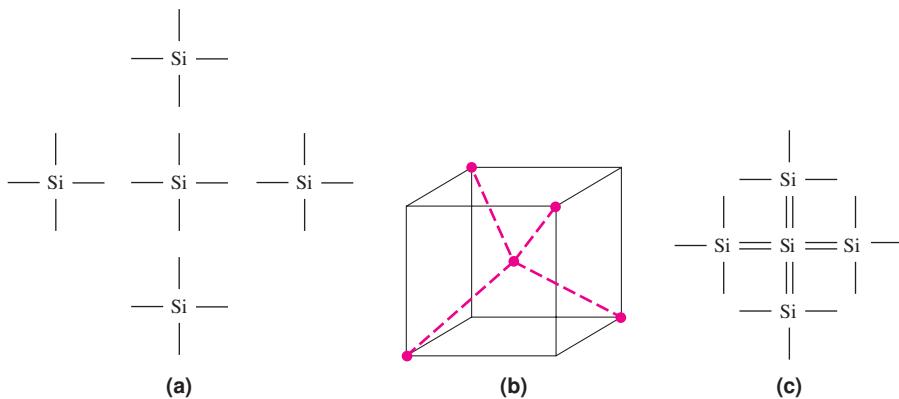


Figure 1.1 Silicon atoms in a crystal matrix: (a) five noninteracting silicon atoms, each with four valence electrons, (b) the tetrahedral configuration, (c) a two-dimensional representation showing the covalent bonding

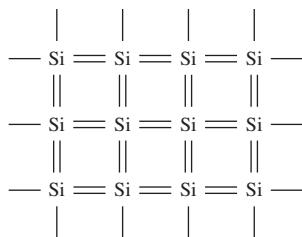


Figure 1.2 Two-dimensional representation of single crystal silicon at $T = 0\text{ K}$; all valence electrons are bound to the silicon atoms by covalent bonding

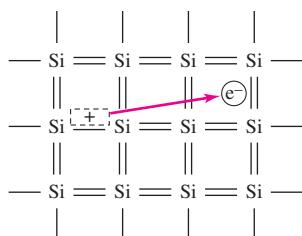


Figure 1.3 The breaking of a covalent bond for $T > 0\text{ K}$ creating an electron in the conduction band and a positively charged “empty state”

A two-dimensional representation of a silicon single crystal is shown in Figure 1.2, for $T = 0\text{ K}$, where T = temperature. Each line between atoms represents a valence electron. At $T = 0\text{ K}$, each electron is in its lowest possible energy state, so each covalent bonding position is filled. If a small electric field is applied to this material, the electrons will not move, because they will still be bound to their individual atoms. Therefore, at $T = 0\text{ K}$, silicon is an **insulator**; that is, no charge flows through it.

When silicon atoms come together to form a crystal, the electrons occupy particular allowed energy bands. At $T = 0\text{ K}$, all valence electrons occupy the valence energy band. If the temperature increases, the valence electrons may gain thermal energy. Any such electron may gain enough thermal energy to break the covalent bond and move away from its original position as schematically shown in Figure 1.3. In order to break the covalent bond, the valence electron must gain a minimum energy, E_g , called the **bandgap energy**. The electrons that gain this minimum energy now exist in the conduction band and are said to be free electrons. These free electrons in the conduction band can move throughout the crystal. The net flow of electrons in the conduction band generates a current.

An energy band diagram is shown in Figure 1.4(a). The energy E_v is the maximum energy of the valence energy band and the energy E_c is the minimum energy of the conduction energy band. The bandgap energy E_g is the difference between E_c and E_v , and the region between these two energies is called the **forbidden bandgap**. Electrons cannot exist within the forbidden bandgap. Figure 1.4(b) qualitatively shows an electron from the valence band gaining enough energy and moving into the conduction band. This process is called generation.

Materials that have large bandgap energies, in the range of 3 to 6 electron–volts¹ (eV), are insulators because, at room temperature, essentially no free electrons exist in the conduction band. In contrast, materials that contain very large numbers of free electrons at room temperature are conductors. In a *semiconductor*, the bandgap energy is on the order of 1 eV.

The net charge in a semiconductor is zero; that is, the semiconductor is neutral. If a negatively charged electron breaks its covalent bond and moves away from its original position, a positively charged “empty” state is created at that position

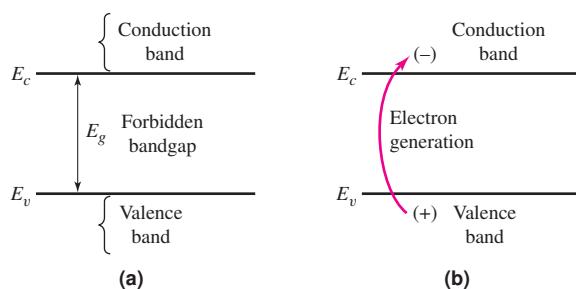


Figure 1.4 (a) Energy band diagram. Vertical scale is electron energy and horizontal scale is distance through the semiconductor, although these scales are normally not explicitly shown. (b) Energy band diagram showing the generation process of creating an electron in the conduction band and the positively charged “empty state” in the valence band.

¹An electron–volt is the energy of an electron that has been accelerated through a potential difference of 1 volt, and $1\text{ eV} = 1.6 \times 10^{-19}\text{ joules}$.

Table 1.3 Semiconductor constants

Material	E_g (eV)	B ($\text{cm}^{-3} \text{K}^{-3/2}$)
Silicon (Si)	1.1	5.23×10^{15}
Gallium arsenide (GaAs)	1.4	2.10×10^{14}
Germanium (Ge)	0.66	1.66×10^{15}

(Figure 1.3). As the temperature increases, more covalent bonds are broken, and more free electrons and positive empty states are created.

A valence electron that has a certain thermal energy and is adjacent to an empty state may move into that position, as shown in Figure 1.5, making it appear as if a positive charge is moving through the semiconductor. This positively charged “particle” is called a **hole**. In semiconductors, then, two types of charged particles contribute to the current: the negatively charged free electron, and the positively charged hole. (This description of a hole is greatly oversimplified, and is meant only to convey the concept of the moving positive charge.) We may note that the charge of a hole has the same magnitude as the charge of an electron.

The concentrations (#/ cm^3) of electrons and holes are important parameters in the characteristics of a semiconductor material, because they directly influence the magnitude of the current. An **intrinsic semiconductor** is a single-crystal semiconductor material with no other types of atoms within the crystal. In an intrinsic semiconductor, the densities of electrons and holes are equal, since the thermally generated electrons and holes are the only source of such particles. Therefore, we use the notation n_i as the **intrinsic carrier concentration** for the concentration of the free electrons, as well as that of the holes. The equation for n_i is as follows:

$$n_i = BT^{3/2} e^{\left(\frac{-E_g}{2kT}\right)} \quad (1.1)$$

where B is a coefficient related to the specific semiconductor material, E_g is the bandgap energy (eV), T is the temperature (K), k is Boltzmann’s constant (86×10^{-6} eV/K), and e , in this context, represents the exponential function. The values for B and E_g for several semiconductor materials are given in Table 1.3. The bandgap energy E_g and coefficient B are not strong functions of temperature. The intrinsic concentration n_i is a parameter that appears often in the current–voltage equations for semiconductor devices.

EXAMPLE 1.1

Objective: Calculate the intrinsic carrier concentration in silicon at $T = 300$ K.

Solution: For silicon at $T = 300$ K, we can write

$$\begin{aligned} n_i &= BT^{3/2} e^{\left(\frac{-E_g}{2kT}\right)} \\ &= (5.23 \times 10^{15})(300)^{3/2} e^{\left(\frac{-1.1}{2(86 \times 10^{-6})(300)}\right)} \end{aligned}$$

or

$$n_i = 1.5 \times 10^{10} \text{ cm}^{-3}$$

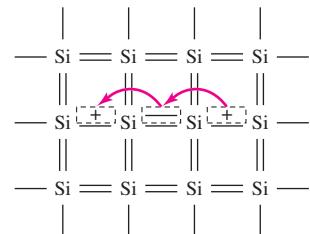


Figure 1.5 A two-dimensional representation of the silicon crystal showing the movement of the positively charged “empty state”

Comment: An intrinsic electron concentration of $1.5 \times 10^{10} \text{ cm}^{-3}$ may appear to be large, but it is relatively small compared to the concentration of silicon atoms, which is $5 \times 10^{22} \text{ cm}^{-3}$.

EXERCISE PROBLEM

Ex 1.1: Calculate the intrinsic carrier concentration in gallium arsenide and germanium at $T = 300 \text{ K}$. (Ans. GaAs, $n_i = 1.80 \times 10^6 \text{ cm}^{-3}$; Ge, $n_i = 2.40 \times 10^{13} \text{ cm}^{-3}$)

1.1.2 Extrinsic Semiconductors

Since the electron and hole concentrations in an intrinsic semiconductor are relatively small, only very small currents are possible. However, these concentrations can be greatly increased by adding controlled amounts of certain impurities. A desirable impurity is one that enters the crystal lattice and replaces (i.e., substitutes for) one of the semiconductor atoms, even though the impurity atom does not have the same valence electron structure. For silicon, the desirable substitutional impurities are from the group III and V elements (see Table 1.2).

The most common group V elements used for this purpose are phosphorus and arsenic. For example, when a phosphorus atom substitutes for a silicon atom, as shown in Figure 1.6(a), four of its valence electrons are used to satisfy the covalent bond requirements. The fifth valence electron is more loosely bound to the phosphorus atom. At room temperature, this electron has enough thermal energy to break the bond, thus being free to move through the crystal and contribute to the electron current in the semiconductor. When the fifth phosphorus valence electron moves into the conduction band, a positively charged phosphorus ion is created as shown in Figure 1.6(b).

The phosphorus atom is called a **donor impurity**, since it donates an electron that is free to move. Although the remaining phosphorus atom has a net positive charge, the atom is immobile in the crystal and cannot contribute to the current. Therefore, when a donor impurity is added to a semiconductor, free electrons are created without generating holes. This process is called **doping**, and it allows us to control the concentration of free electrons in a semiconductor.

A semiconductor that contains donor impurity atoms is called an **n-type semiconductor** (for the negatively charged electrons) and has a preponderance of electrons compared to holes.

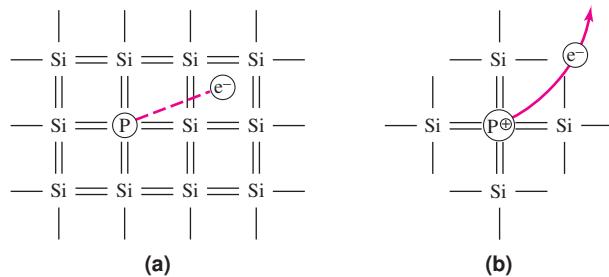


Figure 1.6 (a) Two-dimensional representation of a silicon lattice doped with a phosphorus atom showing the fifth phosphorus valence electron, (b) the resulting positively charged phosphorus ion after the fifth valence electron has moved into the conduction band

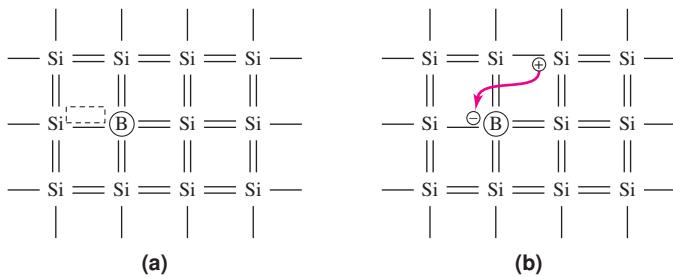


Figure 1.7 (a) Two-dimensional representation of a silicon lattice doped with a boron atom showing the vacant covalent bond position, (b) the resulting negatively charged boron ion after it has accepted an electron from the valence band. A positively charged hole is created.

The most common group III element used for silicon doping is boron. When a boron atom replaces a silicon atom, its three valence electrons are used to satisfy the covalent bond requirements for three of the four nearest silicon atoms (Figure 1.7(a)). This leaves one bond position open. At room temperature, adjacent silicon valence electrons have sufficient thermal energy to move into this position, thereby creating a hole. This effect is shown in Figure 1.7(b). The boron atom then has a net negative charge, but cannot move, and a hole is created that can contribute to a hole current.

Because the boron atom has accepted a valence electron, the boron is therefore called an **acceptor impurity**. Acceptor atoms lead to the creation of holes without electrons being generated. This process, also called doping, can be used to control the concentration of holes in a semiconductor.

A semiconductor that contains acceptor impurity atoms is called a **p-type semiconductor** (for the positively charged holes created) and has a preponderance of holes compared to electrons.

The materials containing impurity atoms are called **extrinsic semiconductors, or doped semiconductors**. The doping process, which allows us to control the concentrations of free electrons and holes, determines the conductivity and currents in the material.

A fundamental relationship between the electron and hole concentrations in a semiconductor *in thermal equilibrium* is given by

$$n_o p_o = n_i^2 \quad (1.2)$$

where n_o is the thermal equilibrium concentration of free electrons, p_o is the thermal equilibrium concentration of holes, and n_i is the intrinsic carrier concentration.

At room temperature ($T = 300$ K), each donor atom donates a free electron to the semiconductor. If the donor concentration N_d is much larger than the intrinsic concentration, we can approximate

$$n_o \approx N_d \quad (1.3)$$

Then, from Equation (1.2), the hole concentration is

$$p_o = \frac{n_i^2}{N_d} \quad (1.4)$$

Similarly, at room temperature, each acceptor atom accepts a valence electron, creating a hole. If the acceptor concentration N_a is much larger than the intrinsic concentration, we can approximate

$$p_o \approx N_a \quad (1.5)$$

Then, from Equation (1.2), the electron concentration is

$$n_o = \frac{n_i^2}{N_a} \quad (1.6)$$

EXAMPLE 1.2

Objective: Calculate the thermal equilibrium electron and hole concentrations.

(a) Consider silicon at $T = 300$ K doped with phosphorus at a concentration of $N_d = 10^{16}$ cm $^{-3}$. Recall from Example 1.1 that $n_i = 1.5 \times 10^{10}$ cm $^{-3}$.

Solution: Since $N_d \gg n_i$, the electron concentration is

$$n_o \cong N_d = 10^{16} \text{ cm}^{-3}$$

and the hole concentration is

$$p_o = \frac{n_i^2}{N_d} = \frac{(1.5 \times 10^{10})^2}{10^{16}} = 2.25 \times 10^4 \text{ cm}^{-3}$$

(b) Consider silicon at $T = 300$ K doped with boron at a concentration of $N_a = 5 \times 10^{16}$ cm $^{-3}$.

Solution: Since $N_a \gg n_i$, the hole concentration is

$$p_o \cong N_a = 5 \times 10^{16} \text{ cm}^{-3}$$

and the electron concentration is

$$n_o = \frac{n_i^2}{N_a} = \frac{(1.5 \times 10^{10})^2}{5 \times 10^{16}} = 4.5 \times 10^3 \text{ cm}^{-3}$$

Comment: We see that in a semiconductor doped with donors, the concentration of electrons is far greater than that of the holes. Conversely, in a semiconductor doped with acceptors, the concentration of holes is far greater than that of the electrons. It is also important to note that the difference in the concentrations between electrons and holes in a particular semiconductor is many orders of magnitude.

EXERCISE PROBLEM

Ex 1.2: (a) Calculate the majority and minority carrier concentrations in silicon at $T = 300$ K for (i) $N_d = 2 \times 10^{16}$ cm $^{-3}$ and (ii) $N_a = 10^{15}$ cm $^{-3}$. (b) Repeat part (a) for GaAs. (Ans. (a) (i) $n_o = 2 \times 10^{16}$ cm $^{-3}$, $p_o = 1.125 \times 10^4$ cm $^{-3}$; (ii) $p_o = 10^{15}$ cm $^{-3}$, $n_o = 2.25 \times 10^{15}$ cm $^{-3}$; (b) (i) $n_o = 2 \times 10^{16}$ cm $^{-3}$, $p_o = 1.62 \times 10^{-4}$ cm $^{-3}$; (ii) $p_o = 10^{15}$ cm $^{-3}$, $n_o = 3.24 \times 10^{-3}$ cm $^{-3}$).

In an n-type semiconductor, electrons are called the **majority carrier** because they far outnumber the holes, which are termed the **minority carrier**. The results obtained in Example 1.2 clarify this definition. In contrast, in a p-type semiconductor, the holes are the majority carrier and the electrons are the minority carrier.

1.1.3 Drift and Diffusion Currents

We've described the creation of negatively charged electrons and positively charged holes in the semiconductor. If these charged particles move, a current is generated. These charged electrons and holes are simply referred to as **carriers**.

The two basic processes which cause electrons and holes to move in a semiconductor are: (a) **drift**, which is the movement caused by electric fields, and (b) **diffusion**, which is the flow caused by variations in the concentration, that is, concentration gradients. Such gradients can be caused by a nonhomogeneous doping distribution, or by the injection of a quantity of electrons or holes into a region, using methods to be discussed later in this chapter.

Drift Current Density

To understand drift, assume an electric field is applied to a semiconductor. The field produces a force that acts on free electrons and holes, which then experience a net drift velocity and net movement. Consider an n-type semiconductor with a large number of free electrons (Figure 1.8(a)). An electric field E applied in one direction produces a force on the electrons in the *opposite* direction, because of the electrons' negative charge. The electrons acquire a drift velocity v_{dn} (in cm/s) which can be written as

$$v_{dn} = -\mu_n E \quad (1.7)$$

where μ_n is a constant called the **electron mobility** and has units of $\text{cm}^2/\text{V}\cdot\text{s}$. For low-doped silicon, the value of μ_n is typically $1350 \text{ cm}^2/\text{V}\cdot\text{s}$. The mobility can be thought of as a parameter indicating how well an electron can move in a semiconductor. The negative sign in Equation (1.7) indicates that the electron drift velocity is opposite to that of the applied electric field as shown in Figure 1.8(a). The electron drift produces a drift current density J_n (A/cm^2) given by

$$J_n = -env_{dn} = -en(-\mu_n E) = +en\mu_n E \quad (1.8)$$

where n is the electron concentration ($\#/ \text{cm}^3$) and e , in this context, is the magnitude of the electronic charge. The conventional drift current is in the opposite direction from the flow of negative charge, which means that the drift current in an n-type semiconductor is in the same direction as the applied electric field.

Next consider a p-type semiconductor with a large number of holes (Figure 1.8(b)). An electric field E applied in one direction produces a force on the holes in the *same* direction, because of the positive charge on the holes. The holes acquire a drift velocity v_{dp} (in cm/s), which can be written as

$$v_{dp} = +\mu_p E \quad (1.9)$$

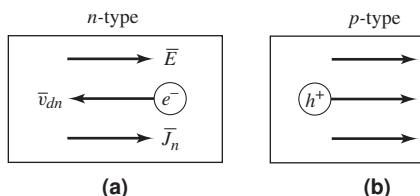


Figure 1.8 Directions of applied electric field and resulting carrier drift velocity and drift current density in (a) an n-type semiconductor and (b) a p-type semiconductor

where μ_p is a constant called the **hole mobility**, and again has units of $\text{cm}^2/\text{V}\cdot\text{s}$. For low-doped silicon, the value of μ_p is typically $480 \text{ cm}^2/\text{V}\cdot\text{s}$, which is less than half the value of the electron mobility. The positive sign in Equation (1.9) indicates that the hole drift velocity is in the same direction as the applied electric field as shown in Figure 1.8(b). The hole drift produces a drift current density J_p (A/cm^2) given by

$$J_p = +epv_{dp} = +ep(+\mu_p E) = +ep\mu_p E \quad (1.10)$$

where p is the hole concentration ($\#/ \text{cm}^3$) and e is again the magnitude of the electronic charge. The conventional drift current is in the same direction as the flow of positive charge, which means that the drift current in a p-type material is also in the same direction as the applied electric field.

Since a semiconductor contains both electrons and holes, the total drift current density is the sum of the electron and hole components. The total drift current density is then written as

$$J = en\mu_n E + ep\mu_p E = \sigma E = \frac{1}{\rho} E \quad (1.11(a))$$

where

$$\sigma = en\mu_n + ep\mu_p \quad (1.11(b))$$

and where σ is the **conductivity** of the semiconductor in $(\Omega\text{-cm})^{-1}$ and where $\rho = 1/\sigma$ is the **resistivity** of the semiconductor in $(\Omega\text{-cm})$. The conductivity is related to the concentration of electrons and holes. If the electric field is the result of applying a voltage to the semiconductor, then Equation (1.11(a)) becomes a linear relationship between current and voltage and is one form of Ohm's law.

From Equation (1.11(b)), we see that the conductivity can be changed from strongly n-type, $n \gg p$, by donor impurity doping to strongly p-type, $p \gg n$, by acceptor impurity doping. **Being able to control the conductivity of a semiconductor by selective doping is what enables us to fabricate the variety of electronic devices that are available.**

EXAMPLE 1.3

Objective: Calculate the drift current density for a given semiconductor.

Consider silicon at $T = 300 \text{ K}$ doped with arsenic atoms at a concentration of $N_d = 8 \times 10^{15} \text{ cm}^{-3}$. Assume mobility values of $\mu_n = 1350 \text{ cm}^2/\text{V}\cdot\text{s}$ and $\mu_p = 480 \text{ cm}^2/\text{V}\cdot\text{s}$. Assume the applied electric field is 100 V/cm .

Solution: The electron and hole concentrations are

$$n \cong N_d = 8 \times 10^{15} \text{ cm}^{-3}$$

and

$$p = \frac{n_i^2}{N_d} = \frac{(1.5 \times 10^{10})^2}{8 \times 10^{15}} = 2.81 \times 10^4 \text{ cm}^{-3}$$

Because of the difference in magnitudes between the two concentrations, the conductivity is given by

$$\sigma = e\mu_n n + e\mu_p p \cong e\mu_n n$$

or

$$\sigma = (1.6 \times 10^{-19})(1350)(8 \times 10^{15}) = 1.73(\Omega\text{-cm})^{-1}$$

The drift current density is then

$$J = \sigma E = (1.73)(100) = 173 \text{ A/cm}^2$$

Comment: Since $n \gg p$, the conductivity is essentially a function of the electron concentration and mobility only. We may note that a current density of a few hundred amperes per square centimeter can be generated in a semiconductor.

EXERCISE PROBLEM

Ex 1.3: Consider n-type GaAs at $T = 300 \text{ K}$ doped to a concentration of $N_d = 2 \times 10^{16} \text{ cm}^{-3}$. Assume mobility values of $\mu_n = 6800 \text{ cm}^2/\text{V}\text{-s}$ and $\mu_p = 300 \text{ cm}^2/\text{V}\text{-s}$. (a) Determine the resistivity of the material. (b) Determine the applied electric field that will induce a drift current density of 175 A/cm^2 . (Ans. (a) $0.0460 \Omega\text{-cm}$, (b) 8.04 V/cm).

Note: Two factors need to be mentioned concerning drift velocity and mobility. Equations (1.7) and (1.9) imply that the carrier drift velocities are linear functions of the applied electric field. This is true for relatively small electric fields. As the electric field increases, the carrier drift velocities will reach a maximum value of approximately 10^7 cm/s . Any further increase in electric field will not produce an increase in drift velocity. This phenomenon is called drift velocity saturation.

Electron and hole mobility values were given in Example 1.3. The mobility values are actually functions of donor and/or acceptor impurity concentrations. As the impurity concentration increases, the mobility values will decrease. This effect then means that the conductivity, Equation (1.11(b)), is not a linear function of impurity doping.

These two factors are important in the design of semiconductor devices, but will not be considered in detail in this text.

Diffusion Current Density

In the diffusion process, particles flow from a region of high concentration to a region of lower concentration. This is a statistical phenomenon related to kinetic theory. To explain, the electrons and holes in a semiconductor are in continuous motion, with an average speed determined by the temperature, and with the directions randomized by interactions with the lattice atoms. Statistically, we can assume that, at any particular instant, approximately half of the particles in the high-concentration region are moving *away* from that region toward the lower-concentration region. We can also assume that, at the same time, approximately half of the particles in the lower-concentration region are moving *toward* the high-concentration region. However, by definition, there are fewer particles in the lower-concentration region than there are in the high-concentration region. Therefore, the net result is a flow of particles away from the high-concentration region and toward the lower-concentration region. This is the basic diffusion process.

For example, consider an electron concentration that varies as a function of distance x , as shown in Figure 1.9(a). The diffusion of electrons from a high-concentration region

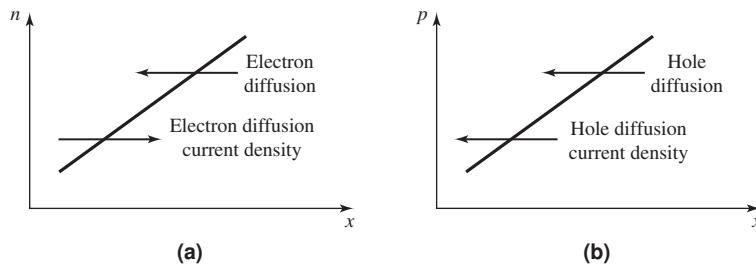


Figure 1.9 (a) Assumed electron concentration versus distance in a semiconductor, and the resulting electron diffusion and electron diffusion current density, (b) assumed hole concentration versus distance in a semiconductor, and the resulting hole diffusion and hole diffusion current density

to a low-concentration region produces a flow of electrons in the negative x direction. Since electrons are negatively charged, the conventional current direction is in the positive x direction.

The diffusion current density due to the diffusion of electrons can be written as (for one dimension)

$$J_n = eD_n \frac{dn}{dx} \quad (1.12)$$

where e , in this context, is the magnitude of the electronic charge, dn/dx is the gradient of the electron concentration, and D_n is the **electron diffusion coefficient**.

In Figure 1.9(b), the hole concentration is a function of distance. The diffusion of holes from a high-concentration region to a low-concentration region produces a flow of holes in the negative x direction. (Conventional current is in the direction of the flow of positive charge.)

The diffusion current density due to the diffusion of holes can be written as (for one dimension)

$$J_p = -eD_p \frac{dp}{dx} \quad (1.13)$$

where e is still the magnitude of the electronic charge, dp/dx is the gradient of the hole concentration, and D_p is the **hole diffusion coefficient**. Note the change in sign between the two diffusion current equations. This change in sign is due to the difference in sign of the electronic charge between the negatively charged electron and the positively charged hole.

EXAMPLE 1.4

Objective: Calculate the diffusion current density for a given semiconductor.

Consider silicon at $T = 300$ K. Assume the electron concentration varies linearly from $n = 10^{12} \text{ cm}^{-3}$ to $n = 10^{16} \text{ cm}^{-3}$ over the distance from $x = 0$ to $x = 3 \mu\text{m}$. Assume $D_n = 35 \text{ cm}^2/\text{s}$.

Solution: We have

$$J_n = eD_n \frac{dn}{dx} = eD_n \frac{\Delta n}{\Delta x} = (1.6 \times 10^{-19})(35) \left(\frac{10^{12} - 10^{16}}{0 - 3 \times 10^{-4}} \right)$$

or

$$J_n = 187 \text{ A/cm}^2$$

Comment: Diffusion current densities on the order of a few hundred amperes per square centimeter can also be generated in a semiconductor.

EXERCISE PROBLEM

Ex 1.4: Consider silicon at $T = 300 \text{ K}$. Assume the hole concentration is given by $p = 10^{16}e^{-x/L_p} (\text{cm}^{-3})$, where $L_p = 10^{-3} \text{ cm}$. Calculate the hole diffusion current density at (a) $x = 0$ and (b) $x = 10^{-3} \text{ cm}$. Assume $D_p = 10 \text{ cm}^2/\text{s}$. (Ans. (a) 16 A/cm^2 , (b) 5.89 A/cm^2)

The mobility values in the drift current equations and the diffusion coefficient values in the diffusion current equations are not independent quantities. They are related by the **Einstein relation**, which is

$$\frac{D_n}{\mu_n} = \frac{D_p}{\mu_p} = \frac{kT}{e} \cong 0.026 \text{ V} \quad (1.14)$$

at room temperature.

The *total* current density is the sum of the drift and diffusion components. Fortunately, in most cases only one component dominates the current at any one time in a given region of a semiconductor.

DESIGN POINTER

In the previous two examples, current densities on the order of 200 A/cm^2 have been calculated. This implies that if a current of 1 mA , for example, is required in a semiconductor device, the size of the device is small. The total current is given by $I = JA$, where A is the cross-sectional area. For $I = 1 \text{ mA} = 1 \times 10^{-3} \text{ A}$ and $J = 200 \text{ A/cm}^2$, the cross-sectional area is $A = 5 \times 10^{-6} \text{ cm}^2$. This simple calculation again shows why semiconductor devices are small in size.

1.1.4 Excess Carriers

Up to this point, we have assumed that the semiconductor is in thermal equilibrium. In the discussion of drift and diffusion currents, we implicitly assumed that equilibrium was not significantly disturbed. Yet, when a voltage is applied to, or a current exists in, a semiconductor device, the semiconductor is really not in equilibrium. In this section, we will discuss the behavior of nonequilibrium electron and hole concentrations.

Valence electrons may acquire sufficient energy to break the covalent bond and become free electrons if they interact with high-energy photons incident on the semiconductor. When this occurs, both an electron and a hole are produced, thus generating

an electron–hole pair. These additional electrons and holes are called **excess electrons** and **excess holes**.

When these excess electrons and holes are created, the concentrations of free electrons and holes increase above their thermal equilibrium values. This may be represented by

$$n = n_o + \delta n \quad (1.15(a))$$

and

$$p = p_o + \delta p \quad (1.15(b))$$

where n_o and p_o are the thermal equilibrium concentrations of electrons and holes, and δn and δp are the excess electron and hole concentrations.

If the semiconductor is in a steady-state condition, the creation of excess electrons and holes will not cause the carrier concentration to increase indefinitely, because a free electron may recombine with a hole, in a process called **electron–hole recombination**. Both the free electron and the hole disappear causing the excess concentration to reach a steady-state value. The mean time over which an excess electron and hole exist before recombination is called the **excess carrier lifetime**.

Excess carriers are involved in the current mechanisms of, for example, solar cells and photodiodes. These devices are discussed in Section 1.5.

Test Your Understanding

TYU 1.1 Determine the intrinsic carrier concentration in silicon, germanium, and GaAs at (a) $T = 400$ K and (b) $T = 250$ K. (Ans. (a) Si: $n_i = 4.76 \times 10^{12} \text{ cm}^{-3}$, Ge: $n_i = 9.06 \times 10^{14} \text{ cm}^{-3}$, GaAs: $n_i = 2.44 \times 10^9 \text{ cm}^{-3}$; (b) Si: $n_i = 1.61 \times 10^8 \text{ cm}^{-3}$, Ge: $n_i = 1.42 \times 10^{12} \text{ cm}^{-3}$, GaAs: $n_i = 6.02 \times 10^3 \text{ cm}^{-3}$)

TYU 1.2 (a) Consider silicon at $T = 300$ K. Assume that $\mu_n = 1350 \text{ cm}^2/\text{V}\cdot\text{s}$ and $\mu_p = 480 \text{ cm}^2/\text{V}\cdot\text{s}$. Determine the conductivity and resistivity if (a) $N_a = 2 \times 10^{15} \text{ cm}^{-3}$ and (b) $N_d = 2 \times 10^{17} \text{ cm}^{-3}$. (Ans. (a) $\sigma = 0.154 (\Omega\text{-cm})^{-1}$, $\rho = 6.51 \Omega\text{-cm}$; (b) $\sigma = 43.2 (\Omega\text{-cm})^{-1}$, $\rho = 0.0231 \Omega\text{-cm}$).

TYU 1.3 Using the results of TYU1.2, determine the drift current density if an electric field of 4 V/cm is applied to the semiconductor. (Ans. (a) 0.616 A/cm^2 , (b) 172.8 A/cm^2).

TYU 1.4 The electron and hole diffusion coefficients in silicon are $D_n = 35 \text{ cm}^2/\text{s}$ and $D_p = 12.5 \text{ cm}^2/\text{s}$, respectively. Calculate the electron and hole diffusion current densities (a) if an electron concentration varies linearly from $n = 10^{15} \text{ cm}^{-3}$ to $n = 10^{16} \text{ cm}^{-3}$ over the distance from $x = 0$ to $x = 2.5 \mu\text{m}$ and (b) if a hole concentration varies linearly from $p = 10^{14} \text{ cm}^{-3}$ to $p = 5 \times 10^{15} \text{ cm}^{-3}$ over the distance from $x = 0$ to $x = 4.0 \mu\text{m}$. (Ans. (a) $J_n = 202 \text{ A/cm}^2$, (b) $J_p = -24.5 \text{ A/cm}^2$)

TYU 1.5 A sample of silicon at $T = 300$ K is doped to $N_d = 8 \times 10^{15} \text{ cm}^{-3}$. (a) Calculate n_o and p_o . (b) If excess holes and electrons are generated such that their respective concentrations are $\delta n = \delta p = 10^{14} \text{ cm}^{-3}$, determine the total concentrations of holes and electrons. (Ans. (a) $n_o = 8 \times 10^{15} \text{ cm}^{-3}$, $p_o = 2.81 \times 10^4 \text{ cm}^{-3}$; (b) $n = 8.1 \times 10^{15} \text{ cm}^{-3}$, $p \approx 10^{14} \text{ cm}^{-3}$)



1.2 THE pn JUNCTION

Objective: • Determine the properties of a pn junction including the ideal current-voltage characteristics of the pn junction diode.

In the preceding sections, we looked at characteristics of semiconductor materials. The real power of semiconductor electronics occurs when p- and n-regions are directly adjacent to each other, forming a **pn junction**. One important concept to remember is that in most integrated circuit applications, the entire semiconductor material is a single crystal, with one region doped to be p-type and the adjacent region doped to be n-type.

1.2.1 The Equilibrium pn Junction

Figure 1.10(a) is a simplified block diagram of a pn junction. Figure 1.10(b) shows the respective p-type and n-type doping concentrations, assuming uniform doping in each region, as well as the minority carrier concentrations in each region, assuming thermal equilibrium. Figure 1.10(c) is a three-dimensional diagram of the pn junction showing the cross-sectional area of the device.

The interface at $x = 0$ is called the **metallurgical junction**. A large density gradient in both the hole and electron concentrations occurs across this junction. Initially, then, there is a diffusion of holes from the p-region into the n-region, and a diffusion of electrons from the n-region into the p-region (Figure 1.11). The flow of holes from the p-region uncovers negatively charged acceptor ions, and the flow of

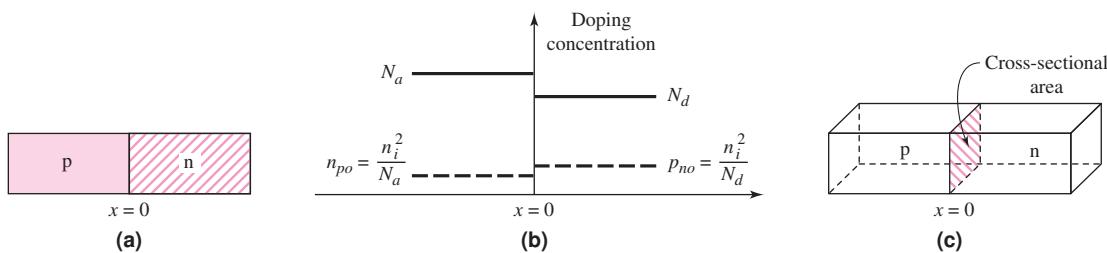


Figure 1.10 (a) The pn junction: (a) simplified one-dimensional geometry, (b) doping profile of an ideal uniformly doped pn junction, and (c) three-dimensional representation showing the cross-sectional area

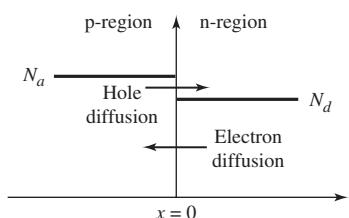


Figure 1.11 Initial diffusion of electrons and holes across the metallurgical junction at the “instant in time” that the p- and n-regions are joined together

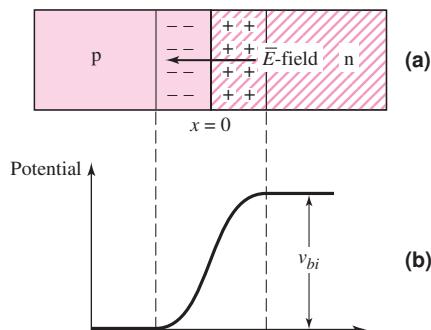


Figure 1.12 The pn junction in thermal equilibrium. (a) The space charge region with negatively charged acceptor ions in the p-region and positively charged donor ions in the n-region; the resulting electric field from the n- to the p-region. (b) The potential through the junction and the built-in potential barrier V_{bi} across the junction.

electrons from the n-region uncovers positively charged donor ions. This action creates a charge separation (Figure 1.12(a)), which sets up an electric field oriented in the direction from the positive charge to the negative charge.

If no voltage is applied to the pn junction, the diffusion of holes and electrons must eventually cease. The direction of the induced electric field will cause the resulting force to repel the diffusion of holes from the p-region and the diffusion of electrons from the n-region. Thermal equilibrium occurs when the force produced by the electric field and the “force” produced by the density gradient exactly balance.

The positively charged region and the negatively charged region comprise the **space-charge region**, or **depletion region**, of the pn junction, in which there are essentially no mobile electrons or holes. Because of the electric field in the space-charge region, there is a potential difference across that region (Figure 1.12(b)). This potential difference is called the **built-in potential barrier**, or built-in voltage, and is given by

$$V_{bi} = \frac{kT}{e} \ln\left(\frac{N_a N_d}{n_i^2}\right) = V_T \ln\left(\frac{N_a N_d}{n_i^2}\right) \quad (1.16)$$

where $V_T \equiv kT/e$, k = Boltzmann’s constant, T = absolute temperature, e = the magnitude of the electronic charge, and N_a and N_d are the net acceptor and donor concentrations in the p- and n-regions, respectively. The parameter V_T is called the **thermal voltage** and is approximately $V_T = 0.026$ V at room temperature, $T = 300$ K.

EXAMPLE 1.5

Objective: Calculate the built-in potential barrier of a pn junction.

Consider a silicon pn junction at $T = 300$ K, doped at $N_a = 10^{16}$ cm $^{-3}$ in the p-region and $N_d = 10^{17}$ cm $^{-3}$ in the n-region.

Solution: From the results of Example 1.1, we have $n_i = 1.5 \times 10^{10}$ cm $^{-3}$ for silicon at room temperature. We then find

$$V_{bi} = V_T \ln\left(\frac{N_a N_d}{n_i^2}\right) = (0.026) \ln\left[\frac{(10^{16})(10^{17})}{(1.5 \times 10^{10})^2}\right] = 0.757 \text{ V}$$

Comment: Because of the log function, the magnitude of V_{bi} is not a strong function of the doping concentrations. Therefore, the value of V_{bi} for silicon pn junctions is usually within 0.1 to 0.2 V of this calculated value.

EXERCISE PROBLEM

Ex 1.5: (a) Calculate V_{bi} for a GaAs pn junction at $T = 300$ K for $N_a = 10^{16} \text{ cm}^{-3}$ and $N_d = 10^{17} \text{ cm}^{-3}$ (b) Repeat part (a) for a Germanium pn junction with the same doping concentrations. (Ans. (a) $V_{bi} = 1.23$ V, (b) $V_{bi} = 0.374$ V).

The potential difference, or built-in potential barrier, across the space-charge region cannot be measured by a voltmeter because new potential barriers form between the probes of the voltmeter and the semiconductor, canceling the effects of V_{bi} . In essence, V_{bi} maintains equilibrium, so no current is produced by this voltage. However, the magnitude of V_{bi} becomes important when we apply a forward-bias voltage, as discussed later in this chapter.

1.2.2 Reverse-Biased pn Junction

Assume a positive voltage is applied to the n-region of a pn junction, as shown in Figure 1.13. The applied voltage V_R induces an applied electric field, E_A , in the semiconductor. The direction of this applied field is the same as that of the E -field in the space-charge region. The magnitude of the electric field *in* the space-charge region increases above the thermal equilibrium value. This increased electric field holds back the holes in the p-region and the electrons in the n-region, so there is essentially no current across the pn junction. By definition, this applied voltage polarity is called **reverse bias**.

When the electric field in the space-charge region increases, the number of positive and negative charges must increase. If the doping concentrations are not changed, the increase in the fixed charge can only occur if the width W of the space-charge region increases. Therefore, with an increasing reverse-bias voltage V_R , space-charge width W also increases. This effect is shown in Figure 1.14.

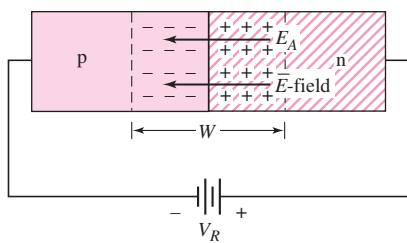


Figure 1.13 A pn junction with an applied reverse-bias voltage, showing the direction of the electric field induced by V_R and the direction of the original space-charge electric field. Both electric fields are in the same direction, resulting in a larger net electric field and a larger barrier between the p- and n-regions.

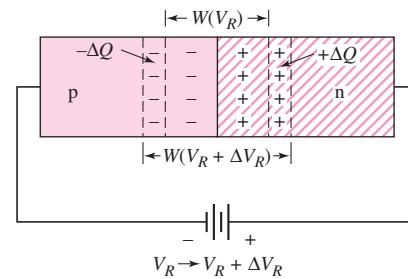


Figure 1.14 Increase in space-charge width with an increase in reverse bias voltage from V_R to $V_R + \Delta V_R$. Creation of additional charges $+ΔQ$ and $-ΔQ$ leads to a junction capacitance.

Because of the additional positive and negative charges induced in the space-charge region with an increase in reverse-bias voltage, a capacitance is associated with the pn junction when a reverse-bias voltage is applied. This **junction capacitance**, or depletion layer capacitance, can be written in the form

$$C_j = C_{jo} \left(1 + \frac{V_R}{V_{bi}} \right)^{-1/2} \quad (1.17)$$

where C_{jo} is the junction capacitance at zero applied voltage.

The junction capacitance will affect the switching characteristics of the pn junction, as we will see later in the chapter. The voltage across a capacitance cannot change instantaneously, so changes in voltages in circuits containing pn junctions will not occur instantaneously.

The capacitance–voltage characteristics can make the pn junction useful for electrically tunable resonant circuits. Junctions fabricated specifically for this purpose are called **varactor diodes**. Varactor diodes can be used in electrically tunable oscillators, such as a Hartley oscillator, discussed in Chapter 15, or in tuned amplifiers, considered in Chapter 8.

EXAMPLE 1.6

Objective: Calculate the junction capacitance of a pn junction.

Consider a silicon pn junction at $T = 300$ K, with doping concentrations of $N_a = 10^{16} \text{ cm}^{-3}$ and $N_d = 10^{15} \text{ cm}^{-3}$. Assume that $n_i = 1.5 \times 10^{10} \text{ cm}^{-3}$ and let $C_{jo} = 0.5 \text{ pF}$. Calculate the junction capacitance at $V_R = 1 \text{ V}$ and $V_R = 5 \text{ V}$.

Solution: The built-in potential is determined by

$$V_{bi} = V_T \ln \left(\frac{N_a N_d}{n_i^2} \right) = (0.026) \ln \left[\frac{(10^{16})(10^{15})}{(1.5 \times 10^{10})^2} \right] = 0.637 \text{ V}$$

The junction capacitance for $V_R = 1 \text{ V}$ is then found to be

$$C_j = C_{jo} \left(1 + \frac{V_R}{V_{bi}} \right)^{-1/2} = (0.5) \left(1 + \frac{1}{0.637} \right)^{-1/2} = 0.312 \text{ pF}$$

For $V_R = 5 \text{ V}$

$$C_j = (0.5) \left(1 + \frac{5}{0.637} \right)^{-1/2} = 0.168 \text{ pF}$$

Comment: The magnitude of the junction capacitance is usually at or below the picofarad range, and it decreases as the reverse-bias voltage increases.

EXERCISE PROBLEM

Ex 1.6: A silicon pn junction at $T = 300$ K is doped at $N_d = 10^{16} \text{ cm}^{-3}$ and $N_a = 10^{17} \text{ cm}^{-3}$. The junction capacitance is to be $C_j = 0.8 \text{ pF}$ when a reverse-bias voltage of $V_R = 5 \text{ V}$ is applied. Find the zero-biased junction capacitance C_{jo} . (Ans. $C_{jo} = 2.21 \text{ pF}$)

As implied in the previous section, the magnitude of the electric field in the space-charge region increases as the reverse-bias voltage increases, and the maximum

electric field occurs at the metallurgical junction. However, neither the electric field in the space-charge region nor the applied reverse-bias voltage can increase indefinitely because at some point, breakdown will occur and a large reverse bias current will be generated. This concept will be described in detail later in this chapter.

1.2.3 Forward-Biased pn Junction

We have seen that the n-region contains many more free electrons than the p-region; similarly, the p-region contains many more holes than the n-region. With zero applied voltage, the built-in potential barrier prevents these majority carriers from diffusing across the space-charge region; thus, the barrier maintains equilibrium between the carrier distributions on either side of the pn junction.

If a positive voltage v_D is applied to the p-region, the potential barrier decreases (Figure 1.15). The electric fields in the space-charge region are very large compared to those in the remainder of the p- and n-regions, so essentially all of the applied voltage exists across the pn junction region. The applied electric field, E_A , induced by the applied voltage is in the opposite direction from that of the thermal equilibrium space-charge E-field. However, the net electric field is *always* from the n- to the p-region. The net result is that the electric field in the space-charge region is lower than the equilibrium value. This upsets the delicate balance between diffusion and the E-field force. Majority carrier electrons from the n-region diffuse into the p-region, and majority carrier holes from the p-region diffuse into the n-region. The process continues as long as the voltage v_D is applied, thus creating a current in the pn junction. This process would be analogous to lowering a dam wall slightly. A slight drop in the wall height can send a large amount of water (current) over the barrier.

This applied voltage polarity (i.e., bias) is known as **forward bias**. The forward-bias voltage v_D must always be less than the built-in potential barrier V_{bi} .

As the majority carriers cross into the opposite regions, they become minority carriers in those regions, causing the minority carrier concentrations to increase. Figure 1.16 shows the resulting excess minority carrier concentrations at the space-charge region edges. These excess minority carriers diffuse into the neutral n- and p-regions, where they recombine with majority carriers, thus establishing a steady-state condition, as shown in Figure 1.16.

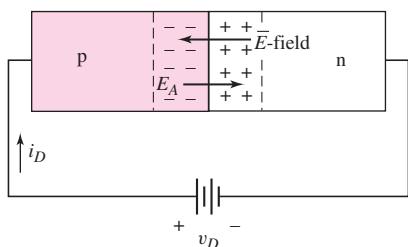


Figure 1.15 A pn junction with an applied forward-bias voltage showing the direction of the electric field induced by v_D and the direction of the original space-charge electric field. The two electric fields are in opposite directions resulting in a smaller net electric field and a smaller barrier between the p- and n-regions. The net electric field is always from the n- to the p-region.

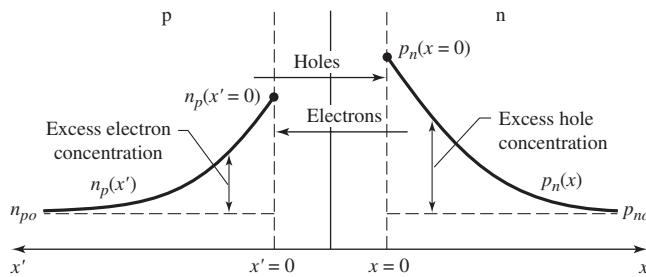


Figure 1.16 Steady-state minority carrier concentrations in a pn junction under forward bias. The gradients in the minority carrier concentrations generate diffusion currents in the device.

1.2.4 Ideal Current–Voltage Relationship

As shown in Figure 1.16, an applied voltage results in a gradient in the minority carrier concentrations, which in turn causes diffusion currents. The theoretical relationship between the voltage and the current in the pn junction is given by

$$i_D = I_S \left[e^{\left(\frac{v_D}{nV_T} \right)} - 1 \right] \quad (1.18)$$

The parameter I_S is the **reverse-bias saturation current**. For silicon pn junctions, typical values of I_S are in the range of 10^{-18} to 10^{-12} A. The actual value depends on the doping concentrations and is also proportional to the cross-sectional area of the junction. The parameter V_T is the thermal voltage, as defined in Equation (1.16), and is approximately $V_T = 0.026$ V at room temperature. The parameter n is usually called the emission coefficient or ideality factor, and its value is in the range $1 \leq n \leq 2$.

The emission coefficient n takes into account any recombination of electrons and holes in the space-charge region. At very low current levels, recombination may be a significant factor and the value of n may be close to 2. At higher current levels, recombination is less a factor, and the value of n will be 1. Unless otherwise stated, we will assume the emission coefficient is $n = 1$.

This pn junction, with nonlinear rectifying current characteristics, is called a **pn junction diode**.

EXAMPLE 1.7

Objective: Determine the current in a pn junction diode.

Consider a pn junction at $T = 300$ K in which $I_S = 10^{-14}$ A and $n = 1$. Find the diode current for $v_D = +0.70$ V and $v_D = -0.70$ V.

Solution: For $v_D = +0.70$ V, the pn junction is forward-biased and we find

$$i_D = I_S \left[e^{\left(\frac{v_D}{V_T} \right)} - 1 \right] = (10^{-14}) \left[e^{\left(\frac{+0.70}{0.026} \right)} - 1 \right] \Rightarrow 4.93 \text{ mA}$$

For $v_D = -0.70$ V, the pn junction is reverse-biased and we find

$$i_D = I_S \left[e^{\left(\frac{v_D}{V_T} \right)} - 1 \right] = (10^{-14}) \left[e^{\left(\frac{-0.70}{0.026} \right)} - 1 \right] \cong -10^{-14} \text{ A}$$

Comment: Although I_S is quite small, even a relatively small value of forward-bias voltage can induce a moderate junction current. With a reverse-bias voltage applied, the junction current is virtually zero.

EXERCISE PROBLEM

Ex 1.7: (a) A silicon pn junction at $T = 300$ K has a reverse-saturation current of $I_S = 2 \times 10^{-14}$ A. Determine the required forward-bias voltage to produce a current of (i) $I_D = 50 \mu\text{A}$ and (ii) $I_D = 1 \text{ mA}$. (b) Repeat part (a) for $I_S = 2 \times 10^{-12}$ A. (Ans. (a) (i) 0.563 V, (ii) 0.641 V; (b) (i) 0.443 V, (ii) 0.521 V).

1.2.5 pn Junction Diode

Figure 1.17 is a plot of the derived current–voltage characteristics of a pn junction. For a forward-bias voltage, the current is an exponential function of voltage. Figure 1.18 depicts the forward-bias current plotted on a log scale. With only a small change in the forward-bias voltage, the corresponding forward-bias current increases by orders of magnitude. For a forward-bias voltage $v_D > +0.1$ V, the (-1) term in Equation (1.18) can be neglected. In the reverse-bias direction, the current is almost zero.

Figure 1.19 shows the diode circuit symbol and the conventional current direction and voltage polarity. The diode can be thought of and used as a voltage controlled switch that is “off” for a reverse-bias voltage and “on” for a forward-bias voltage. In the forward-bias or “on” state, a relatively large current is produced by a fairly small applied voltage; in the reverse-bias, or “off” state, only a very small current is created.

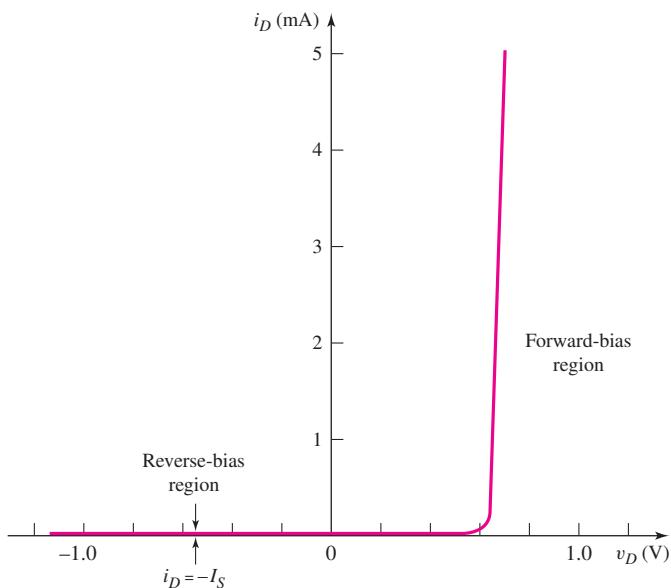


Figure 1.17 Ideal I - V characteristics of a pn junction diode for $I_S = 10^{-14}$ A. The diode current is an exponential function of diode voltage in the forward-bias region and is very nearly zero in the reverse-bias region. The pn junction diode is a nonlinear electronic device.

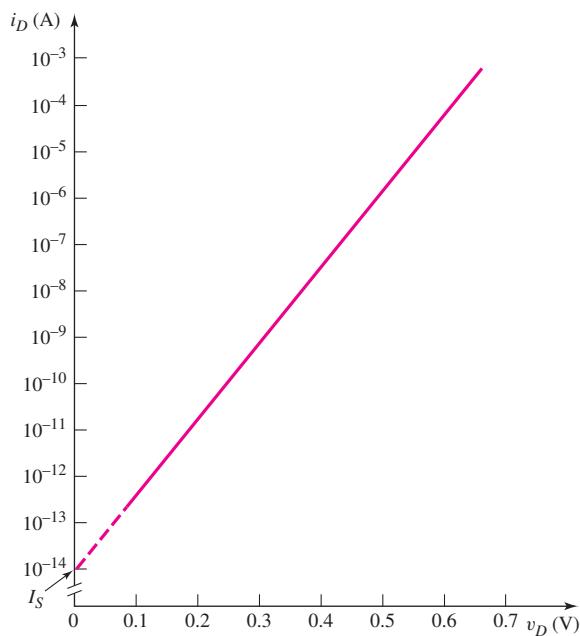


Figure 1.18 Ideal forward-biased I - V characteristics of a pn junction diode, with the current plotted on a log scale for $I_S = 10^{-14}$ A and $n = 1$. The diode current increases approximately one order of magnitude for every 60-mV increase in diode voltage.

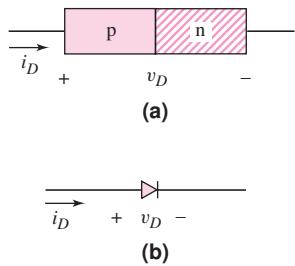


Figure 1.19 The basic pn junction diode: (a) simplified geometry and (b) circuit symbol, and conventional current direction and voltage polarity

When a diode is reverse-biased by at least 0.1 V, the diode current is $i_D = -I_S$. The current is in the reverse direction and is a constant, hence the name reverse-bias saturation current. Real diodes, however, exhibit reverse-bias currents that are considerably larger than I_S . This additional current is called a generation current and is due to electrons and holes being generated within the space-charge region. Whereas a typical value of I_S may be 10^{-14} A, a typical value of reverse-bias current may be 10^{-9} A or 1 nA. Even though this current is much larger than I_S , it is still small and negligible in most cases.

Temperature Effects

Since both I_S and V_T are functions of temperature, the diode characteristics also vary with temperature. The temperature-related variations in forward-bias characteristics are illustrated in Figure 1.20. For a given current, the required forward-bias voltage decreases as temperature increases. For silicon diodes, the change is approximately $2 \text{ mV}/^\circ\text{C}$.

The parameter I_S is a function of the intrinsic carrier concentration n_i , which in turn is strongly dependent on temperature. Consequently, the value of I_S approximately doubles for every 5°C increase in temperature. The actual reverse-bias diode current, as a general rule, doubles for every 10°C rise in temperature. As an example of the importance of this effect, the relative value of n_i in germanium, is large, resulting in a large reverse-saturation current in germanium-based diodes. Increases in this reverse current with increases in the temperature make the germanium diode highly impractical for most circuit applications.

Breakdown Voltage

When a reverse-bias voltage is applied to a pn junction, the electric field in the space-charge region increases. The electric field may become large enough that covalent

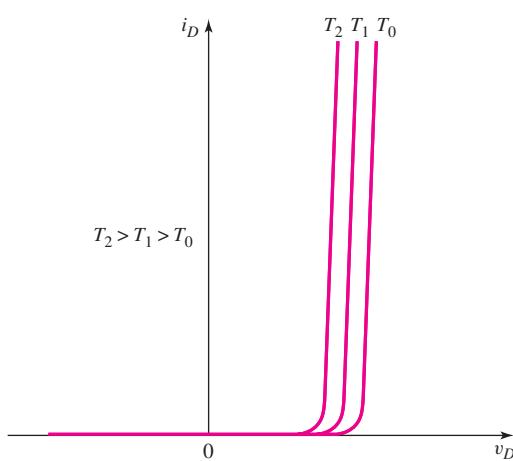


Figure 1.20 Forward-biased pn junction characteristics versus temperature. The required diode voltage to produce a given current decreases with an increase in temperature.

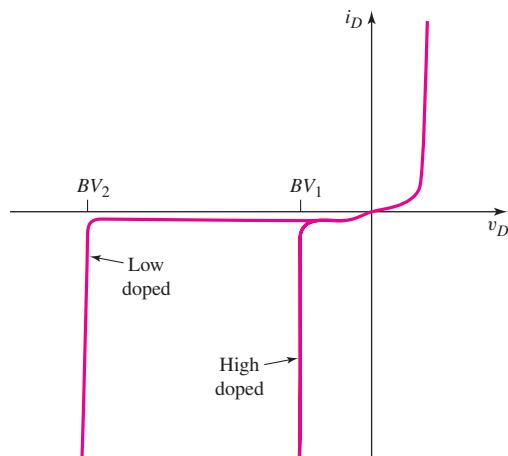


Figure 1.21 Reverse-biased diode characteristics showing breakdown for a low-doped pn junction and a high-doped pn junction. The reverse-bias current increases rapidly once breakdown has occurred.

bonds are broken and electron–hole pairs are created. Electrons are swept into the n-region and holes are swept into the p-region by the electric field, generating a large reverse bias current. This phenomenon is called **breakdown**. The reverse-bias current created by the breakdown mechanism is limited only by the external circuit. If the current is not sufficiently limited, a large power can be dissipated in the junction that may damage the device and cause burnout. The current–voltage characteristic of a diode in breakdown is shown in Figure 1.21.

The most common breakdown mechanism is called **avalanche breakdown**, which occurs when carriers crossing the space charge region gain sufficient kinetic energy from the high electric field to be able to break covalent bonds during a collision process. The basic avalanche multiplication process is demonstrated in Figure 1.22. The generated electron–hole pairs can themselves be involved in a collision process generating additional electron–hole pairs, thus the avalanche process. The breakdown voltage is a function of the doping concentrations in the

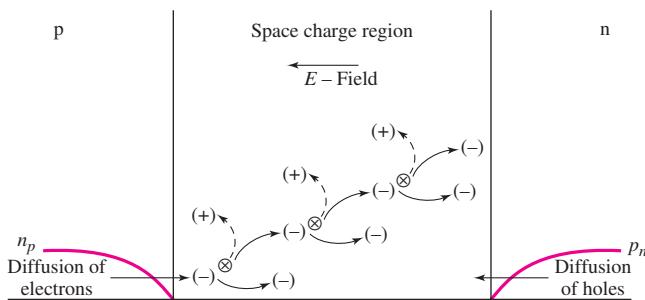


Figure 1.22 The avalanche multiplication process in the space charge region. Shown are the collisions of electrons creating additional electron–hole pairs. Holes can also be involved in collisions creating additional electron–hole pairs.

n- and p-regions of the junction. Larger doping concentrations result in smaller breakdown voltages.

A second breakdown mechanism is called **Zener breakdown** and is a result of tunneling of carriers across the junction. This effect is prominent at very high doping concentrations and results in breakdown voltages less than 5 V.

The voltage at which breakdown occurs depends on fabrication parameters of the pn junction, but is usually in the range of 50 to 200 V for discrete devices, although breakdown voltages outside this range are possible—in excess of 1000 V, for example. A pn junction is usually rated in terms of its **peak inverse voltage** or **PIV**. The PIV of a diode must never be exceeded in circuit operation if reverse breakdown is to be avoided.

Diodes can be fabricated with a specifically designed breakdown voltage and are designed to operate in the breakdown region. These diodes are called Zener diodes and are discussed later in this chapter as well as in the next chapter.

Switching Transient

Since the pn junction diode can be used as an electrical switch, an important parameter is its transient response, that is, its speed and characteristics, as it is switched from one state to the other. Assume, for example, that the diode is switched from the forward-bias “on” state to the reverse-bias “off” state. Figure 1.23 shows a simple circuit that will switch the applied voltage at time $t = 0$. For $t < 0$, the forward-bias current i_D is

$$i_D = I_F = \frac{V_F - v_D}{R_F} \quad (1.19)$$

The minority carrier concentrations for an applied forward-bias voltage and an applied reverse-bias voltage are shown in Figure 1.24. Here, we neglect the change in the space charge region width. When a forward-bias voltage is applied, excess minority carrier charge is stored in both the p- and n-regions. The excess charge is the difference between the minority carrier concentrations for a forward-bias voltage and those for a reverse-bias voltage as indicated in the figure. This charge must be removed when the diode is switched from the forward to the reverse bias.

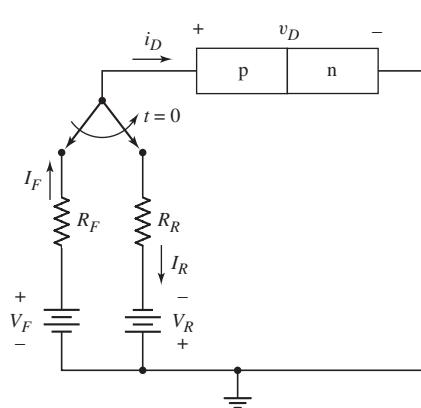


Figure 1.23 Simple circuit for switching a diode from forward to reverse bias

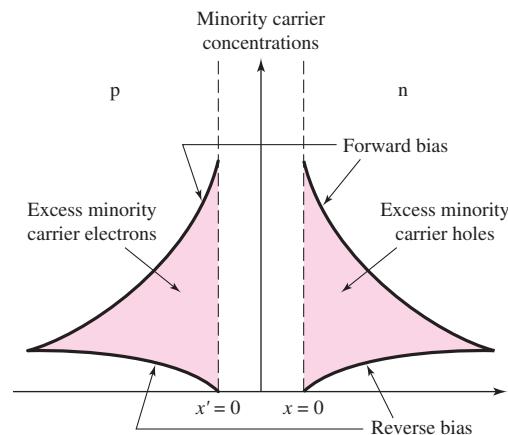


Figure 1.24 Stored excess minority carrier charge under forward bias compared to reverse bias. This charge must be removed as the diode is switched from forward to reverse bias.

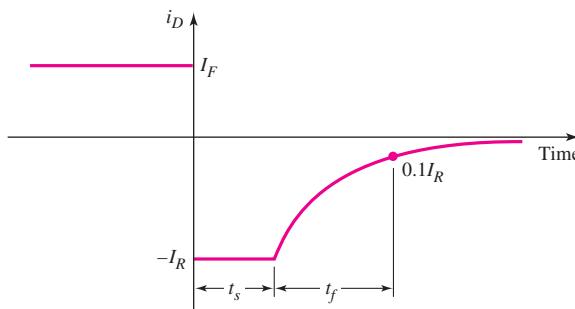


Figure 1.25 Current characteristics versus time during diode switching

As the forward-bias voltage is removed, relatively large diffusion currents are created in the reverse-bias direction. This happens because the excess minority carrier electrons flow back across the junction into the n-region, and the excess minority carrier holes flow back across the junction into the p-region.

The large reverse-bias current is initially limited by resistor R_R to approximately

$$i_D = -I_R \cong \frac{-V_R}{R_R} \quad (1.20)$$

The junction capacitances do not allow the junction voltage to change instantaneously. The reverse current I_R is approximately constant for $0^+ < t < t_s$, where t_s is the **storage time**, which is the length of time required for the minority carrier concentrations at the space-charge region edges to reach the thermal equilibrium values. After this time, the voltage across the junction begins to change. The fall time t_f is typically defined as the time required for the current to fall to 10 percent of its initial value. The total **turn-off time** is the sum of the storage time and the fall time. Figure 1.25 shows the current characteristics as this entire process takes place.

In order to switch a diode quickly, the diode must have a small excess minority carrier lifetime, and we must be able to produce a large reverse current pulse. Therefore, in the design of diode circuits, we must provide a path for the transient reverse-bias current pulse. These same transient effects impact the switching of transistors. For example, the switching speed of transistors in digital circuits will affect the speed of computers.

The turn-on transient occurs when the diode is switched from the “off” state to the forward-bias “on” state, which can be initiated by applying a forward-bias current pulse. The transient **turn-on time** is the time required to establish the forward-bias minority carrier distributions. During this time, the voltage across the junction gradually increases toward its steady-state value. Although the turn-on time for the pn junction diode is not zero, it is usually less than the transient turn-off time.

Test Your Understanding

TYU 1.6 (a) Determine V_{bi} for a silicon pn junction at $T = 300$ K for $N_a = 10^{15} \text{ cm}^{-3}$ and $N_d = 5 \times 10^{16} \text{ cm}^{-3}$. (b) Repeat part (a) for a GaAs pn junction. (c) Repeat part (a) for a Ge pn junction. (Ans. (a) 0.679 V, (b) 1.15 V, (c) 0.296 V).

TYU 1.7 A silicon pn junction diode at $T = 300$ K has a reverse-saturation current of $I_S = 10^{-16} \text{ A}$. (a) Determine the forward-bias diode current for (i) $V_D = 0.55$ V,

(ii) $V_D = 0.65$ V, and (iii) $V_D = 0.75$ V, (b) Find the reverse-bias diode current for (i) $V_D = -0.55$ V and (ii) $V_D = -2.5$ V. (Ans. (a) (i) $0.154 \mu\text{A}$, (ii) $7.20 \mu\text{A}$, (iii) 0.337 mA ; (b) (i) -10^{-16} A , (ii) -10^{-16} A).

TYU 1.8 Recall that the forward-bias diode voltage decreases approximately by $2 \text{ mV}/^\circ\text{C}$ for silicon diodes with a given current. If $V_D = 0.650$ V at $I_D = 1 \text{ mA}$ for a temperature of 25°C , determine the diode voltage at $I_D = 1 \text{ mA}$ for $T = 125^\circ\text{C}$. (Ans. $V_D = 0.450$ V)



1.3 DIODE CIRCUITS: DC ANALYSIS AND MODELS

Objective: • Examine dc analysis techniques for diode circuits using various models to describe the diode characteristics.

In this section, we begin to study the diode in various circuit configurations. As we have seen, the diode is a two-terminal device with nonlinear $i-v$ characteristics, as opposed to a two-terminal resistor, which has a linear relationship between current and voltage. The analysis of nonlinear electronic circuits is not as straightforward as the analysis of linear electric circuits. However, there are electronic functions that can be implemented only by nonlinear circuits. Examples include the generation of dc voltages from sinusoidal voltages and the implementation of logic functions.

Mathematical relationships, or **models**, that describe the current–voltage characteristics of electrical elements allow us to analyze and design circuits without having to fabricate and test them in the laboratory. An example is Ohm’s law, which describes the properties of a resistor. In this section, we will develop the dc analysis and modeling techniques of diode circuits.

This section considers the current–voltage characteristics of the pn junction diode in order to construct various circuit models. Large-signal models are initially developed that describe the behavior of the device with relatively large changes in voltages and currents. These models simplify the analysis of diode circuits and make the analysis of relatively complex circuits much easier. In the next section, we will consider a small-signal model of the diode that will describe the behavior of the pn junction with small changes in voltages and currents. It is important to understand the difference between large-signal and small-signal models and the conditions when they are used.

To begin to understand diode circuits, consider a simple diode application. The current–voltage characteristics of the pn junction diode were given in Figure 1.17. An **ideal diode** (as opposed to a diode with ideal $I-V$ characteristics) has the characteristics shown in Figure 1.26(a). When a reverse-bias voltage is applied, the current through the diode is zero (Figure 1.26(b)); when current through the diode is greater than zero, the voltage across the diode is zero (Figure 1.26(c)). An external circuit connected to the diode must be designed to control the forward current through the diode.

One diode circuit is the **rectifier** circuit shown in Figure 1.27(a). Assume that the input voltage v_I is a sinusoidal signal, as shown in Figure 1.27(b), and the diode is an ideal diode (see Figure 1.26(a)). During the positive half-cycle of the sinusoidal

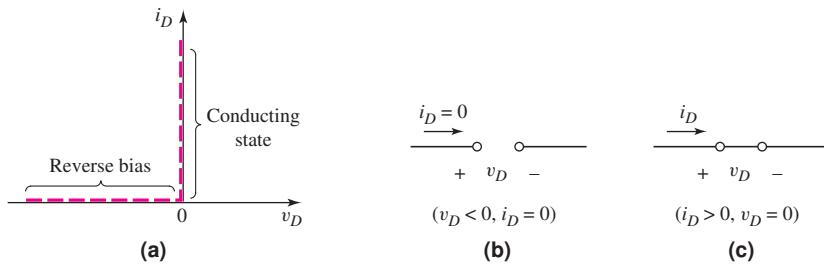


Figure 1.26 The ideal diode: (a) the I - V characteristics of the ideal diode, (b) equivalent circuit under reverse bias (an open circuit), and (c) equivalent circuit in the conducting state (a short circuit)

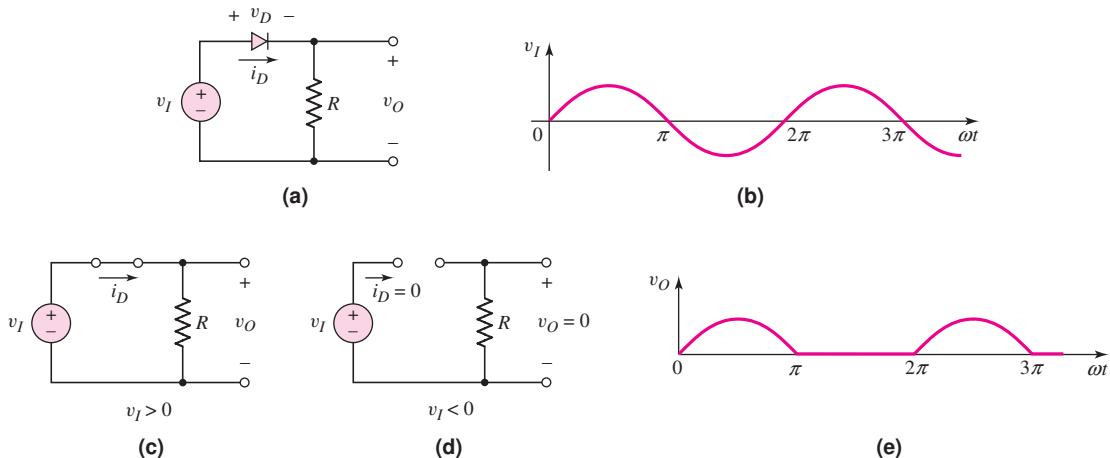


Figure 1.27 The diode rectifier: (a) circuit, (b) sinusoidal input signal, (c) equivalent circuit for $v_I > 0$, (d) equivalent circuit for $v_I < 0$, and (e) rectified output signal

input, a forward-bias current exists in the diode and the voltage across the diode is zero. The equivalent circuit for this condition is shown in Figure 1.27(c). The output voltage v_O is then equal to the input voltage. During the negative half-cycle of the sinusoidal input, the diode is reverse biased. The equivalent circuit for this condition is shown in Figure 1.27(d). In this part of the cycle, the diode acts as an open circuit, the current is zero, and the output voltage is zero. The output voltage of the circuit is shown in Figure 1.27(e).

Over the entire cycle, the input signal is sinusoidal and has a zero average value; however, the output signal contains only positive values and therefore has a positive average value. Consequently, this circuit is said to **rectify** the input signal, which is the first step in generating a dc voltage from a sinusoidal (ac) voltage. A dc voltage is required in virtually all electronic circuits.

As mentioned, the analysis of nonlinear circuits is not as straightforward as that of linear circuits. In this section, we will look at four approaches to the dc analysis of diode circuits: (a) iteration; (b) graphical techniques; (c) a piecewise linear modeling method; and (d) a computer analysis. Methods (a) and (b) are closely related and are therefore presented together.

1.3.1

Iteration and Graphical Analysis Techniques

Iteration means using trial and error to find a solution to a problem. The graphical analysis technique involves plotting two simultaneous equations and locating their point of intersection, which is the solution to the two equations. We will use both techniques to solve the circuit equations, which include the diode equation. These equations are difficult to solve by hand because they contain both linear and exponential terms.

Consider, for example, the circuit shown in Figure 1.28, with a dc voltage V_{PS} applied across a resistor and a diode. **Kirchhoff's voltage law** applies both to nonlinear and linear circuits, so we can write

$$V_{PS} = I_D R + V_D \quad (1.21(a))$$

which can be rewritten as

$$I_D = \frac{V_{PS}}{R} - \frac{V_D}{R} \quad (1.21(b))$$

[Note: In the remainder of this section in which dc analysis is emphasized, the dc variables are denoted by uppercase letters and uppercase subscripts.]

The diode voltage V_D and current I_D are related by the ideal diode equation as

$$I_D = I_S \left[e^{\left(\frac{V_D}{V_T} \right)} - 1 \right] \quad (1.22)$$

where I_S is assumed to be known for a particular diode.

Combining Equations (1.21(a)) and (1.22), we obtain

$$V_{PS} = I_S R \left[e^{\left(\frac{V_D}{V_T} \right)} - 1 \right] + V_D \quad (1.23)$$

which contains only one unknown, V_D . However, Equation (1.23) is a transcendental equation and cannot be solved directly. The use of iteration to find a solution to this equation is demonstrated in the following example.

EXAMPLE 1.8

Objective: Determine the diode voltage and current for the circuit shown in Figure 1.28.

Consider a diode with a given reverse-saturation current of $I_S = 10^{-13}$ A.

Solution: We can write Equation (1.23) as

$$5 = (10^{-13})(2 \times 10^3) \left[e^{\left(\frac{V_D}{0.026} \right)} - 1 \right] + V_D \quad (1.24)$$

If we first try $V_D = 0.60$ V, the right side of Equation (1.24) is 2.7 V, so the equation is not balanced and we must try again. If we next try $V_D = 0.65$ V, the right side of Equation (1.24) is 15.1 V. Again, the equation is not balanced, but we can see that the solution for V_D is between 0.6 and 0.65 V. If we continue refining our guesses, we will be able to show that, when $V_D = 0.619$ V, the right side of Equation (1.29) is 4.99 V, which is essentially equal to the value of the left side of the equation.

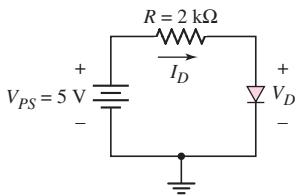


Figure 1.28 A simple diode circuit

The current in the circuit can then be determined by dividing the voltage difference across the resistor by the resistance, or

$$I_D = \frac{V_{PS} - V_D}{R} = \frac{5 - 0.619}{2} = 2.19 \text{ mA}$$

Comment: Once the diode voltage is known, the current can also be determined from the ideal diode equation. However, dividing the voltage difference across a resistor by the resistance is usually easier, and this approach is used extensively in the analysis of diode and transistor circuits.

EXERCISE PROBLEM

Ex 1.8: Consider the circuit in Figure 1.28. Let $V_{PS} = 4 \text{ V}$, $R = 4 \text{ k}\Omega$, and $I_S = 10^{-12} \text{ A}$. Determine V_D and I_D , using the ideal diode equation and the iteration method. (Ans. $V_D = 0.535 \text{ V}$, $I_D = 0.866 \text{ mA}$)

To use a graphical approach to analyze the circuit, we go back to Kirchhoff's voltage law, as expressed by Equation (1.21(a)), which was $V_{PS} = I_D R + V_D$. Solving for the current I_D , we have

$$I_D = \frac{V_{PS}}{R} - \frac{V_D}{R}$$

which was also given by Equation (1.21(b)). This equation gives a linear relation between the diode current I_D and the diode voltage V_D for a given power supply voltage V_{PS} and resistance R . This equation is referred to as the circuit **load line**, and is usually plotted on a graph with the current I_D as the vertical axis and the voltage V_D as the horizontal axis.

From Equation (1.21(b)), we see that if $I_D = 0$, then $V_D = V_{PS}$ which is the horizontal axis intercept. Also from this equation, if $V_D = 0$, then $I_D = V_{PS}/R$ which is the vertical axis intercept. The load line can be drawn between these two points. From Equation (1.21(b)), we see that the slope of the load line is $-1/R$.

Using the values given in Example (1.8), we can plot the straight line shown in Figure 1.29. The second plot in the figure is that of Equation (1.22), which is the ideal diode equation relating the diode current and voltage. The intersection of the load line and the device characteristics curve provides the dc current $I_D \approx 2.2 \text{ mA}$ through the diode and the dc voltage $V_D \approx 0.62 \text{ V}$ across the diode. This point is referred to as the **quiescent point**, or the **Q-point**.

The graphical analysis method can yield accurate results, but it is somewhat cumbersome. However, the concept of the load line and the graphical approach are useful for “visualizing” the response of a circuit, and the load line is used extensively in the evaluation of electronic circuits.

1.3.2 Piecewise Linear Model

Another, simpler way to analyze diode circuits is to *approximate* the diode's current–voltage characteristics, using linear relationships or straight lines. Figure 1.30, for example, shows the ideal current–voltage characteristics and two linear approximations.

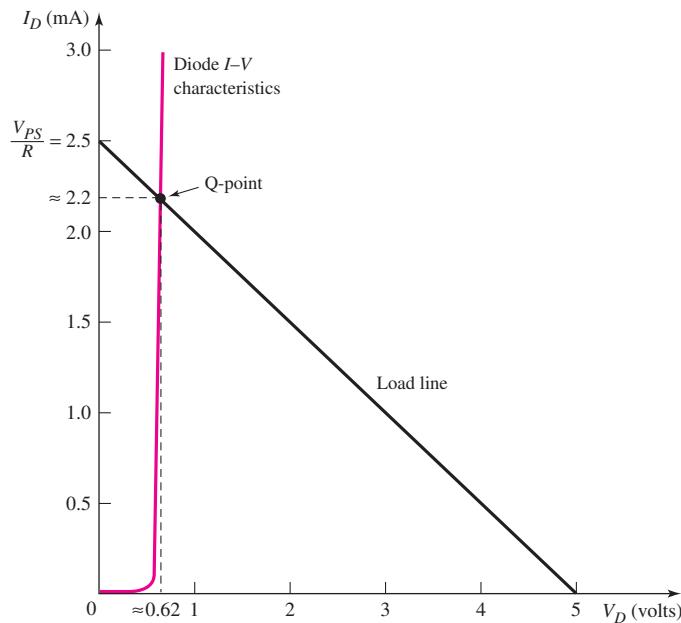


Figure 1.29 The diode and load line characteristics for the circuit shown in Figure 1.28

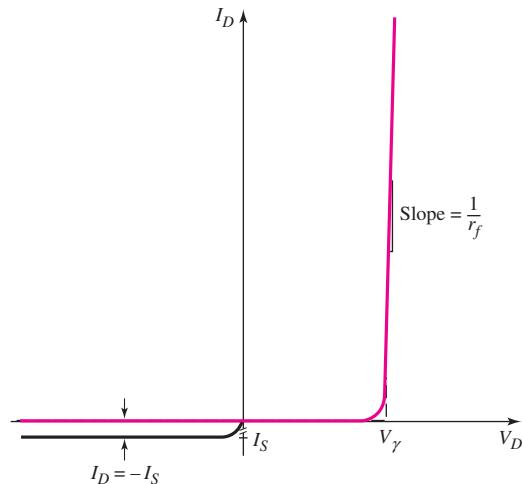


Figure 1.30 The diode \$I\$-\$V\$ characteristics and two linear approximations. The linear approximations form the piecewise linear model of the diode.

For $V_D \geq V_\gamma$, we assume a straight-line approximation whose slope is $1/r_f$, where V_γ is the **turn-on**, or **cut-in, voltage** of the diode, and r_f is the **forward diode resistance**. The equivalent circuit for this linear approximation is a constant-voltage source in series with a resistor (Figure 1.31(a)).² For $V_D < V_\gamma$, we assume a straight-line

²It is important to keep in mind that the voltage source in Figure 1.31(a) only represents a voltage drop for $V_D \geq V_\gamma$. When $V_D < V_\gamma$, the V_γ source does *not* produce a negative diode current. For $V_D < V_\gamma$, the equivalent circuit in Figure 1.31(b) must be used.

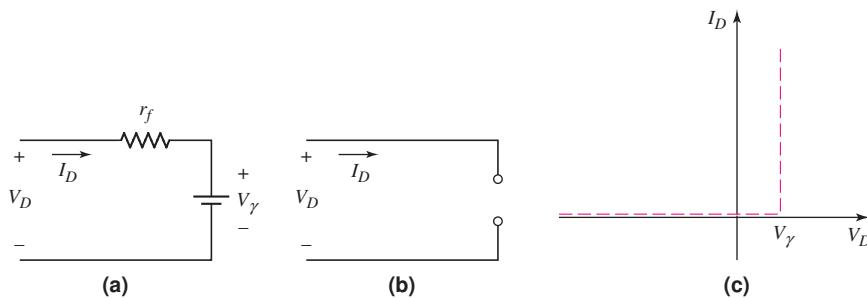


Figure 1.31 The diode piecewise equivalent circuit (a) in the “on” condition when $V_D \geq V_\gamma$, (b) in the “off” condition when $V_D < V_\gamma$, and (c) piecewise linear approximation when $r_f = 0$. When $r_f = 0$, the voltage across the diode is a constant at $V_D = V_\gamma$ when the diode is conducting.

approximation parallel with the V_D axis at the zero current level. In this case, the equivalent circuit is an open circuit (Figure 1.31(b)).

This method models the diode with segments of straight lines; thus the name **piecewise linear model**. If we assume $r_f = 0$, the piecewise linear diode characteristics are shown in Figure 1.31(c).

EXAMPLE 1.9

Objective: Determine the diode voltage and current in the circuit shown in Figure 1.28, using a piecewise linear model. Also determine the power **dissipated** in the diode.

Assume piecewise linear diode parameters of $V_\gamma = 0.6$ V and $r_f = 10 \Omega$.

Solution: With the given input voltage polarity, the diode is forward biased or “turned on,” so $I_D > 0$. The equivalent circuit is shown in Figure 1.31(a). The diode current is determined by

$$I_D = \frac{V_{PS} - V_\gamma}{R + r_f} = \frac{5 - 0.6}{2 \times 10^3 + 10} \Rightarrow 2.19 \text{ mA}$$

and the diode voltage is

$$V_D = V_\gamma + I_D r_f = 0.6 + (2.19 \times 10^{-3})(10) = 0.622 \text{ V}$$

The power dissipated in the diode is given by

$$P_D = I_D V_D$$

We then find

$$P_D = (2.19)(0.622) = 1.36 \text{ mW}$$

Comment: This solution, obtained using the piecewise linear model, is nearly equal to the solution obtained in Example 1.8, in which the ideal diode equation was used. However, the analysis using the piecewise-linear model in this example is by far easier than using the actual diode I - V characteristics as was done in Example 1.8. In general, we are willing to accept some slight analysis inaccuracy for ease of analysis.

EXERCISE PROBLEM

Ex 1.9: (a) Consider the circuit shown in Figure 1.28. Let $V_{PS} = 8\text{ V}$ and $V_\gamma = 0.7\text{ V}$. Assume $r_f = 0$. Calculate the value of R such that $I_D = 1.20\text{ mA}$. (b) If V_{PS} is reduced to 4 V and R is changed to $3.5\text{ k}\Omega$, determine the power dissipated in the diode. (Ans. (a) $6.08\text{ k}\Omega$, (b) 0.66 mW).

Because the forward diode resistance r_f in Example 1.9 is much smaller than the circuit resistance R , the diode current I_D is essentially independent of the value of r_f . In addition, if the cut-in voltage is 0.7 V instead of 0.6 V , the calculated diode current will be 2.15 mA , which is not significantly different from the previous results. Therefore, the calculated diode current is not a strong function of the cut-in voltage. Consequently, we will often assume a cut-in voltage of 0.7 V for silicon pn junction diodes.

The concept of the load line and the piecewise linear model can be combined in diode circuit analyses. From Kirchhoff's voltage law, the load line for the circuit shown in Figure 1.28 and for the piecewise linear model of the diode can be written as

$$V_{PS} = I_D R + V_\gamma$$

where V_γ is the diode cut-in voltage. We can assume $V_\gamma = 0.7\text{ V}$. Various load lines can be determined and plotted for the following circuit conditions:

- A: $V_{PS} = 5\text{ V}$, $R = 2\text{ k}\Omega$
- B: $V_{PS} = 5\text{ V}$, $R = 4\text{ k}\Omega$
- C: $V_{PS} = 2.5\text{ V}$, $R = 2\text{ k}\Omega$
- D: $V_{PS} = 2.5\text{ V}$, $R = 4\text{ k}\Omega$

The load line for condition A is plotted in Figure 1.32(a). Also plotted in the figure are the piecewise linear characteristics of the diode. The intersection of the two curves corresponds to the Q-point. For this case, the quiescent diode current is $I_{DQ} \cong 2.15\text{ mA}$.

Figure 1.32(b) shows the same piecewise linear characteristics of the diode. In addition, all four load lines, defined by the conditions listed above in A, B, C, and D are plotted on the figure. We see that the Q-point of the diode is a function of the load line. The Q-point changes for each load line.

The load line concept is also useful when the diode is reverse biased. Figure 1.33 (a) shows the same diode circuit as before, but with the direction of the diode reversed.

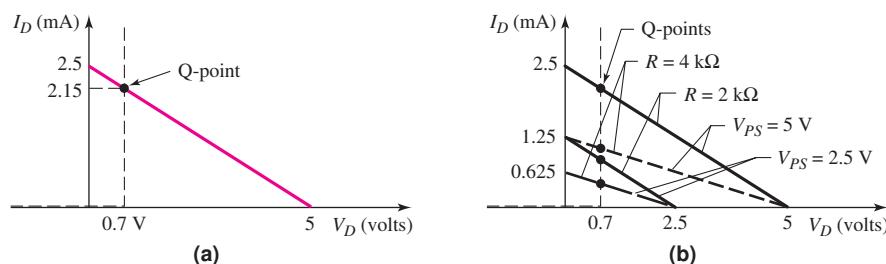


Figure 1.32 Piecewise linear diode approximation superimposed on (a) load line for $V_{PS} = 5\text{ V}$, $R = 2\text{ k}\Omega$ and (b) several load lines. The Q-point of the diode changes when the load line changes.

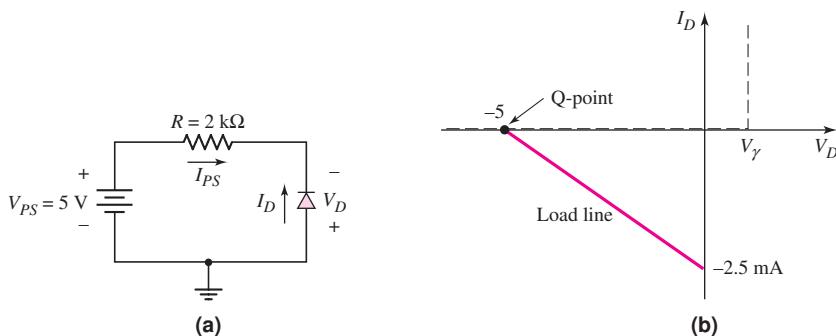


Figure 1.33 Reverse-biased diode (a) circuit and (b) piecewise linear approximation and load line

The diode current I_D and voltage V_D shown are the usual forward-biased parameters. Applying Kirchhoff's voltage law, we can write

$$V_{PS} = I_{PS}R - V_D = -I_DR - V_D \quad (1.25(a))$$

or

$$I_D = -\frac{V_{PS}}{R} - \frac{V_D}{R} \quad (1.25(b))$$

where $I_D = -I_{PS}$. Equation (1.25(b)) is the load line equation. The two end points are found by setting $I_D = 0$, which yields $V_D = -V_{PS} = -5$ V, and by setting $V_D = 0$, which yields $I_D = -V_{PS}/R = -5/2 = -2.5$ mA. The diode characteristics and the load line are plotted in Figure 1.33(b). We see that the load line is now in the third quadrant, where it intersects the diode characteristics curve at $V_D = -5$ V and $I_D = 0$, demonstrating that the diode is reverse biased.

Although the piecewise linear model may yield solutions that are less accurate than those obtained with the ideal diode equation, the analysis is much easier.

1.3.3 Computer Simulation and Analysis

Today's computers are capable of using detailed simulation models of various components and performing complex circuit analyses quickly and relatively easily. Such models can factor in many diverse conditions, such as the temperature dependence of various parameters. One of the earliest, and now the most widely used, circuit analysis programs is the *simulation program with integrated circuit emphasis* (SPICE). This program, developed at the University of California at Berkeley, was first released about 1973, and has been continuously refined since that time. One outgrowth of SPICE is PSpice, which is designed for use on personal computers.

EXAMPLE 1.10

Objective: Determine the diode current and voltage characteristics of the circuit shown in Figure 1.28 using a PSpice analysis.

Solution: Figure 1.34(a) is the PSpice circuit schematic diagram. A standard 1N4002 diode from the PSpice library was used in the analysis. The input voltage V_1 was varied (dc sweep) from 0 to 5 V. Figure 1.34(b) and (c) shows the diode voltage and diode current characteristics versus the input voltage.

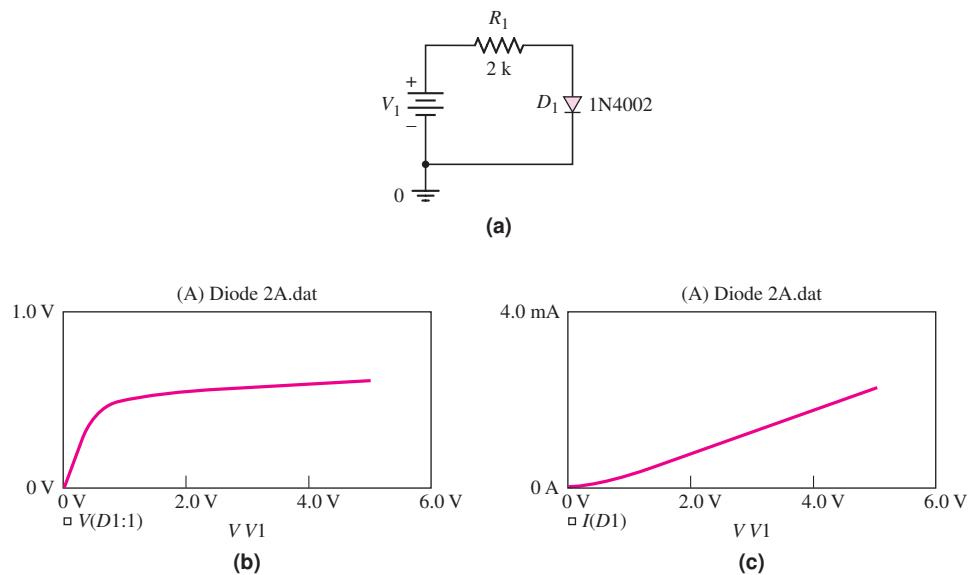


Figure 1.34 (a) PSpice circuit schematic, (b) diode voltage, and (c) diode current for Example 1.10

Discussion: Several observations may be made from the results. The diode voltage increases at almost a linear rate up to approximately 400 mV without any discernible (mA) current being measured. For an input voltage greater than approximately 500 mV, the diode voltage increases gradually to a value of about 610 mV at the maximum input voltage. The current also increases to a maximum value of approximately 2.2 mA at the maximum input voltage. The piecewise linear model predicts quite accurate results at the maximum input voltage. However, these results show that there is definitely a non-linear relation between the diode current and diode voltage. We must keep in mind that the piecewise linear model is an approximation technique that works very well in many applications.

EXERCISE PROBLEM

Ex 1.10: The resistor parameter in the circuit shown in Figure 1.28 is changed to $R = 20 \text{ k}\Omega$. Using a PSpice analysis, plot the diode current I_D and diode voltage V_D versus the power supply voltage V_{PS} over the range $0 \leq V_{PS} \leq 10 \text{ V}$.

1.3.4 Summary of Diode Models

The two dc diode models used in the hand analysis of diode circuits are: the ideal diode equation and the piecewise linear approximation. For the ideal diode equation, the reverse-saturation current I_S must be specified. For the piecewise linear model, the cut-in voltage V_y and forward diode resistance r_f must be specified. In most cases, however, r_f is assumed to be zero unless otherwise given.

The particular model that should be used in a specific application or situation is a compromise between accuracy and ease of calculation. This decision comes with experience. In general, a simple model can be used in an initial design for ease of

calculation. In a final design, we may want to use a computer simulation for better accuracy. However, it is very important to understand that the diode model or diode parameters used in the computer simulation must correspond to the actual diode parameters used in the circuit to ensure that the results are meaningful.

Test Your Understanding

TYU 1.9 Consider the diode and circuit in Exercise EX 1.8. Determine V_D and I_D , using the graphical technique. (Ans. $V_D \cong 0.54$ V, $I_D \cong 0.87$ mA)

TYU 1.10 Consider the circuit in Figure 1.28. Let $R = 4\text{ k}\Omega$ and $V_\gamma = 0.7$ V. Determine I_D for (a) $V_{PS} = 0.5$ V, (b) $V_{PS} = 2$ V, (c) $V_{PS} = 5$ V, (d) $V_{PS} = -1$ V, and (e) $V_{PS} = -5$ V. (Ans. (a) 0, (b) 0.325 mA, (c) 1.075 mA, (d) 0, (e) 0).

TYU 1.11 The power supply (input) voltage in the circuit of Figure 1.28 is $V_{PS} = 10$ V and the diode cut-in voltage is $V_\gamma = 0.7$ V (assume $r_f = 0$). The power dissipated in the diode is to be no more than 1.05 mW. Determine the maximum diode current and the minimum value of R to meet the power specification. (Ans. $I_D = 1.5$ mA, $R = 6.2\text{ k}\Omega$)



1.4

DIODE CIRCUITS: AC EQUIVALENT CIRCUIT

Objective: • Develop an equivalent circuit for a diode that is used when a small, time-varying signal is applied to a diode circuit.

Up to this point, we have only looked at the dc characteristics of the pn junction diode. When semiconductor devices with pn junctions are used in linear amplifier circuits, the time-varying, or ac, characteristics of the pn junction become important, because sinusoidal signals may be superimposed on the dc currents and voltages. The following sections examine these ac characteristics.

1.4.1

Sinusoidal Analysis

In the circuit shown in Figure 1.35(a), the voltage source v_i is assumed to be a sinusoidal, or time-varying, signal. The total input voltage v_I is composed of a dc component V_{PS} and an ac component v_i superimposed on the dc value. To investigate this circuit, we will look at two types of analyses: a dc analysis involving only the dc voltages and currents, and an ac analysis involving only the ac voltages and currents.

Current-Voltage Relationships

Since the input voltage contains a dc component with an ac signal superimposed, the diode current will also contain a dc component with an ac signal superimposed, as shown in Figure 1.35(b). Here, I_{DQ} is the dc quiescent diode current. In addition, the diode voltage will contain a dc value with an ac signal superimposed, as shown in Figure 1.35(c). For this analysis, assume that the ac signal is small compared to the dc component, so that a linear ac model can be developed from the nonlinear diode.

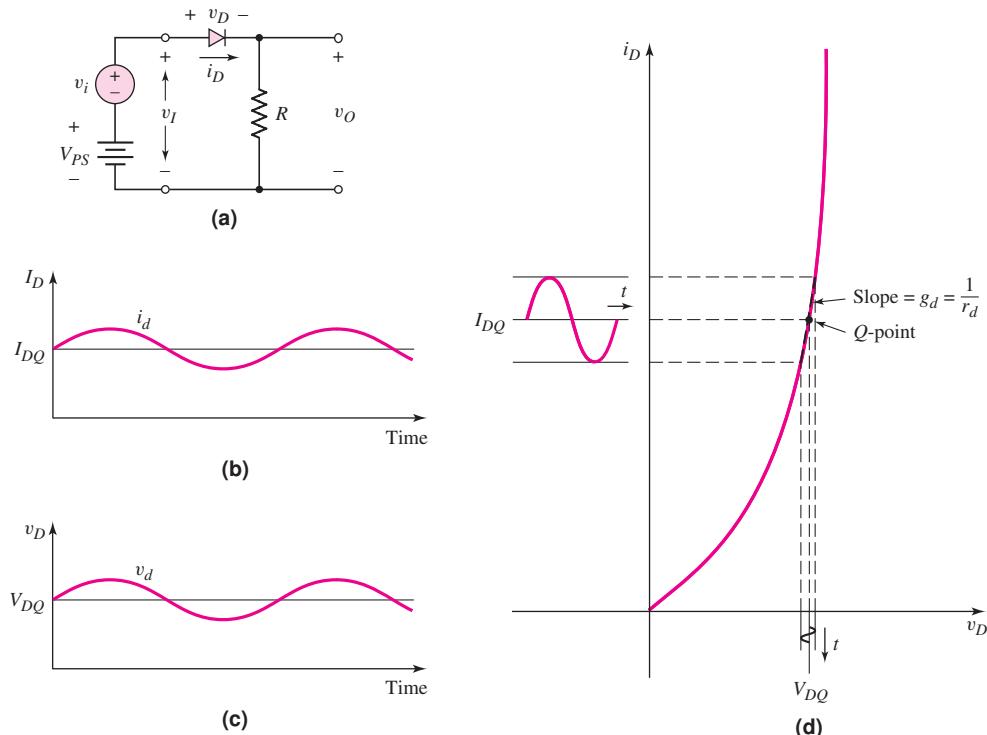


Figure 1.35 AC circuit analysis: (a) circuit with combined dc and sinusoidal input voltages, (b) sinusoidal diode current superimposed on the quiescent current, (c) sinusoidal diode voltage superimposed on the quiescent value, and (d) forward-biased diode I - V characteristics with a sinusoidal current and voltage superimposed on the quiescent values

The relationship between the diode current and voltage can be written as

$$i_D \cong I_S e^{\left(\frac{v_D}{V_T}\right)} = I_S e^{\left(\frac{V_{DQ}+v_d}{V_T}\right)} \quad (1.26)$$

where V_{DQ} is the dc quiescent voltage and v_d is the ac component. We are neglecting the -1 term in the diode equation given by Equation (1.22). Equation (1.26) can be rewritten as

$$i_D = I_S \left[e^{\left(\frac{V_{DQ}}{V_T}\right)} \right] \cdot \left[e^{\left(\frac{v_d}{V_T}\right)} \right] \quad (1.27)$$

If the ac signal is “small,” then $v_d \ll V_T$, and we can expand the exponential function into a linear series, as follows:

$$e^{\left(\frac{v_d}{V_T}\right)} \cong 1 + \frac{v_d}{V_T} \quad (1.28)$$

We may also write the quiescent diode current as

$$I_{DQ} = I_S e^{\left(\frac{V_{DQ}}{V_T}\right)} \quad (1.29)$$

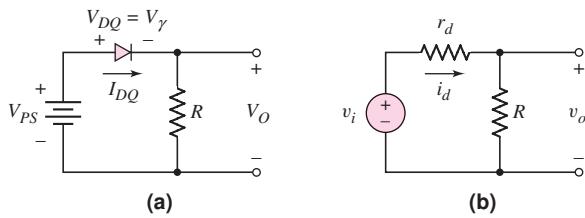


Figure 1.36 Equivalent circuits: (a) dc and (b) ac

The diode current–voltage relationship from Equation (1.27) can then be written as

$$i_D = I_{DQ} \left(1 + \frac{v_d}{V_T} \right) = I_{DQ} + \frac{I_{DQ}}{V_T} \cdot v_d = I_{DQ} + i_d \quad (1.30)$$

where i_d is the ac component of the diode current. The relationship between the ac components of the diode voltage and current is then

$$i_d = \left(\frac{I_{DQ}}{V_T} \right) \cdot v_d = g_d \cdot v_d \quad (1.31(a))$$

or

$$v_d = \left(\frac{V_T}{I_{DO}} \right) \cdot i_d = r_d \cdot i_d \quad (1.31(b))$$

The parameters g_d and r_d , respectively, are the diode **small-signal incremental conductance** and **resistance**, also called the **diffusion conductance** and **diffusion resistance**. We see from these two equations that

$$r_d = \frac{1}{g_d} = \frac{V_T}{I_{DO}} \quad (1.32)$$

This equation tells us that the incremental resistance is a function of the dc bias current I_{DQ} and is inversely proportional to the slope of the I - V characteristics curve, as shown in Figure 1.35(d).

Circuit Analysis

To analyze the circuit shown in Figure 1.35(a), we first perform a dc analysis and then an ac analysis. These two types of analyses will use two equivalent circuits. Figure 1.36(a) is the dc equivalent circuit that we have seen previously. If the diode is forward biased, then the voltage across the diode is the piecewise linear turn-on voltage.

Figure 1.36(b) is the ac equivalent circuit. The diode has been replaced by its equivalent resistance r_d . All parameters in this circuit are the small-signal time-varying parameters.

EXAMPLE 1.11

Objective: Analyze the circuit shown in Figure 1.35(a).

Assume circuit and diode parameters of $V_{PS} = 5$ V, $R = 5$ k Ω , $V_\gamma = 0.6$ V, and $v_i = 0.1 \sin \omega t$ (V).

Solution: Divide the analysis into two parts: the dc analysis and the ac analysis.

For the dc analysis, we set $v_i = 0$ and then determine the dc quiescent current from Figure 1.36(a) as

$$I_{DQ} = \frac{V_{PS} - V_Y}{R} = \frac{5 - 0.6}{5} = 0.88 \text{ mA}$$

The dc value of the output voltage is

$$V_o = I_{DQ} R = (0.88)(5) = 4.4 \text{ V}$$

For the ac analysis, we consider only the ac signals and parameters in the circuit in Figure 1.36(b). In other words, we effectively set $V_{PS} = 0$. The ac Kirchhoff voltage law (KVL) equation becomes

$$v_i = i_d r_d + i_d R = i_d(r_d + R)$$

where r_d is again the small-signal diode diffusion resistance. From Equation (1.32), we have

$$r_d = \frac{V_T}{I_{DQ}} = \frac{0.026}{0.88} = 0.0295 \text{ k}\Omega$$

The ac diode current is

$$i_d = \frac{v_i}{r_d + R} = \frac{0.1 \sin \omega t}{0.0295 + 5} \Rightarrow 19.9 \sin \omega t (\mu\text{A})$$

The ac component of the output voltage is

$$v_o = i_d R = 0.0995 \sin \omega t (\text{V})$$

Comment: Throughout the text, we will divide the circuit analysis into a dc analysis and an ac analysis. To do so, we will use separate equivalent circuit models for each analysis.

EXERCISE PROBLEM

Ex 1.11: (a) The circuit and diode parameters for the circuit shown in Figure 1.35(a) are $V_{PS} = 8 \text{ V}$, $R = 20 \text{ k}\Omega$, $V_Y = 0.7 \text{ V}$, and $v_i = 0.25 \sin \omega t (\text{V})$. Determine the quiescent diode current and the time-varying diode current. (b) Repeat part (a) if the resistor is changed to $R = 10 \text{ k}\Omega$. (Ans. (a) $I_{DQ} = 0.365 \text{ mA}$, $i_d = 12.5 \sin \omega t (\mu\text{A})$; (b) $I_{DQ} = 0.730 \text{ mA}$, $i_d = 24.9 \sin \omega t (\mu\text{A})$).

Frequency Response

In the previous analysis, we implicitly assumed that the frequency of the ac signal was small enough that capacitance effects in the circuit would be negligible. If the frequency of the ac input signal increases, the **diffusion capacitance** associated with a forward-biased pn junction becomes important. The source of the diffusion capacitance is shown in Figure 1.37.

Consider the minority carrier hole concentration on the right side of the figure. At the quiescent diode voltage, V_{DQ} , the minority carrier hole concentration is shown as the solid line and indicated by $p_{n|V_{DQ}}$.

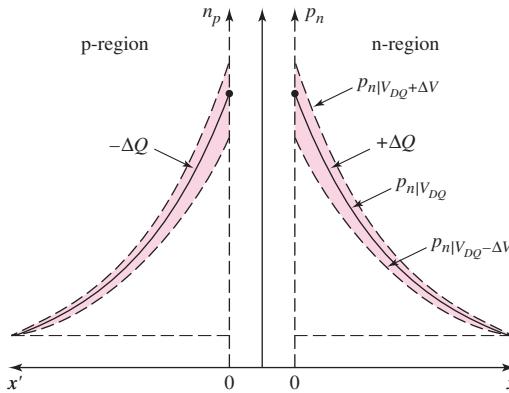


Figure 1.37 Change in minority carrier stored charge with a time-varying voltage superimposed on a dc quiescent diode voltage. The change in stored charge leads to a diode diffusion capacitance.

If the total diode voltage increases by ΔV during the positive half cycle of a sinusoidal signal superimposed on the quiescent value, the hole concentration will increase to that shown by the dotted line indicated by $p_n|V_{DQ} + \Delta V$. Now, if the total diode voltage decreases by ΔV during the negative half cycle of a sinusoidal signal superimposed on the quiescent value, the hole concentration will decrease to that shown by the dotted line indicated by $p_n|V_{DQ} - \Delta V$. The $+\Delta Q$ charge is alternately being charged and discharged through the pn junction as the voltage across the junction changes.

The same process is occurring with the minority carrier electrons in the p-region.

The diffusion capacitance is the change in the stored minority carrier charge that is caused by a change in the voltage, or

$$C_d = \frac{dQ}{dV_D} \quad (1.33)$$

The diffusion capacitance C_d is normally much larger than the junction capacitance C_j , because of the magnitude of the charges involved.

1.4.2 Small-Signal Equivalent Circuit

The small-signal equivalent circuit of the forward-biased pn junction is shown in Figure 1.38 and is developed partially from the equation for the **admittance**, which is given by

$$Y = g_d + j\omega C_d \quad (1.34)$$

where g_d and C_d are the diffusion conductance and capacitance, respectively. We must also add the junction capacitance, which is in parallel with the diffusion resistance and capacitance, and a series resistance, which is required because of the finite resistances in the neutral n- and p-regions.

The small-signal equivalent circuit of the pn junction is used to obtain the ac response of a diode circuit subjected to ac signals superimposed on the Q-point values. Small-signal equivalent circuits of pn junctions are also used to develop small-signal models of transistors, and these models are used in the analysis and design of transistor amplifiers.

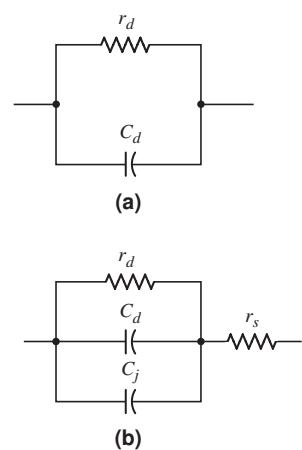


Figure 1.38 Small-signal equivalent circuit of the diode: (a) simplified version and (b) complete circuit

Test Your Understanding

TYU 1.12 Determine the diffusion conductance of a pn junction diode at $T = 300\text{ K}$ and biased at a current of 0.8 mA . (Ans. $g_d = 30.8\text{ mS}$)

TYU 1.13 Determine the small-signal diffusion resistance of a pn junction diode at $I_D = 10\text{ }\mu\text{A}$, $100\text{ }\mu\text{A}$, and 1 mA . (Ans. $2.6\text{ k}\Omega$, $260\text{ }\Omega$, $26\text{ }\Omega$).

TYU 1.14 The diffusion resistance of a pn junction diode at $T = 300\text{ K}$ is determined to be $r_d = 50\text{ }\Omega$. What is the quiescent diode current? (Ans. $I_{DQ} = 0.52\text{ mA}$)



1.5 OTHER DIODE TYPES

Objective: • Gain an understanding of the properties and characteristics of a few specialized diodes.

There are many other types of diodes with specialized characteristics that are useful in particular applications. We will briefly consider only a few of these diodes. We will consider the solar cell, photodiode, light-emitting diode, Schottky diode, and Zener diode.

1.5.1 Solar Cell

A **solar cell** is a pn junction device with no voltage directly applied across the junction. The pn junction, which converts solar energy into electrical energy, is connected to a load as indicated in Figure 1.39. When light hits the space-charge region, electrons and holes are generated. They are quickly separated and swept out of the space-charge region by the electric field, thus creating a **photocurrent**. The generated photocurrent will produce a voltage across the load, which means that the solar cell has supplied power. Solar cells are usually fabricated from silicon, but may be made from GaAs or other III–V compound semiconductors.

Solar cells have long been used to power the electronics in satellites and space vehicles, and also as the power supply to some calculators. Solar cells are also used to power race cars in a Sunrayce event. Collegiate teams in the United States design,

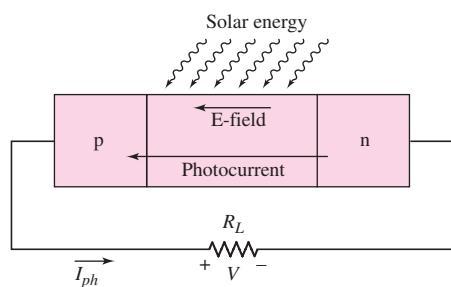


Figure 1.39 A pn junction solar cell connected to load

build and drive the race cars. Typically, a Sunrayce car has 8 m^2 of solar cell arrays that can produce 800 W of power on a sunny day at noon. The power from the solar array can be used either to power an electric motor or to charge a battery pack.

1.5.2 Photodiode

Photodetectors are devices that convert optical signals into electrical signals. An example is the **photodiode**, which is similar to a solar cell except that the pn junction is operated with a reverse-bias voltage. Incident photons or light waves create excess electrons and holes in the space-charge region. These excess carriers are quickly separated and swept out of the space-charge region by the electric field, thus creating a “photocurrent.” This generated photocurrent is directly proportional to the incident photon flux.

1.5.3 Light-Emitting Diode

The **light-emitting diode (LED)** converts current to light. As previously explained, when a forward-bias voltage is applied across a pn junction, electrons and holes flow across the space-charge region and become excess minority carriers. These excess minority carriers diffuse into the neutral semiconductor regions, where they recombine with majority carriers. If the semiconductor is a **direct bandgap material**, such as GaAs, the electron and hole can recombine with no change in momentum, and a photon or light wave can be emitted. Conversely, in an **indirect bandgap material**, such as silicon, when an electron and hole recombine, both energy and momentum must be conserved, so the emission of a photon is very unlikely. Therefore, LEDs are fabricated from GaAs or other compound semiconductor materials. In an LED, the diode current is directly proportional to the recombination rate, which means that the output light intensity is also proportional to the diode current.

Monolithic arrays of LEDs are fabricated for numeric and alphanumeric displays, such as the readout of a digital voltmeter.

An LED may be integrated into an optical cavity to produce a coherent photon output with a very narrow bandwidth. Such a device is a laser diode, which is used in optical communications applications.

The LED can be used in conjunction with a photodiode to create an optical system such as that shown in Figure 1.40. The light signal created may travel over

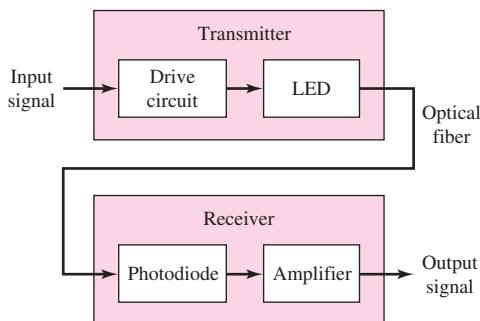


Figure 1.40 Basic elements in an optical transmission system

relatively long distances through the optical fiber, because of the low optical absorption in high-quality optical fibers.

1.5.4 Schottky Barrier Diode

A **Schottky barrier diode**, or simply a Schottky diode, is formed when a metal, such as aluminum, is brought into contact with a *moderately* doped n-type semiconductor to form a rectifying junction. Figure 1.41(a) shows the metal-semiconductor contact, and Figure 1.41(b) shows the circuit symbol with the current direction and voltage polarity.

The current–voltage characteristics of a Schottky diode are very similar to those of a pn junction diode. The same ideal diode equation can be used for both devices. However, there are two important differences between the two diodes that directly affect the response of the Schottky diode.

First, the current mechanism in the two devices is different. The current in a pn junction diode is controlled by the diffusion of minority carriers. The current in a Schottky diode results from the flow of majority carriers over the potential barrier at the metallurgical junction. This means that there is no minority carrier storage in the Schottky diode, so the switching time from a forward bias to a reverse bias is very short compared to that of a pn junction diode. The storage time, t_s , for a Schottky diode is essentially zero.

Second, the reverse-saturation current I_S for a Schottky diode is larger than that of a pn junction diode for comparable device areas. This property means that it takes less forward bias voltage to induce a particular current compared to a pn junction diode. We will see an application of this in Chapter 17.

Figure 1.42 compares the characteristics of the two diodes. Applying the piecewise linear model, we can determine that the Schottky diode has a smaller turn-on

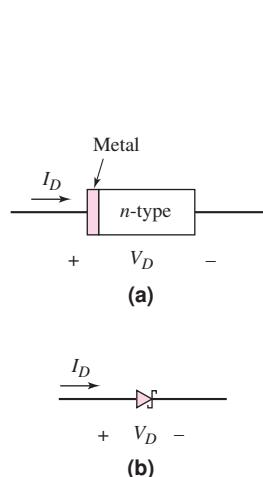


Figure 1.41 Schottky barrier diode: (a) simplified geometry and (b) circuit symbol

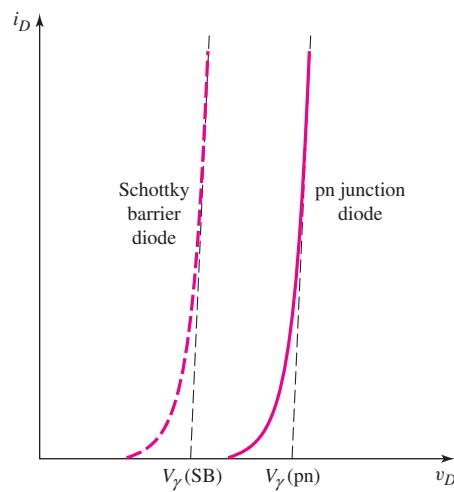


Figure 1.42 Comparison of the forward-bias I – V characteristics of a pn junction diode and a Schottky barrier diode

voltage than the pn junction diode. In later chapters, we will see how this lower turn-on voltage and the shorter switching time make the Schottky diode useful in integrated-circuit applications.

EXAMPLE 1.12

Objective: Determine diode voltages.

The reverse saturation currents of a pn junction diode and a Schottky diode are $I_S = 10^{-12}$ A and 10^{-8} A, respectively. Determine the forward-bias voltages required to produce 1 mA in each diode.

Solution: The diode current-voltage relationship is given by

$$I_D = I_S e^{V_D/V_T}$$

Solving for the diode voltage, we obtain

$$V_D = V_T \ln\left(\frac{I_D}{I_S}\right)$$

We then find, for the pn junction diode

$$V_D = (0.026) \ln\left(\frac{1 \times 10^{-3}}{10^{-12}}\right) = 0.539 \text{ V}$$

and, for the Schottky diode

$$V_D = (0.026) \ln\left(\frac{1 \times 10^{-3}}{10^{-8}}\right) = 0.299 \text{ V}$$

Comment: Since the reverse-saturation current for the Schottky diode is relatively large, less voltage across this diode is required to produce a given current compared to the pn junction diode.

EXERCISE PROBLEM

Ex 1.12: A pn junction diode and a Schottky diode both have forward-bias currents of 1.2 mA. The reverse-saturation current of the pn junction diode is $I_S = 4 \times 10^{-15}$ A. The difference in forward-bias voltages is 0.265 V. Determine the reverse-saturation current of the Schottky diode. (Ans. $I_S = 1.07 \times 10^{-10}$ A)

Another type of metal–semiconductor junction is also possible. A metal applied to a heavily doped semiconductor forms, in most cases, an *ohmic contact*: that is, a contact that conducts current equally in both directions, with very little voltage drop across the junction. Ohmic contacts are used to connect one semiconductor device to another on an IC, or to connect an IC to its external terminals.

1.5.5 Zener Diode

As mentioned earlier in this chapter, the applied reverse-bias voltage cannot increase without limit. At some point, breakdown occurs and the current in the reverse-bias

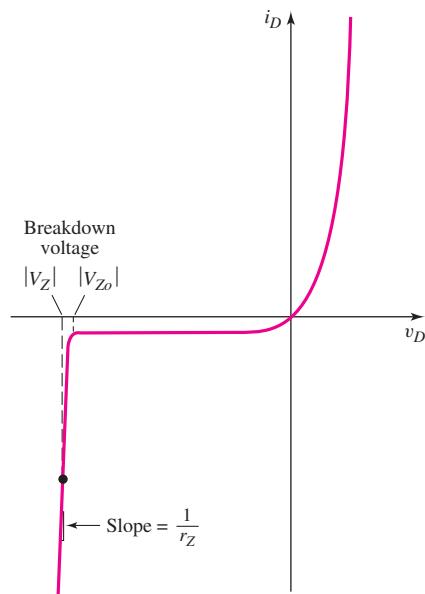


Figure 1.43 Diode I - V characteristics showing breakdown effects

direction increases rapidly. The voltage at this point is called the breakdown voltage. The diode I - V characteristics, including breakdown, are shown in Figure 1.43.

Diodes, called **Zener diodes**, can be designed and fabricated to provide a specified breakdown voltage V_{Zo} . (Although the breakdown voltage is on the negative voltage axis (reverse-bias), its value is given as a positive quantity.) The large current that may exist at breakdown can cause heating effects and catastrophic failure of the diode due to the large power dissipation in the device. However, diodes can be operated in the breakdown region by limiting the current to a value within the capabilities of the device. Such a diode can be used as a constant-voltage reference in a circuit. The diode breakdown voltage is essentially constant over a wide range of currents and temperatures. The slope of the I - V characteristics curve in breakdown is quite large, so the incremental resistance r_z is small. Typically, r_z is in the range of a few ohms or tens of ohms.

The circuit symbol of the Zener diode is shown in Figure 1.44. (Note the subtle difference between this symbol and the Schottky diode symbol.) The voltage V_Z is the Zener breakdown voltage, and the current I_Z is the reverse-bias current when the diode is operating in the breakdown region. We will see applications of the Zener diode in the next chapter.

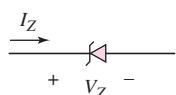


Figure 1.44 Circuit symbol of the Zener diode

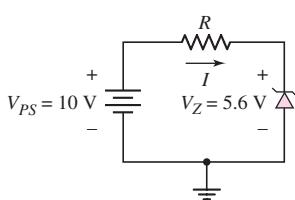


Figure 1.45 Simple circuit containing a Zener diode in which the Zener diode is biased in the breakdown region

DESIGN EXAMPLE 1.13

Objective: Consider a simple constant-voltage reference circuit and design the value of resistance required to limit the current in this circuit.

Consider the circuit shown in Figure 1.45. Assume that the Zener diode breakdown voltage is $V_Z = 5.6$ V and the Zener resistance is $r_z = 0$. The current in the diode is to be limited to 3 mA.

Solution: As before, we can determine the current from the voltage difference across R divided by the resistance. That is,

$$I = \frac{V_{PS} - V_Z}{R}$$

The resistance is then

$$R = \frac{V_{PS} - V_Z}{I} = \frac{10 - 5.6}{3} = 1.47 \text{ k}\Omega$$

The power dissipated in the Zener diode is

$$P_Z = I_Z V_Z = (3)(5.6) = 16.8 \text{ mW}$$

The Zener diode must be able to dissipate 16.8 mW of power without being damaged.

Comment: The resistance external to the Zener diode limits the current when the diode is operating in the breakdown region. In the circuit shown in the figure, the output voltage will remain constant at 5.6 V, even though the power supply voltage and the resistance may change over a limited range. Hence, this circuit provides a constant output voltage. We will see further applications of the Zener diode in the next chapter.

EXERCISE PROBLEM

Ex 1.13: Consider the circuit shown in Figure 1.45. Determine the value of resistance R required to limit the power dissipated in the Zener diode to 10 mW.
(Ans. $R = 2.46 \text{ k}\Omega$)

Test Your Understanding

TYU 1.15 Consider the circuit shown in Figure 1.46. The diode can be either a pn junction diode or a Schottky diode. Assume the cut-in voltages are $V_Y = 0.7 \text{ V}$ and $V_Y = 0.3 \text{ V}$ for the pn junction diode and Schottky diode, respectively. Let $r_f = 0$ for both diodes. Calculate the current I_D when each diode is inserted in the circuit.
(Ans. pn diode, 0.825 mA; Schottky diode, 0.925 mA).

TYU 1.16 A Zener diode has an equivalent series resistance of 20Ω . If the voltage across the Zener diode is 5.20 V at $I_Z = 1 \text{ mA}$, determine the voltage across the diode at $I_Z = 10 \text{ mA}$. (Ans. $V_Z = 5.38 \text{ V}$)

TYU 1.17 The resistor in the circuit shown in Figure 1.45 has a value of $R = 4 \text{ k}\Omega$, the Zener diode breakdown voltage is $V_Z = 3.6 \text{ V}$, and the power rating of the Zener diode is $P = 6.5 \text{ mW}$. Determine the maximum diode current and the maximum power supply voltage that can be applied without damaging the diode. (Ans. 1.81 mA, 10.8 V).

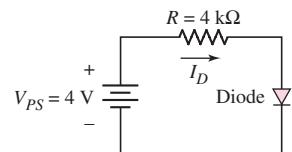


Figure 1.46 Circuit for exercise problem TYU 1.15. The diode can be either a pn junction diode or a Schottky diode.



1.6 DESIGN APPLICATION: DIODE THERMOMETER

Objective: • Design a simple electronic thermometer using the temperature characteristics of a diode.

Specifications: The temperature range is to be 0 to 100 °F.

Design Approach: We will use the forward-bias diode temperature characteristics as shown in Figure 1.20. If the diode current is held constant, the variation in diode voltage is a function of temperature.

Choices: Assume that a silicon pn junction diode with a reverse-saturation current of $I_S = 10^{-13}$ A at $T = 300$ K is available.

Solution: Neglecting the (-1) term in the diode $I-V$ relation, we have

$$I_D = I_S e^{V_D/V_T} \propto n_i^2 e^{V_D/V_T} \propto e^{-E_g/kT} \cdot e^{V_D/V_T}$$

The reverse-saturation current I_S is proportional to n_i^2 and in turn n_i^2 is proportional to the exponential function involving the bandgap energy E_g and temperature.

Taking the ratio of the diode current at two temperature values and using the definition of thermal voltage, we have³

$$\frac{I_{D1}}{I_{D2}} = \frac{e^{-E_g/kT_1} \cdot e^{eV_{D1}/kT_1}}{e^{-E_g/kT_2} \cdot e^{eV_{D2}/kT_2}} \quad (1.35)$$

where V_{D1} and V_{D2} are the diode voltages at temperatures T_1 and T_2 , respectively. If the diode current is held constant at the different temperatures, Equation (1.35) can be written as

$$e^{eV_{D2}/kT_2} = e^{-E_g/kT_1} e^{+E_g/kT_2} e^{eV_{D1}/kT_1} \quad (1.36)$$

Taking the natural logarithm of both sides, we obtain

$$\frac{eV_{D2}}{kT_2} = \frac{-E_g}{kT_1} + \frac{E_g}{kT_2} + \frac{eV_{D1}}{kT_1} \quad (1.37)$$

or

$$V_{D2} = \frac{-E_g}{e} \left(\frac{T_2}{T_1} \right) + \frac{E_g}{e} + V_{D1} \left(\frac{T_2}{T_1} \right) \quad (1.38)$$

For silicon, the bandgap energy is $E_g/e = 1.12$ V. Then, assuming the bandgap energy does not vary over the temperature range, we have

$$V_{D2} = 1.12 \left(1 - \frac{T_2}{T_1} \right) + V_{D1} \left(\frac{T_2}{T_1} \right) \quad (1.39)$$

³Note that e in, for example, $e^{-E_g/kT}$ represents the exponential function whereas e in the exponent, for example, eV_{D1}/kT_1 is the magnitude of the electronic charge. The context in which e is used should make the meaning clear.

Consider the circuit shown in Figure 1.47. Assume that the diode has a reverse-saturation current of $I_S = 10^{-13}$ A at $T = 300$ K. From the circuit, we can write

$$I_D = \frac{15 - V_D}{R} = I_S e^{V_D/V_T}$$

or

$$\frac{15 - V_D}{15 \times 10^3} = 10^{-13} e^{V_D/0.026}$$

By trial and error, we find

$$V_D = 0.5976 \text{ V}$$

and

$$I_D = \frac{15 - 0.5976}{15 \times 10^3} \Rightarrow 0.960 \text{ mA}$$

In Equation (1.39), we can set $T_1 = 300$ K and let $T_2 \equiv T$ be a variable temperature. We find

$$V_D = 1.12 - 0.522 \left(\frac{T}{300} \right) \quad (1.40)$$

so the diode voltage is a linear function of temperature. If the temperature range is to be from 0 to 100 °F, for example, the corresponding change in kelvins is from 255.2 to 310.8. The diode voltage versus temperature is plotted in Figure 1.48.

A simple circuit that can be used was shown in Figure 1.47. With a power supply voltage of 15 V, a change in diode voltage of approximately 0.1 V over the temperature range produces only an approximately 0.67 percent change in diode current. Thus the preceding analysis is valid.

Comment: This design example shows that a diode connected in a simple circuit can be used as a sensing element in an electronic thermometer. We assumed a diode reverse-saturation current of $I_S = 10^{-13}$ A at $T = 300$ K(80 °F). The actual reverse-saturation current of a particular diode may be different. This difference simply means that the diode voltage versus temperature curve shown in Figure 1.48 would slide up or down to match the actual diode voltage at room temperature.

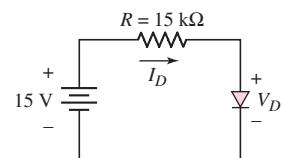


Figure 1.47 Circuit of diode thermometer

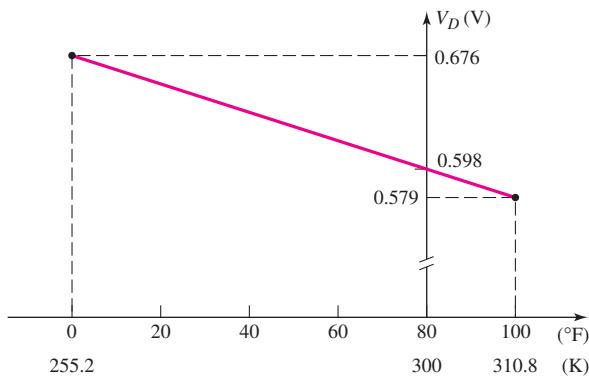


Figure 1.48 Diode voltage versus temperature

Design Pointer: In order to complete this design, two additional components or electronic systems must be added to the circuit shown in Figure 1.47. First, we must add a circuit to measure the diode voltage. Adding this circuit must not alter the diode characteristics and there must be no loading effects. An op-amp circuit that will be described in Chapter 9 can be used for this purpose. A second electronic system required is to convert the diode voltage to a temperature reading. An analog-to-digital converter that will be described in Chapter 16 can be used to provide a digital temperature reading.

1.7 SUMMARY

- We initially considered some of the characteristics and properties of semiconductor materials, such as the concept of electrons (negative charge) and holes (positive charge) as two distinct charge carriers in a semiconductor. The doping process produces either n-type or p-type semiconductor materials. The concepts of n-type and p-type materials are used throughout the text.
- A pn junction diode is formed when an n-doped region and a p-doped region are directly adjacent to each other. The diode current is an exponential function of voltage in the forward-bias condition, and is essentially zero in the reverse-bias condition.
- A piecewise-linear model of the diode was developed so that approximate hand calculation results can be easily obtained. The $i - v$ characteristics of the diode are broken into linear segments, which are valid over particular regions of operation. The concept of a diode turn-on voltage was introduced.
- Time-varying, or ac signals, may be superimposed on a dc diode current and voltage. A small-signal linear equivalent circuit was developed and is used to determine the relationship between the ac current and ac voltage. This same equivalent circuit will be applied extensively when the frequency response of transistors is discussed.
- Specialized pn junction devices were discussed. In particular, pn junction solar cells are used to convert solar energy to electrical energy. Schottky barrier diodes are metal–semiconductor rectifying junctions that, in general, have smaller turn-on voltages than pn junctions. Zener diodes operate in the reverse breakdown region and are used in constant-voltage circuits. Photodiodes and LEDs were also briefly discussed.
- As an application, a simple diode thermometer was designed, based on the temperature properties of the pn junction.

CHECKPOINT

After studying this chapter, the reader should have the ability to:

- ✓ Understand the concept of intrinsic carrier concentration, the difference between n-type and p-type materials, and the concept of drift and diffusion currents.
- ✓ Analyze a simple diode circuit using the ideal diode current–voltage characteristics and using the iteration analysis technique.
- ✓ Analyze a diode circuit using the piecewise linear approximation model for the diode.

- ✓ Determine the small-signal characteristics of a diode using the small-signal equivalent circuit.
- ✓ Understand the general characteristics of a solar cell, light-emitting diode, Schottky barrier diode, and Zener diode.

REVIEW QUESTIONS

1. Describe an intrinsic semiconductor material. What is meant by the intrinsic carrier concentration?
2. Describe the concept of an electron and a hole as charge carriers in the semiconductor material.
3. Describe an extrinsic semiconductor material. What is the electron concentration in terms of the donor impurity concentration? What is the hole concentration in terms of the acceptor impurity concentration?
4. Describe the concepts of drift current and diffusion current in a semiconductor material.
5. How is a pn junction formed? What is meant by a built-in potential barrier, and how is it formed?
6. How is a junction capacitance created in a reverse-biased pn junction diode?
7. Write the ideal diode current–voltage relationship. Describe the meaning of I_S and V_T .
8. Describe the iteration method of analysis and when it must be used to analyze a diode circuit.
9. Describe the piecewise linear model of a diode and why it is useful. What is the diode turn-on voltage?
10. Define a load line in a simple diode circuit.
11. Under what conditions is the small-signal model of a diode used in the analysis of a diode circuit?
12. Describe the operation of a simple solar cell circuit.
13. How do the i – v characteristics of a Schottky barrier diode differ from those of a pn junction diode?
14. What characteristic of a Zener diode is used in the design of a Zener diode circuit?
15. Describe the characteristics of a photodiode and a photodiode circuit.

PROBLEMS

[Note: Unless otherwise specified, assume that $T = 300\text{ K}$ in the following problems. Also, assume the emission coefficient is $n = 1$ unless otherwise stated.]

Section 1.1 Semiconductor Materials and Properties

- 1.1 (a) Calculate the intrinsic carrier concentration in silicon at (i) $T = 250\text{ K}$ and (ii) $T = 350\text{ K}$. (b) Repeat part (a) for gallium arsenide.
- 1.2 (a) The intrinsic carrier concentration in silicon is to be no larger than $n_i = 10^{12}\text{ cm}^{-3}$. Determine the maximum allowable temperature. (b) Repeat part (a) for $n_i = 10^9\text{ cm}^{-3}$.
- 1.3 Calculate the intrinsic carrier concentration in silicon and germanium at (a) $T = 100\text{ K}$, (b) $T = 300\text{ K}$, and (c) $T = 500\text{ K}$.

- 1.4 (a) Find the concentration of electrons and holes in a sample of germanium that has a concentration of donor atoms equal to 10^{15} cm^{-3} . Is the semiconductor n-type or p-type? (b) Repeat part (a) for silicon.
- 1.5 Gallium arsenide is doped with acceptor impurity atoms at a concentration of 10^{16} cm^{-3} . (a) Find the concentration of electrons and holes. Is the semiconductor n-type or p-type? (b) Repeat part (a) for germanium.
- 1.6 Silicon is doped with 5×10^{16} arsenic atoms/cm³. (a) Is the material n- or p-type? (b) Calculate the electron and hole concentrations at $T = 300 \text{ K}$. (c) Repeat part (b) for $T = 350 \text{ K}$.
- 1.7 (a) Calculate the concentration of electrons and holes in silicon that has a concentration of acceptor atoms equal to $5 \times 10^{16} \text{ cm}^{-3}$. Is the semiconductor n-type or p-type? (b) Repeat part (a) for GaAs.
- 1.8 A silicon sample is fabricated such that the hole concentration is $p_o = 2 \times 10^{17} \text{ cm}^{-3}$. (a) Should boron or arsenic atoms be added to the intrinsic silicon? (b) What concentration of impurity atoms must be added? (c) What is the concentration of electrons?
- 1.9 The electron concentration in silicon at $T = 300 \text{ K}$ is $n_o = 5 \times 10^{15} \text{ cm}^{-3}$. (a) Determine the hole concentration. (b) Is the material n-type or p-type? (c) What is the impurity doping concentration?
- 1.10 (a) A silicon semiconductor material is to be designed such that the majority carrier electron concentration is $n_o = 7 \times 10^{15} \text{ cm}^{-3}$. Should donor or acceptor impurity atoms be added to intrinsic silicon to achieve this electron concentration? What concentration of dopant impurity atoms is required? (b) In this silicon material, the minority carrier hole concentration is to be no larger than $p_o = 10^6 \text{ cm}^{-3}$. Determine the maximum allowable temperature.
- 1.11 (a) The applied electric field in p-type silicon is $E = 10 \text{ V/cm}$. The semiconductor conductivity is $\sigma = 1.5 (\Omega\text{-cm})^{-1}$ and the cross-sectional area is $A = 10^{-5} \text{ cm}^2$. Determine the drift current. (b) The cross-sectional area of a semiconductor is $A = 2 \times 10^{-4} \text{ cm}^2$ and the resistivity is $\rho = 0.4 \text{ } (\Omega\text{-cm})$. If the drift current is $I = 1.2 \text{ mA}$, what applied electric field must be applied?
- 1.12 A drift current density of 120 A/cm^2 is established in n-type silicon with an applied electric field of 18 V/cm . If the electron and hole mobilities are $\mu_n = 1250 \text{ cm}^2/\text{V}\cdot\text{s}$ and $\mu_p = 450 \text{ cm}^2/\text{V}\cdot\text{s}$, respectively, determine the required doping concentration.
- 1.13 An n-type silicon material has a resistivity of $\rho = 0.65 \text{ } \Omega\text{-cm}$. (a) If the electron mobility is $\mu_n = 1250 \text{ cm}^2/\text{V}\cdot\text{s}$, what is the concentration of donor atoms? (b) Determine the required electric field to establish a drift current density of $J = 160 \text{ A/cm}^2$.
- 1.14 (a) The required conductivity of a silicon material must be $\sigma = 1.5 (\Omega\text{-cm})^{-1}$. If $\mu_n = 1000 \text{ cm}^2/\text{V}\cdot\text{s}$ and $\mu_p = 375 \text{ cm}^2/\text{V}\cdot\text{s}$, what concentration of donor atoms must be added? (b) The required conductivity of a silicon material must be $\sigma = 0.8 (\Omega\text{-cm})^{-1}$. If $\mu_n = 1200 \text{ cm}^2/\text{V}\cdot\text{s}$ and $\mu_p = 400 \text{ cm}^2/\text{V}\cdot\text{s}$, what concentration of acceptor atoms must be added?
- 1.15 In GaAs, the mobilities are $\mu_n = 8500 \text{ cm}^2/\text{V}\cdot\text{s}$ and $\mu_p = 400 \text{ cm}^2/\text{V}\cdot\text{s}$. (a) Determine the range in conductivity for a range in donor concentration of $10^{15} \leq N_d \leq 10^{19} \text{ cm}^{-3}$. (b) Using the results of part (a), determine the range in drift current density if the applied electric field is $E = 0.10 \text{ V/cm}$.

- 1.16 The electron and hole concentrations in a sample of silicon are shown in Figure P1.16. Assume the electron and hole mobilities are the same as in Problem 1.12. Determine the total diffusion current density versus distance x for $0 \leq x \leq 0.001$ cm.

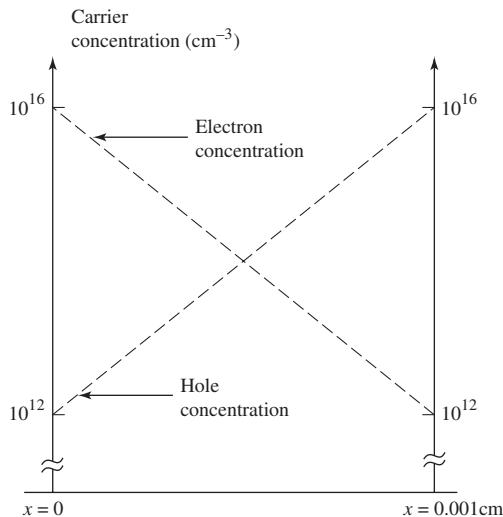


Figure P1.16

- 1.17 The hole concentration in silicon is given by

$$p(x) = 10^4 + 10^{15} \exp(-x/L_p) \quad x \geq 0$$

The value of L_p is $10 \mu\text{m}$. The hole diffusion coefficient is $D_p = 15 \text{ cm}^2/\text{s}$. Determine the hole diffusion current density at (a) $x = 0$, (b) $x = 10 \mu\text{m}$, and (c) $x = 30 \mu\text{m}$.

- 1.18 GaAs is doped to $N_a = 10^{17} \text{ cm}^{-3}$. (a) Calculate n_o and p_o . (b) Excess electrons and holes are generated such that $\delta n = \delta p = 10^{15} \text{ cm}^{-3}$. Determine the total concentration of electrons and holes.

Section 1.2 The pn Junction

- 1.19 (a) Determine the built-in potential barrier V_{bi} in a silicon pn junction for (i) $N_d = N_a = 5 \times 10^{15} \text{ cm}^{-3}$; (ii) $N_d = 5 \times 10^{17} \text{ cm}^{-3}$ and $N_a = 10^{15} \text{ cm}^{-3}$; (iii) $N_a = N_d = 10^{18} \text{ cm}^{-3}$. (b) Repeat part (a) for GaAs.
- 1.20 Consider a silicon pn junction. The n-region is doped to a value of $N_d = 10^{16} \text{ cm}^{-3}$. The built-in potential barrier is to be $V_{bi} = 0.712 \text{ V}$. Determine the required p-type doping concentration.
- 1.21 The donor concentration in the n-region of a silicon pn junction is $N_d = 10^{16} \text{ cm}^{-3}$. Plot V_{bi} versus N_a over the range $10^{15} \leq N_a \leq 10^{18} \text{ cm}^{-3}$ where N_a is the acceptor concentration in the p-region.
- 1.22 Consider a uniformly doped GaAs pn junction with doping concentrations of $N_a = 5 \times 10^{18} \text{ cm}^{-3}$ and $N_d = 5 \times 10^{16} \text{ cm}^{-3}$. Plot the built-in potential barrier V_{bi} versus temperature for $200 \text{ K} \leq T \leq 500 \text{ K}$.
- 1.23 The zero-biased junction capacitance of a silicon pn junction is $C_{jo} = 0.4 \text{ pF}$. The doping concentrations are $N_a = 1.5 \times 10^{16} \text{ cm}^{-3}$ and

$N_d = 4 \times 10^{15} \text{ cm}^{-3}$. Determine the junction capacitance at (a) $V_R = 1 \text{ V}$, (b) $V_R = 3 \text{ V}$, and (c) $V_R = 5 \text{ V}$.

- *1.24 The zero-bias capacitance of a silicon pn junction diode is $C_{jo} = 0.02 \text{ pF}$ and the built-in potential is $V_{bi} = 0.80 \text{ V}$. The diode is reverse biased through a $47\text{-k}\Omega$ resistor and a voltage source. (a) For $t < 0$, the applied voltage is 5 V and, at $t = 0$, the applied voltage drops to zero volts. Estimate the time it takes for the diode voltage to change from 5 V to 1.5 V . (As an approximation, use the average diode capacitance between the two voltage levels.) (b) Repeat part (a) for an input voltage change from 0 V to 5 V and a diode voltage change from 0 V to 3.5 V . (Use the average diode capacitance between these two voltage levels.)
- 1.25 The doping concentrations in a silicon pn junction are $N_d = 5 \times 10^{15} \text{ cm}^{-3}$ and $N_a = 10^{17} \text{ cm}^{-3}$. The zero-bias junction capacitance is $C_{jo} = 0.60 \text{ pF}$. An inductance of 1.50 mH is connected in parallel with the pn junction. Calculate the resonant frequency f_o of the circuit for reverse-bias voltages of (a) $V_R = 1 \text{ V}$, (b) $V_R = 3 \text{ V}$, and (c) $V_R = 5 \text{ V}$.
- 1.26 (a) At what reverse-bias voltage does the reverse-bias current in a silicon pn junction diode reach 90 percent of its saturation value? (b) What is the ratio of the current for a forward-bias voltage of 0.2 V to the current for a reverse-bias voltage of 0.2 V ?
- 1.27 (a) The reverse-saturation current of a pn junction diode is $I_S = 10^{-11} \text{ A}$. Determine the diode current for diode voltages of $0.3, 0.5, 0.7, -0.02, -0.2$, and -2 V . (b) Repeat part (a) for $I_S = 10^{-13} \text{ A}$.
- 1.28 (a) The reverse-saturation current of a pn junction diode is $I_S = 10^{-11} \text{ A}$. Determine the diode voltage to produce currents of (i) $10 \mu\text{A}$, $100 \mu\text{A}$, 1 mA , and (ii) $-5 \times 10^{-12} \text{ A}$. (b) Repeat part (a) for $I_S = 10^{-13} \text{ A}$ and part (a) (ii) for -10^{-14} A .
- 1.29 A silicon pn junction diode has an emission coefficient of $n = 1$. The diode current is $I_D = 1 \text{ mA}$ when $V_D = 0.7 \text{ V}$. (a) What is the reverse-bias saturation current? (b) Plot, on the same graph, $\log_{10} I_D$ versus V_D over the range $0.1 \leq V_D \leq 0.7 \text{ V}$ when the emission coefficient is (i) $n = 1$ and (ii) $n = 2$.
- 1.30 Plot $\log_{10} I_D$ versus V_D over the range $0.1 \leq V_D \leq 0.7 \text{ V}$ for (a) $I_S = 10^{-12}$ and (b) $I_S = 10^{-14} \text{ A}$.
- 1.31 (a) Consider a silicon pn junction diode operating in the forward-bias region. Determine the increase in forward-bias voltage that will cause a factor of 10 increase in current. (b) Repeat part (a) for a factor of 100 increase in current.
- 1.32 A pn junction diode has $I_S = 2 \text{ nA}$. (a) Determine the diode voltage if (i) $I_D = 2 \text{ A}$ and (ii) $I_D = 20 \text{ A}$. (b) Determine the diode current if (i) $V_D = 0.4 \text{ V}$ and (ii) $V_D = 0.65 \text{ V}$.
- 1.33 The reverse-bias saturation current for a set of diodes varies between $5 \times 10^{-14} \leq I_S \leq 5 \times 10^{-12} \text{ A}$. The diodes are all to be biased at $I_D = 2 \text{ mA}$. What is the range of forward-bias voltages that must be applied?
- 1.34 (a) A germanium pn junction has a diode current of $I_D = 1.5 \text{ mA}$ when biased at $V_D = 0.30 \text{ V}$. What is the reverse-bias saturation current? (b) Using the results of part (a), determine the diode current when the diode is biased at (i) $V_D = 0.35 \text{ V}$ and (ii) $V_D = 0.25 \text{ V}$.
- 1.35 (a) The reverse-saturation current of a gallium arsenide pn junction diode is $I_S = 10^{-22} \text{ A}$. Determine the diode current for diode voltages of $0.8, 1.0, 1.2, -0.02, -0.2$, and -2 V . (b) Repeat part (a) for $I_S = 5 \times 10^{-24} \text{ A}$.

- *1.36 The reverse-saturation current of a silicon pn junction diode at $T = 300\text{ K}$ is $I_S = 10^{-12}\text{ A}$. Determine the temperature range over which I_S varies from $0.5 \times 10^{-12}\text{ A}$ to $50 \times 10^{-12}\text{ A}$.
- *1.37 A silicon pn junction diode has an applied forward-bias voltage of 0.6 V . Determine the ratio of current at 100°C to that at -55°C .

Section 1.3 DC Diode Analysis

- 1.38 A pn junction diode is in series with a $1\text{ M}\Omega$ resistor and a 2.8 V power supply. The reverse-saturation current of the diode is $I_S = 5 \times 10^{-11}\text{ A}$.
- Determine the diode current and voltage if the diode is forward biased.
 - Repeat part (a) if the diode is reverse biased.
- 1.39 Consider the diode circuit shown in Figure P1.39. The diode reverse-saturation current is $I_S = 10^{-12}\text{ A}$. Determine the diode current I_D and diode voltage V_D .
- *1.40 The diode in the circuit shown in Figure P1.40 has a reverse-saturation current of $I_S = 5 \times 10^{-13}\text{ A}$. Determine the diode voltage and current.
- 1.41 (a) For the circuit shown in Figure P1.41(a), determine I_{D1}, I_{D2}, V_{D1} , and V_{D2} for (i) $I_{S1} = I_{S2} = 10^{-13}\text{ A}$ and (ii) $I_{S1} = 5 \times 10^{-14}\text{ A}, I_{S2} = 5 \times 10^{-13}\text{ A}$.
- (b) Repeat part (a) for the circuit shown in Figure P1.41(b).

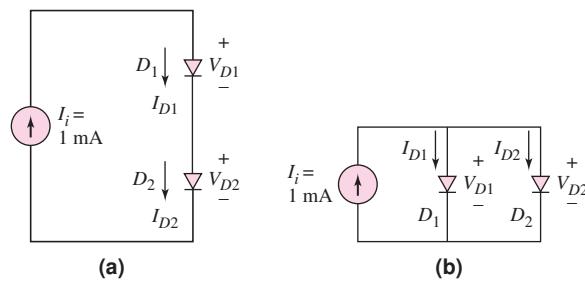


Figure P1.41

- 1.42 (a) The reverse-saturation current of each diode in the circuit shown in Figure P1.42 is $I_S = 6 \times 10^{-14}\text{ A}$. Determine the input voltage V_I required to produce an output voltage of $V_O = 0.635\text{ V}$. (b) Repeat part (a) if the $1\text{ k}\Omega$ resistor is changed to $R = 500\text{ }\Omega$.
- 1.43 (a) Consider the circuit shown in Figure P1.40. The value of R_1 is reduced to $R_1 = 10\text{ k}\Omega$ and the cut-in voltage of the diode is $V_\gamma = 0.7\text{ V}$. Determine I_D and V_D . (b) Repeat part (a) if $R_1 = 50\text{ k}\Omega$.
- 1.44 Consider the circuit shown in Figure P1.44. Determine the diode current I_D and diode voltage V_D for (a) $V_\gamma = 0.6\text{ V}$ and (b) $V_\gamma = 0.7\text{ V}$.
- 1.45 The diode cut-in voltage is $V_\gamma = 0.7\text{ V}$ for the circuits shown in Figure P1.45. Plot V_O and I_D versus I_I over the range $0 \leq I_I \leq 2\text{ mA}$ for the circuit shown in (a) Figure P1.45(a), (b) Figure P1.45(b), and (c) Figure P1.45(c).

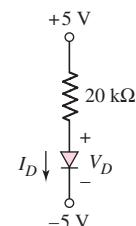


Figure P1.39

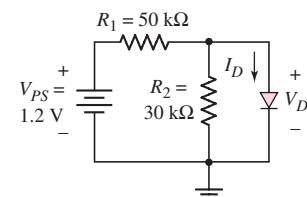


Figure P1.40

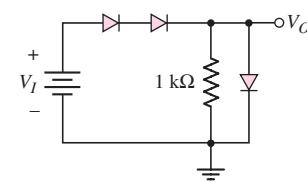


Figure P1.42

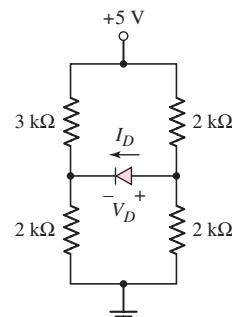


Figure P1.44

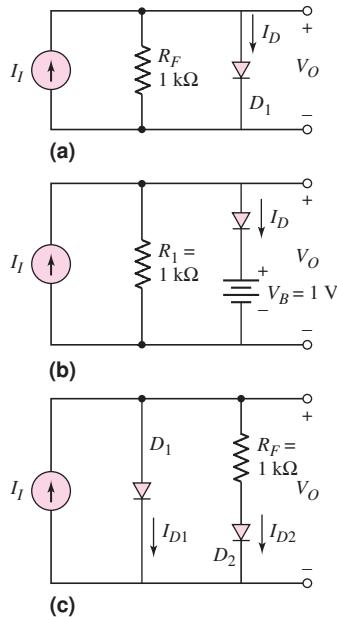


Figure P1.45

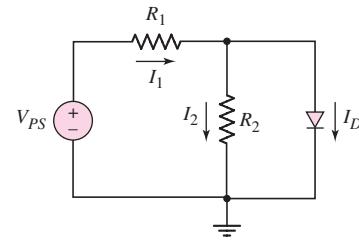


Figure P1.46

- *1.46 The cut-in voltage of the diode shown in the circuit in Figure P1.46 is $V_y = 0.7$ V. The diode is to remain biased “on” for a power supply voltage in the range $5 \leq V_{PS} \leq 10$ V. The minimum diode current is to be $I_D(\min) = 2$ mA. The maximum power dissipated in the diode is to be no more than 10 mW. Determine appropriate values of R_1 and R_2 .
- 1.47 Find I and V_O in each circuit shown in Figure P1.47 if (i) $V_y = 0.7$ V and (ii) $V_y = 0.6$ V.

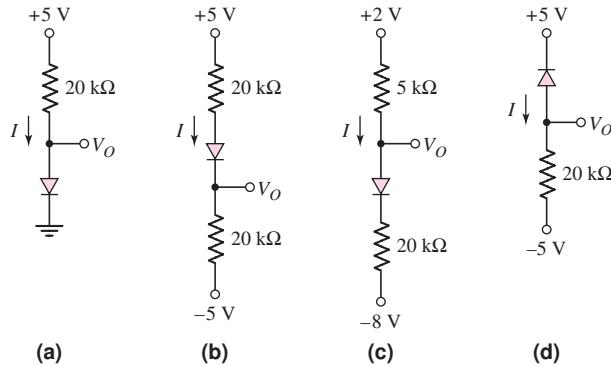


Figure P1.47

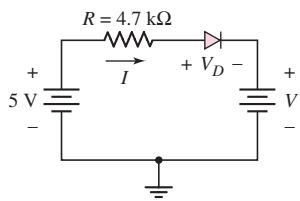


Figure P1.49

- *1.48 Repeat Problem 1.47 if the reverse-saturation current for each diode is $I_S = 5 \times 10^{-14}$ A. What is the voltage across each diode?
- 1.49 (a) In the circuit shown in Figure P1.49, find the diode voltage V_D and the supply voltage V such that the current is $I_D = 0.4$ mA. Assume the diode cut-in voltage is $V_y = 0.7$ V. (b) Using the results of part (a), determine the power dissipated in the diode.

- 1.50 Assume each diode in the circuit shown in Figure P1.50 has a cut-in voltage of $V_\gamma = 0.65$ V. (a) The input voltage is $V_I = 5$ V. Determine the value of R_1 required such that I_{D1} is one-half the value of I_{D2} . What are the values of I_{D1} and I_{D2} ? (b) If $V_I = 8$ V and $R_1 = 2\text{ k}\Omega$, determine I_{D1} and I_{D2} .

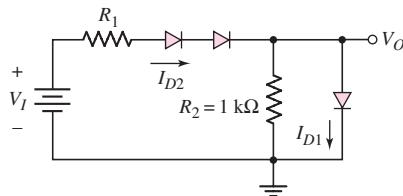


Figure P1.50

Section 1.4 Small-Signal Diode Analysis

- 1.51 (a) Consider a pn junction diode biased at $I_{DQ} = 1$ mA. A sinusoidal voltage is superimposed on V_{DQ} such that the peak-to-peak sinusoidal current is $0.05I_{DQ}$. Find the value of the applied peak-to-peak sinusoidal voltage.
 (b) Repeat part (a) if $I_{DQ} = 0.1$ mA.
- 1.52 Determine the small-signal diffusion resistance r_d for a diode biased at (a) $I_D = 26\ \mu\text{A}$, (b) $I_D = 260\ \mu\text{A}$, and (c) $I_D = 2.6$ mA.
- *1.53 The diode in the circuit shown in Figure P1.53 is biased with a constant current source I . A sinusoidal signal v_s is coupled through R_S and C . Assume that C is large so that it acts as a short circuit to the signal. (a) Show that the sinusoidal component of the diode voltage is given by

$$v_o = v_s \left(\frac{V_T}{V_T + IR_S} \right)$$

- (b) If $R_S = 260\ \Omega$, find v_o/v_s , for $I = 1\ \text{mA}$, $I = 0.1\ \text{mA}$, and $I = 0.01\ \text{mA}$.

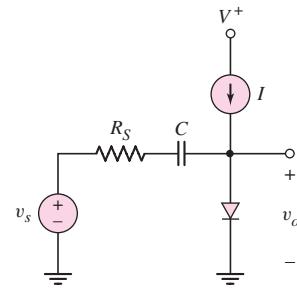


Figure P1.53

Section 1.5 Other Types of Diodes

- 1.54 The forward-bias currents in a pn junction diode and a Schottky diode are 0.72 mA. The reverse-saturation currents are $I_S = 5 \times 10^{-13}$ A and $I_S = 5 \times 10^{-8}$ A, respectively. Determine the forward-bias voltage across each diode.
- 1.55 A pn junction diode and a Schottky diode have equal cross-sectional areas and have forward-bias currents of 0.5 mA. The reverse-saturation current of the Schottky diode is $I_S = 5 \times 10^{-7}$ A. The difference in forward-bias voltages between the two diodes is 0.30 V. Determine the reverse-saturation current of the pn junction diode.
- 1.56 The reverse-saturation currents of a Schottky diode and a pn junction diode are $I_S = 5 \times 10^{-8}$ A and 10^{-12} A, respectively. (a) The diodes are connected in parallel and the parallel combination is driven by a constant current of 0.5 mA. (i) Determine the current in each diode. (ii) Determine the voltage across each diode. (b) Repeat part (a) for the diodes connected in series, with a voltage of 0.90 V connected across the series combination.

- *1.57 Consider the Zener diode circuit shown in Figure P1.57. The Zener breakdown voltage is $V_Z = 5.6\text{ V}$ at $I_Z = 0.1\text{ mA}$, and the incremental Zener resistance is $r_z = 10\Omega$. (a) Determine V_O with no load ($R_L = \infty$). (b) Find the change in the output voltage if V_{PS} changes by $\pm 1\text{ V}$. (c) Find V_O if $V_{PS} = 10\text{ V}$ and $R_L = 2\text{ k}\Omega$.

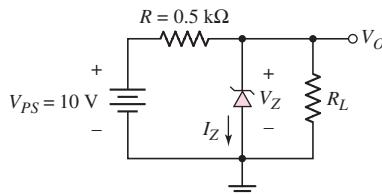


Figure P1.57

- 1.58 (a) The Zener diode in Figure P1.57 is ideal with $V_Z = 6.8\text{ V}$. Determine the maximum current and maximum power dissipated in the diode ($R_L = \infty$). (b) Determine the value of R_L such that I_Z is reduced to 0.1 of its maximum value.
- *1.59 Consider the Zener diode circuit shown in Figure P1.57. The Zener diode voltage is $V_Z = 6.8\text{ V}$ at $I_Z = 0.1\text{ mA}$ and the incremental Zener resistance is $r_z = 20\Omega$. (a) Calculate V_O with no load ($R_L = \infty$). (b) Find the change in the output voltage when a load resistance of $R_L = 1\text{ k}\Omega$ is connected.
- 1.60 The output current of a pn junction diode used as a solar cell can be given by

$$I_D = 0.2 - 5 \times 10^{-14} \left[\exp \left(\frac{V_D}{V_T} \right) - 1 \right] \quad \text{A}$$

The short-circuit current is defined as $I_{SC} = I_D$ when $V_D = 0$ and the open-circuit voltage is defined as $V_{OC} = V_D$ when $I_D = 0$. Find the values of I_{SC} and V_{OC} .

- 1.61 Using the current–voltage characteristics of the solar cell described in Problem 1.60, plot I_D versus V_D .
- 1.62 (a) Using the current–voltage characteristics of the solar cell described in problem 1.60, determine V_D when $I_D = 0.8I_{SC}$. (b) Using the results of part (a), determine the power supplied by the solar cell.



COMPUTER SIMULATION PROBLEMS

- 1.63 Use a computer simulation to generate the ideal current–voltage characteristics of a diode from a reverse-bias voltage of 5 V to a forward-bias current of 1 mA, for an I_S parameter value of (a) 10^{-15} A and (b) 10^{-13} A . Use the default values for all other parameters.
- 1.64 Use a computer simulation to find the diode current and voltage for the circuit described in Problem 1.38.

- 1.65 The reverse-saturation current for each diode in Figure P1.42 is $I_S = 10^{-14}$ A. Use a computer simulation to plot the output voltage V_O versus the input voltage V_I over the range $0 \leq V_I \leq 2.0$ V.
- 1.66 Use a computer simulation to find the diode current, diode voltage, and output voltage for each circuit shown in Figure P1.47. Assume $I_S = 10^{-13}$ A for each diode.

DESIGN PROBLEMS

[Note: Each design should be verified by a computer simulation.]

- *D1.67 Design a diode circuit to produce the load line and Q -point shown in Figure P1.67. Assume diode piecewise linear parameters of $V_\gamma = 0.7$ V and $r_f = 0$.

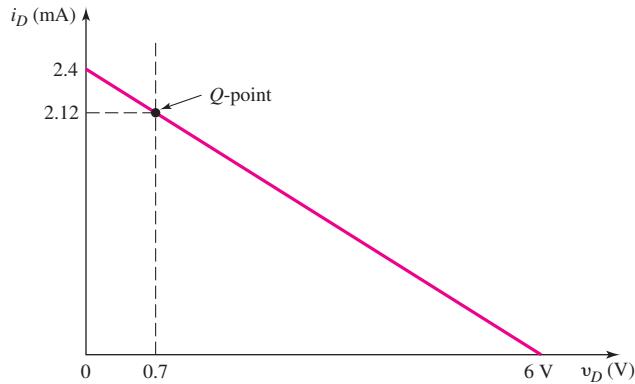


Figure P1.67

- *D1.68 Design a circuit to produce the characteristics shown in Figure P1.68, where i_D is the diode current and v_I is the input voltage. Assume diode piecewise linear parameters of $V_\gamma = 0.7$ V and $r_f = 0$.

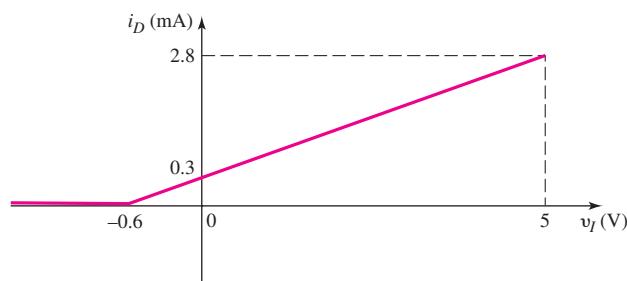


Figure P1.68

- *D1.69 Design a circuit to produce the characteristics shown in Figure P1.69, where v_O is the output voltage and v_I is the input voltage. Assume diode piecewise linear parameters of $V_\gamma = 0.7$ V and $r_f = 0$.

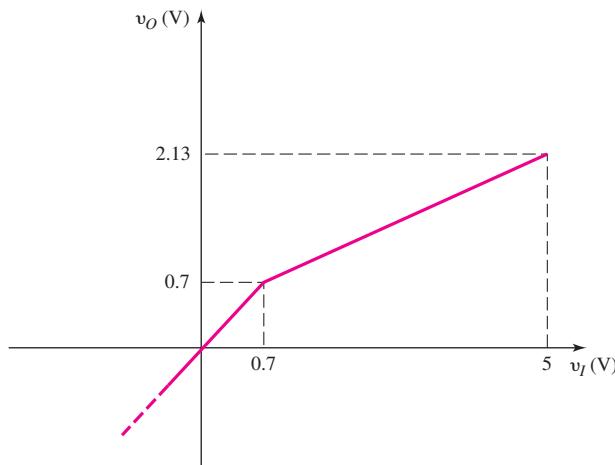


Figure P1.69

- *D1.70 Design a circuit to produce the characteristics shown in Figure P1.70, where v_O is the output voltage and v_I is the input voltage. Assume diode piecewise linear parameters of $V_\gamma = 0.7$ V and $r_f = 0$.

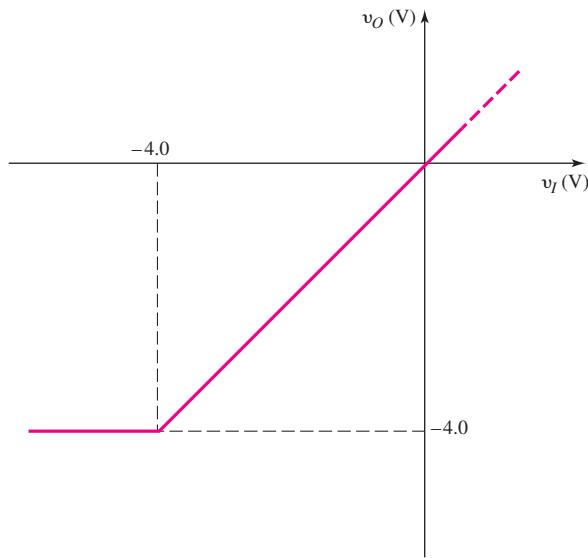


Figure P1.70

Diode Circuits

In the last chapter, we discussed some of the properties of semiconductor materials and introduced the diode. We presented the ideal current–voltage relationship of the diode and considered the piecewise linear model, which simplifies the dc analysis of diode circuits. In this chapter, the techniques and concepts developed in Chapter 1 are used to analyze and design electronic circuits containing diodes. A general goal of this chapter is to develop the ability to use the piecewise linear model and approximation techniques in the hand analysis and design of various diode circuits.

Each circuit to be considered accepts an input signal at a set of input terminals and produces an output signal at a set of output terminals. This process is called **signal processing**. The circuit “processes” the input signal and produces an output signal that is a different shape or a different function compared to the input signal. We will see in this chapter how diodes are used to perform these various signal processing functions.

Although diodes are useful electronic devices, we will begin to see the limitations of these devices and the desirability of having some type of “amplifying” device.

PREVIEW

In this chapter, we will:

- Determine the operation and characteristics of diode rectifier circuits, which, in general, form the first stage of the process of converting an ac signal into a dc signal in the electronic power supply.
- Apply the characteristics of the Zener diode to a Zener diode voltage regulator circuit.
- Apply the nonlinear characteristics of diodes to create waveshaping circuits known as clippers and clampers.
- Examine the techniques used to analyze circuits that contain more than one diode.
- Understand the operation and characteristics of specialized photodiode and light-emitting diode circuits.
- Design a basic dc power supply incorporating a filtered rectifier circuit and a Zener diode.

2.1 RECTIFIER CIRCUITS

Objective: • Determine the operation and characteristics of diode rectifier circuits, which form the first stage in the process of converting an ac signal into a dc signal in the electronic power supply.

One application of diodes is in the design of rectifier circuits. A diode rectifier forms the first stage of a dc power supply. A dc voltage is required to power essentially every electronic device, including personal computers, televisions, and stereo systems. An electrical cord that is plugged into a wall socket and attached to a television, for example, is connected to a rectifier circuit inside the TV. In addition, battery chargers for portable electronic devices such as cell phones and laptop computers contain rectifier circuits.

Figure 2.1 is a diagram of a dc power supply. The output voltage¹ v_O is usually in the range of 3 to 24 V depending on the particular electronics application. Throughout the first part of this chapter, we will analyze and design various stages in the power supply circuit.

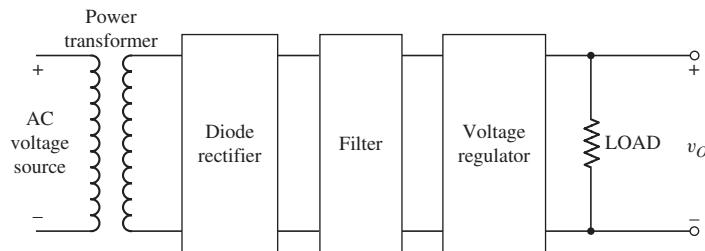


Figure 2.1 Diagram of an electronic power supply. The circuits that characterize each block diagram are considered in this chapter.

Rectification is the process of converting an alternating (ac) voltage into one that is limited to one polarity. The diode is useful for this function because of its nonlinear characteristics, that is, current exists for one voltage polarity, but is essentially zero for the opposite polarity. Rectification is classified as **half-wave** or **full-wave**, with half-wave being the simpler and full-wave being more efficient.

2.1.1 Half-Wave Rectification

Figure 2.2(a) shows a power transformer with a diode and resistor connected to the secondary of the transformer. We will use the piecewise linear approach in analyzing this circuit, assuming the diode forward resistance is $r_f = 0$ when the diode is “on.”

The input signal, v_I , is, in general, a 120 V (rms), 60 Hz ac signal. Recall that the secondary voltage, v_S , and primary voltage, v_I , of an ideal transformer are related by

$$\frac{v_I}{v_S} = \frac{N_1}{N_2} \quad (2.1)$$

¹Ideally, the output voltage of a rectifier circuit is a dc voltage. However, as we will see, there may be an ac ripple voltage superimposed on a dc value. For this reason, we will use the notation v_O as the instantaneous value of output voltage.

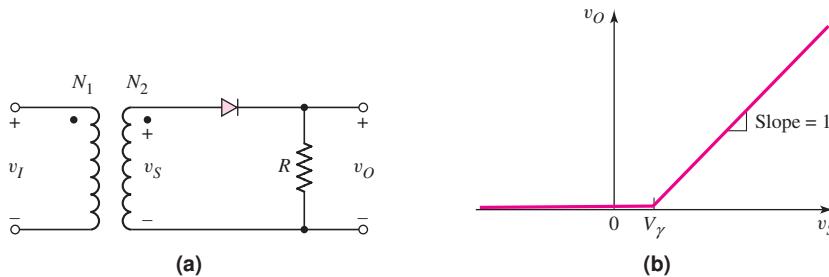


Figure 2.2 Half-wave rectifier (a) circuit and (b) voltage transfer characteristics

where N_1 and N_2 are the number of primary and secondary turns, respectively. The ratio N_1/N_2 is called the **transformer turns ratio**. The transformer turns ratio will be designed to provide a particular secondary voltage, v_S , which in turn will produce a particular output voltage v_O .

Problem-Solving Technique: Diode Circuits

In using the piecewise linear model of the diode, the first objective is to determine the linear region (conducting or not conducting) in which the diode is operating. To do this, we can:

1. Determine the input voltage condition such that a diode is conducting (on). Then find the output signal for this condition.
2. Determine the input voltage condition such that a diode is not conducting (off). Then find the output signal for this condition.

[Note: Item 2 can be performed before item 1 if desired.]

Figure 2.2(b) shows the voltage transfer characteristics, v_O versus v_S , for the circuit. For $v_S < 0$, the diode is reverse biased, which means that the current is zero and the output voltage is zero. As long as $v_S < V_\gamma$, the diode will be nonconducting, so the output voltage will remain zero. When $v_S > V_\gamma$, the diode becomes forward biased and a current is induced in the circuit. In this case, we can write

$$i_D = \frac{v_S - V_\gamma}{R} \quad (2.2(a))$$

and

$$v_O = i_D R = v_S - V_\gamma \quad (2.2(b))$$

For $v_S > V_\gamma$, the slope of the transfer curve is 1.

If v_S is a sinusoidal signal, as shown in Figure 2.3(a), the output voltage can be found using the voltage transfer curve in Figure 2.2(b). For $v_S \leq V_\gamma$ the output voltage is zero; for $v_S > V_\gamma$, the output is given by Equation (2.2(b)), or

$$v_O = v_S - V_\gamma$$

and is shown in Figure 2.3(b). We can see that while the input signal v_S alternates polarity and has a time-average value of zero, the output voltage v_O is unidirectional and has an average value that is not zero. The input signal is therefore rectified. Also, since the output voltage appears only during the positive cycle of the input signal, the circuit is called a **half-wave rectifier**.

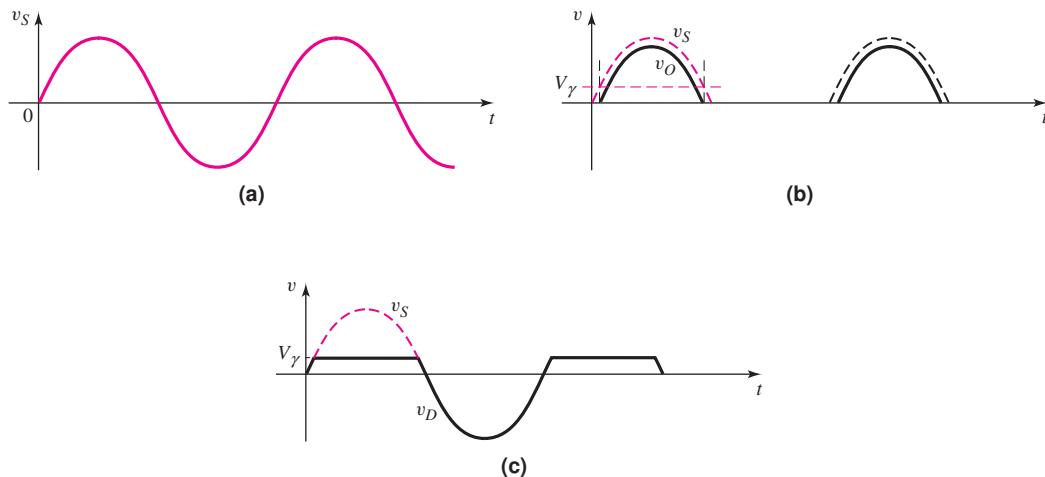


Figure 2.3 Signals of the half-wave rectifier circuit: (a) sinusoidal input voltage, (b) rectified output voltage, and (c) diode voltage

When the diode is cut off and nonconducting, no voltage drop occurs across the resistor R ; therefore, the entire input signal voltage appears across the diode (Figure 2.3(c)). Consequently, the diode must be capable of handling the peak current in the forward direction and sustaining the largest peak inverse voltage (PIV) without breakdown. For the circuit shown in Figure 2.2(a), the value of PIV is equal to the peak value of v_S .

We can use a half-wave rectifier circuit to charge a battery as shown in Figure 2.4(a). Charging current exists whenever the instantaneous ac source voltage is greater than the battery voltage plus the diode cut-in voltage as shown in Figure 2.4(b). The resistance R in the circuit is to limit the current. When the ac source voltage is less than V_B , the current is zero. Thus current flows only in the direction to charge the battery. One disadvantage of the half-wave rectifier is that we “waste” the negative half-cycles. The current is zero during the negative half-cycles, so there is no energy dissipated, but at the same time, we are not making use of any possible available energy.

EXAMPLE 2.1

Objective: Determine the currents and voltages in a half-wave rectifier circuit.

Consider the circuit shown in Figure 2.4. Assume $V_B = 12$ V, $R = 100 \Omega$, and $V_\gamma = 0.6$ V. Also assume $v_S(t) = 24 \sin \omega t$. Determine the peak diode current,

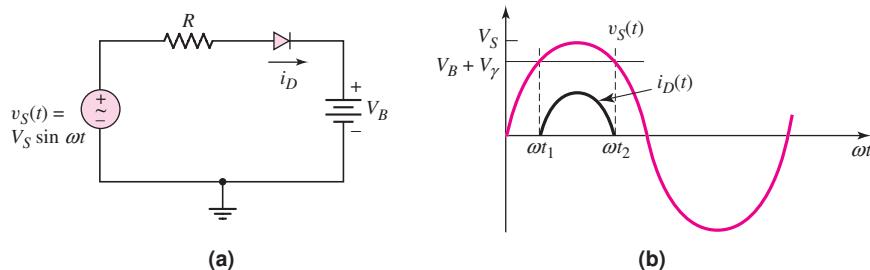


Figure 2.4 (a) Half-wave rectifier used as a battery charger; (b) input voltage and diode current waveforms

maximum reverse-bias diode voltage, and the fraction of the cycle over which the diode is conducting.

Solution: Peak diode current:

$$i_D(\text{peak}) = \frac{V_S - V_B - V_\gamma}{R} = \frac{24 - 12 - 0.6}{0.10} = 114 \text{ mA}$$

Maximum reverse-bias diode voltage:

$$v_R(\text{max}) = V_S + V_B = 24 + 12 = 36 \text{ V}$$

Diode conduction cycle:

$$v_I = 24 \sin \omega t_1 = 12.6$$

or

$$\omega t_1 = \sin^{-1} \left(\frac{12.6}{24} \right) \Rightarrow 31.7^\circ$$

By symmetry,

$$\omega t_2 = 180 - 31.7 = 148.3^\circ$$

Then

$$\text{Percent time} = \frac{148.3 - 31.7}{360} \times 100\% = 32.4\%$$

Comment: This example shows that the diode conducts only approximately one-third of the time, which means that the efficiency of this battery charger is quite low.

EXERCISE PROBLEM

Ex 2.1: Repeat Example 2.1 if the input voltage is $v_s(t) = 12 \sin \omega t$ (V), $V_B = 4.5$ V, and $R = 250 \Omega$. (Ans. $i_D(\text{peak}) = 27.6$ mA, $v_R(\text{max}) = 16.5$ V, 36.0 %)

2.1.2 Full-Wave Rectification

The full-wave rectifier inverts the negative portions of the sine wave so that a unipolar output signal is generated during both halves of the input sinusoid. One example of a full-wave rectifier circuit appears in Figure 2.5(a). The input to the rectifier consists of a power transformer, in which the input is normally a 120 V (rms), 60 Hz ac signal, and the two outputs are from a center-tapped secondary winding that provides equal voltages v_S , with the polarities shown. When the input line voltage is positive, both output signal voltages v_S are also positive.

The primary winding connected to the 120 V ac source has N_1 windings, and each half of the secondary winding has N_2 windings. The value of the v_S output voltage is $120 (N_2/N_1)$ volts (rms). The **turns ratio** of the transformer, usually designated (N_1/N_2) can be designed to “step down” the input line voltage to a value that will produce a particular dc output voltage from the rectifier.

The input power transformer also provides electrical isolation between the powerline circuit and the electronic circuits to be biased by the rectifier circuit. This isolation reduces the risk of electrical shock.

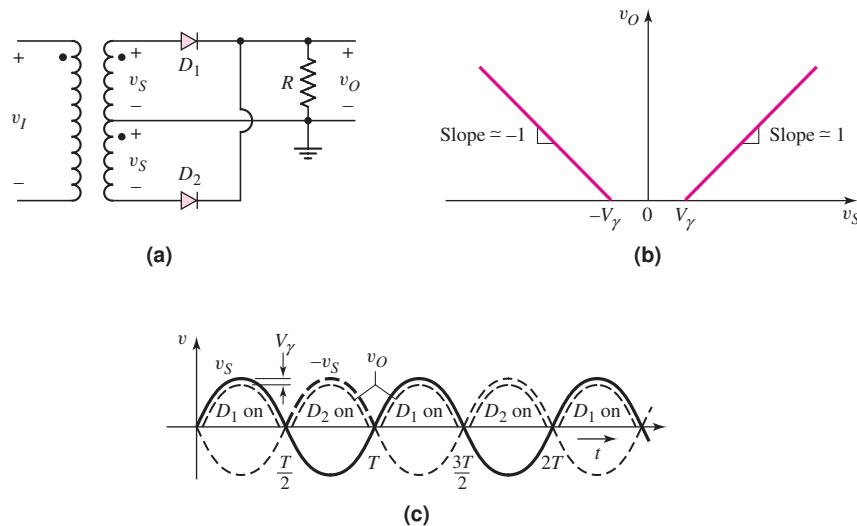


Figure 2.5 Full-wave rectifier: (a) circuit with center-tapped transformer, (b) voltage transfer characteristics, and (c) input and output waveforms

During the positive half of the input voltage cycle, both output voltages v_S are positive; therefore, diode D_1 is forward biased and conducting and D_2 is reverse biased and cut off. The current through D_1 and the output resistance produce a positive output voltage. During the negative half cycle, D_1 is cut off and D_2 is forward biased, or “on,” and the current through the output resistance again produces a positive output voltage. If we assume that the forward diode resistance r_f of each diode is small and negligible, we obtain the voltage transfer characteristics, v_O versus v_S , shown in Figure 2.5(b).

For a sinusoidal input voltage, we can determine the output voltage versus time by using the voltage transfer curve shown in Figure 2.5(b). When $v_S > V_\gamma$, D_1 is on and the output voltage is $v_O = v_S - V_\gamma$. When v_S is negative, then for $v_S < -V_\gamma$ or $-v_S > V_\gamma$, D_2 is on and the output voltage is $v_O = -v_S - V_\gamma$. The corresponding input and output voltage signals are shown in Figure 2.5(c). Since a rectified output voltage occurs during both the positive and negative cycles of the input signal, this circuit is called a **full-wave rectifier**.

Another example of a full-wave rectifier circuit appears in Figure 2.6(a). This circuit is a **bridge rectifier**, which still provides electrical isolation between the input ac powerline and the rectifier output, but does not require a center-tapped secondary winding. However, it does use four diodes, compared to only two in the previous circuit.

During the positive half of the input voltage cycle, v_S is positive, D_1 and D_2 are forward biased, D_3 and D_4 are reverse biased, and the direction of the current is as shown in Figure 2.6(a). During the negative half-cycle of the input voltage, v_S is negative, and D_3 and D_4 are forward biased. The direction of the current, shown in Figure 2.6(b), produces the same output voltage polarity as before.

Figure 2.6(c) shows the sinusoidal voltage v_S and the rectified output voltage v_O . Because two diodes are in series in the conduction path, the magnitude of v_O is two diode drops less than the magnitude of v_S .

One difference to be noted in the bridge rectifier circuit in Figure 2.6(a) and the rectifier in Figure 2.5(a) is the ground connection. Whereas the center tap of the secondary winding of the circuit in Figure 2.5(a) is at ground potential, the secondary

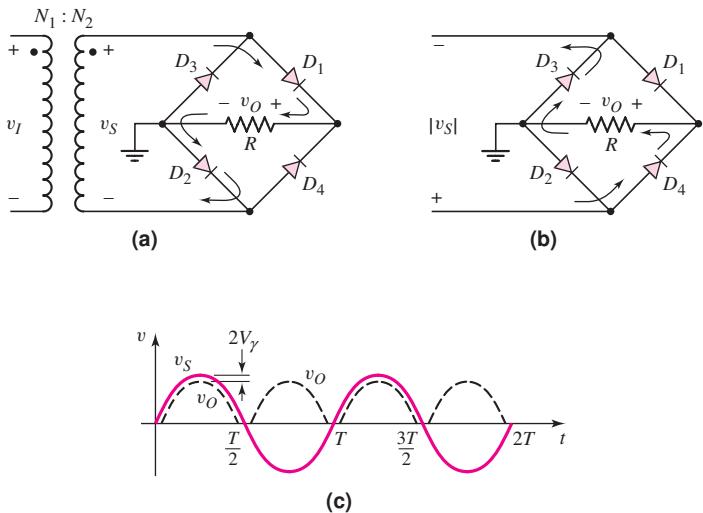


Figure 2.6 A full-wave bridge rectifier: (a) circuit showing the current direction for a positive input cycle, (b) current direction for a negative input cycle, and (c) input and output voltage waveforms

winding of the bridge circuit (Figure 2.6(a)) is not directly grounded. One side of the load R is grounded, but the secondary of the transformer is not.

EXAMPLE 2.2

Objective: Compare voltages and the transformer turns ratio in two full-wave rectifier circuits.

Consider the rectifier circuits shown in Figures 2.5(a) and 2.6(a). Assume the input voltage is from a 120 V (rms), 60 Hz ac source. The desired peak output voltage v_O is 9 V, and the diode cut-in voltage is assumed to be $V_\gamma = 0.7$ V.

Solution: For the center-tapped transformer circuit shown in Figure 2.5(a), a peak voltage of $v_O(\text{max}) = 9$ V means that the peak value of v_S is

$$v_S(\text{max}) = v_O(\text{max}) + V_\gamma = 9 + 0.7 = 9.7 \text{ V}$$

For a sinusoidal signal, this produces an rms value of

$$v_{S,\text{rms}} = \frac{9.7}{\sqrt{2}} = 6.86 \text{ V}$$

The turns ratio of the primary to each secondary winding must then be

$$\frac{N_1}{N_2} = \frac{120}{6.86} \cong 17.5$$

For the bridge circuit shown in Figure 2.6(a), a peak voltage of $v_O(\text{max}) = 9$ V means that the peak value of v_S is

$$v_S(\text{max}) = v_O(\text{max}) + 2V_\gamma = 9 + 2(0.7) = 10.4 \text{ V}$$

For a sinusoidal signal, this produces an rms value of

$$v_{S,\text{rms}} = \frac{10.4}{\sqrt{2}} = 7.35 \text{ V}$$

The turns ratio should then be

$$\frac{N_1}{N_2} = \frac{120}{7.35} \cong 16.3$$

For the center-tapped rectifier, the peak inverse voltage (PIV) of a diode is

$$\text{PIV} = v_R(\text{max}) = 2v_S(\text{max}) - V_\gamma = 2(9.7) - 0.7 = 18.7 \text{ V}$$

For the bridge rectifier, the peak inverse voltage of a diode is

$$\text{PIV} = v_R(\text{max}) = v_S(\text{max}) - V_\gamma = 10.4 - 0.7 = 9.7 \text{ V}$$

Comment: These calculations demonstrate the advantages of the bridge rectifier over the center-tapped transformer circuit. First, only half as many turns are required for the secondary winding in the bridge rectifier. This is true because only half of the secondary winding of the center-tapped transformer is utilized at any one time. Second, for the bridge circuit, the peak inverse voltage that any diode must sustain without breakdown is only half that of the center-tapped transformer circuit.

EXERCISE PROBLEM

Ex 2.2: Consider the bridge circuit shown in Figure 2.6(a) with an input voltage $v_S = V_M \sin \omega t$. Assume a diode cut-in voltage of $V_\gamma = 0.7 \text{ V}$. Determine the fraction (percent) of time that the diode D_1 is conducting for peak sinusoidal voltages of (a) $V_M = 12 \text{ V}$ and (b) $V_M = 4 \text{ V}$. (Ans. (a) 46.3% (b) 38.6%)

Because of the advantages demonstrated in Example 2.2 the bridge rectifier circuit is used more often than the center-tapped transformer circuit.

Both full-wave rectifier circuits discussed (Figures 2.5 and 2.6) produce a positive output voltage. As we will see in the next chapter discussing transistor circuits, there are times when a negative dc voltage is also required. We can produce negative rectification by reversing the direction of the diodes in either circuit. Figure 2.7(a) shows the bridge circuit with the diodes reversed compared to those in Figure 2.6. The direction of current is shown during the positive half cycle of v_S . The output voltage is now negative with respect to ground potential. During the negative half cycle of v_S , the complementary diodes turn on and the direction of current through the load is the same, producing a negative output voltage. The input and output voltages are shown in Figure 2.7(b).

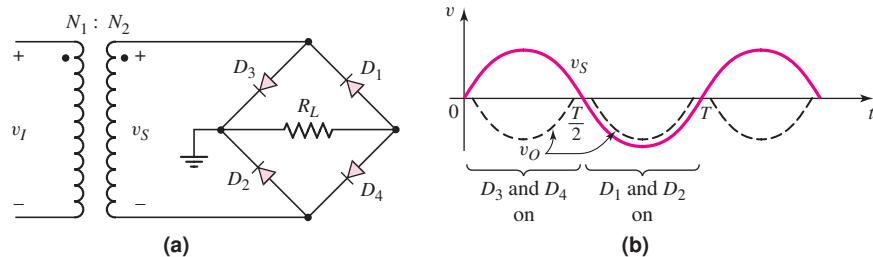


Figure 2.7 (a) Full-wave bridge rectifier circuit to produce negative output voltages.
(b) Input and output waveforms.

2.1.3 Filters, Ripple Voltage, and Diode Current

If a capacitor is added in parallel with the load resistor of a half-wave rectifier to form a simple filter circuit (Figure 2.8(a)), we can begin to transform the half-wave sinusoidal output into a dc voltage. Figure 2.8(b) shows the positive half of the output sine wave, and the beginning portion of the voltage across the capacitor, assuming the capacitor is initially uncharged. If we assume that the diode forward resistance is $r_f = 0$, which means that the $r_f C$ time constant is zero, the voltage across the capacitor follows this initial portion of the signal voltage. When the signal voltage reaches its peak and begins to decrease, the voltage across the capacitor also starts to decrease, which means the capacitor starts to discharge. The only discharge current path is through the resistor. If the RC time constant is large, the voltage across the capacitor discharges exponentially with time (Figure 2.8(c)). During this time period, the diode is cut off.

A more detailed analysis of the circuit response when the input voltage is near its peak value indicates a subtle difference between actual circuit operation and the qualitative description. If we assume that the diode turns off immediately when the input voltage starts to decrease from its peak value, then the output voltage will decrease exponentially with time, as previously indicated. An exaggerated sketch of these two voltages is shown in Figure 2.8(d). The output voltage decreases at a faster rate than the input voltage, which means that at time t_1 , the difference between v_I and

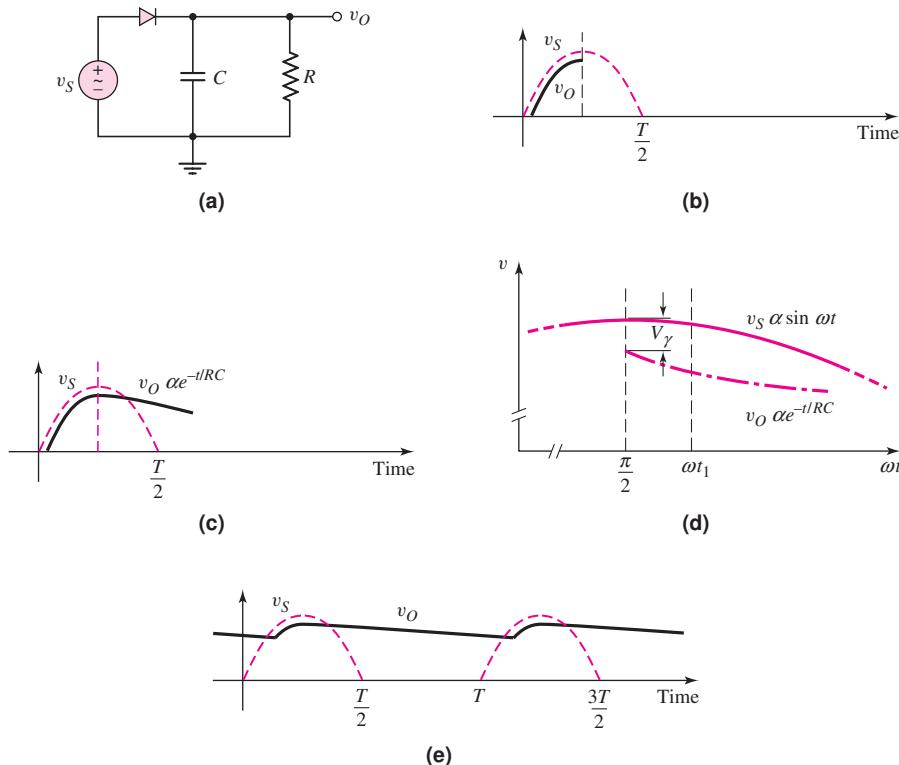


Figure 2.8 Simple filter circuit: (a) half-wave rectifier with an RC filter, (b) positive input voltage and initial portion of output voltage, (c) output voltage resulting from capacitor discharge, (d) expanded view of input and output voltages assuming capacitor discharge begins at $\omega t = \pi/2$, and (e) steady-state input and output voltages

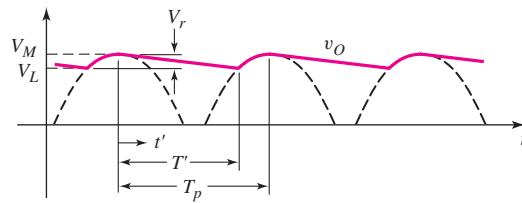


Figure 2.9 Output voltage of a full-wave rectifier with an RC filter showing the ripple voltage

v_O , that is, the voltage across the diode, is greater than V_γ . However, this condition cannot exist. Therefore, the diode does not turn off immediately. If the RC time constant is large, there is only a small difference between the time of the peak input voltage and the time the diode turns off. In this situation, a computer analysis may provide more accurate results than an approximate hand analysis.

During the next positive cycle of the input voltage, there is a point at which the input voltage is greater than the capacitor voltage, and the diode turns back on. The diode remains on until the input reaches its peak value and the capacitor voltage is completely recharged.

Since the capacitor filters out a large portion of the sinusoidal signal, it is called a **filter capacitor**. The steady-state output voltage of the RC filter is shown in Figure 2.8(e).

The ripple effect in the output from a full-wave filtered rectifier circuit can be seen in the output waveform in Figure 2.9. The capacitor charges to its peak voltage value when the input signal is at its peak value. As the input decreases, the diode becomes reverse biased and the capacitor discharges through the output resistance R . Determining the ripple voltage is necessary for the design of a circuit with an acceptable amount of ripple.

To a good approximation, the output voltage, that is, the voltage across the capacitor or the RC circuit, can be written as

$$v_O(t) = V_M e^{-t'/\tau} = V_M e^{-t'/RC} \quad (2.3)$$

where t' is the time after the output has reached its peak value, and RC is the time constant of the circuit.

The smallest output voltage is

$$V_L = V_M e^{-T'/RC} \quad (2.4)$$

where T' is the discharge time, as indicated in Figure 2.9.

The **ripple voltage** V_r is defined as the difference between V_M and V_L , and is determined by

$$V_r = V_M - V_L = V_M (1 - e^{-T'/RC}) \quad (2.5)$$

Normally, we will want the discharge time T' to be small compared to the time constant, or $T' \ll RC$. Expanding the exponential in a series and keeping only the linear terms of that expansion, we have the approximation²

$$e^{-T'/RC} \cong 1 - \frac{T'}{RC} \quad (2.6)$$

²We can show that the difference between the exponential function and the linear approximation given by Equation (2.6) is less than 0.5 percent for $RC = 10T'$. We need a relatively large RC time constant for this application.

The ripple voltage can now be written as

$$V_r \cong V_M \left(\frac{T'}{RC} \right) \quad (2.7)$$

Since the discharge time T' depends on the RC time constant, Equation (2.7) is difficult to solve. However, if the ripple effect is small, then as an approximation, we can let $T' = T_p$, so that

$$V_r \cong V_M \left(\frac{T_p}{RC} \right) \quad (2.8)$$

where T_p is the time between peak values of the output voltage. For a full-wave rectifier, T_p is one-half the signal period. Therefore, we can relate T_p to the signal frequency,

$$f = \frac{1}{2T_p}$$

The ripple voltage is then

$$V_r = \frac{V_M}{2fRC} \quad (2.9)$$

For a half-wave rectifier, the time T_p corresponds to a full period (not a half period) of the signal, so the factor 2 does not appear in Equation (2.9). The factor of 2 shows that the full-wave rectifier has half the ripple voltage of the half-wave rectifier.

Equation (2.9) can be used to determine the capacitor value required for a particular ripple voltage.

EXAMPLE 2.3

Objective: Determine the capacitance required to yield a particular ripple voltage.

Consider a full-wave rectifier circuit with a 60 Hz input signal and a peak output voltage of $V_M = 10$ V. Assume the output load resistance is $R = 10$ k Ω and the ripple voltage is to be limited to $V_r = 0.2$ V.

Solution: From Equation (2.9), we can write

$$C = \frac{V_M}{2fRV_r} = \frac{10}{2(60)(10 \times 10^3)(0.2)} \Rightarrow 41.7 \mu\text{F}$$

Comment: If the ripple voltage is to be limited to a smaller value, a larger filter capacitor must be used. Note that the size of the ripple voltage and the size of filter capacitor are related to the load resistance R .

EXERCISE PROBLEM

Ex 2.3: Assume the input signal to a rectifier circuit has a peak value of $V_M = 12$ V and is at a frequency of 60 Hz. Assume the output load resistance is $R = 2$ k Ω and the ripple voltage is to be limited to $V_r = 0.4$ V. Determine the capacitance required to yield this specification for a (a) full-wave rectifier and (b) half-wave rectifier. (Ans. (a) $C = 125 \mu\text{F}$, (b) $C = 250 \mu\text{F}$).

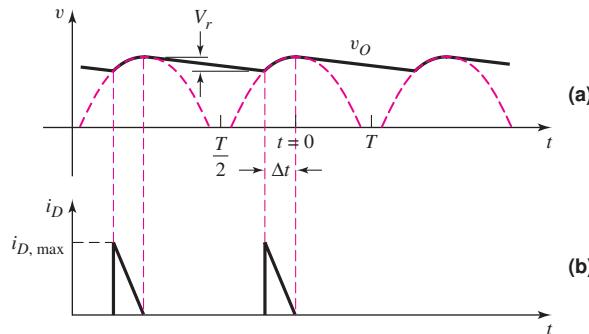


Figure 2.10 Output of a full-wave rectifier with an RC filter: (a) diode conduction time and (b) diode current

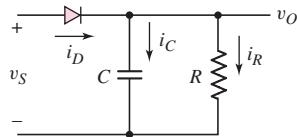


Figure 2.11 Equivalent circuit of a full-wave rectifier during capacitor charging cycle

The diode in a filtered rectifier circuit conducts for a brief interval Δt near the peak of the sinusoidal input signal. The diode current supplies the charge lost by the capacitor during the discharge time. Figure 2.10 shows the rectified output of a full-wave rectifier and the filtered output assuming ideal diodes ($V_y = 0$) in the rectifier circuit. We will use this approximate model to estimate the diode current during the diode conduction time. Figure 2.11 shows the equivalent circuit of the full-wave rectifier during the charging time. We see that

$$i_D = i_C + i_R = C \frac{dv_O}{dt} + \frac{v_O}{R} \quad (2.10)$$

During the diode conduction time near $t = 0$ (Figure 2.10), we can write

$$v_O = V_M \cos \omega t \quad (2.11)$$

For small ripple voltages, the diode conduction time is small, so we can approximate the output voltage as

$$v_O = V_M \cos \omega t \cong V_M \left[1 - \frac{1}{2} (\omega t)^2 \right] \quad (2.12)$$

The charging current through the capacitor is

$$i_C = C \frac{dv_O}{dt} = CV_M \left[-\frac{1}{2} (2)(\omega t)(\omega) \right] = -\omega CV_M \omega t \quad (2.13)$$

From Figure 2.10, the diode conduction occurs during the time $-\Delta t < t < 0$, so that the capacitor current is positive and is a linear function of time. We note that at $t = 0$, the capacitor current is $i_C = 0$. At $t = -\Delta t$, the capacitor charging current is at a peak value and is given by

$$i_{C,\text{peak}} = -\omega CV_M [\omega(-\Delta t)] = +\omega CV_M \omega \Delta t \quad (2.14)$$

The capacitor current during the diode charging time is approximately triangular and is shown in Figure 2.10(b).

From Equation (2.11), we can write that the voltage V_L is given by

$$V_L = V_M \cos[\omega(-\Delta t)] \cong V_M \left[1 - \frac{1}{2} (\omega \Delta t)^2 \right] \quad (2.15)$$

Solving for $\omega\Delta t$, we find

$$\omega\Delta t = \sqrt{\frac{2V_r}{V_M}} \quad (2.16)$$

where $V_r = V_M - V_L$.

From Equation (2.9), we can write

$$fC = \frac{V_M}{2RV_r} \quad (2.17(a))$$

or

$$2\pi fC = \omega C = \frac{\pi V_M}{RV_r} \quad (2.17(b))$$

Substituting Equations (2.17(b)) and (2.16) into Equation (2.14), we have

$$i_{C,\text{peak}} = \left(\frac{\pi V_M}{RV_r} \right) V_M \left(\sqrt{\frac{2V_r}{V_M}} \right) \quad (2.18(a))$$

or

$$i_{C,\text{peak}} = \pi \frac{V_M}{R} \sqrt{\frac{2V_M}{V_r}} \quad (2.18(b))$$

Since the charging current through the capacitor is triangular, we have that the average capacitor current during the diode charging time is

$$i_{C,\text{avg}} = \frac{\pi}{2} \frac{V_M}{R} \sqrt{\frac{2V_M}{V_r}} \quad (2.19)$$

During the capacitor charging time, there is still a current through the load. This current is also being supplied through the diode. Neglecting the ripple voltage, the load current is approximately

$$i_L \cong \frac{V_M}{R} \quad (2.20)$$

Therefore, the peak diode current during the diode conduction time for a full-wave rectifier is approximately

$$i_{D,\text{peak}} \cong \frac{V_M}{R} \left(1 + \pi \sqrt{\frac{2V_M}{V_r}} \right) \quad (2.21)$$

and the average diode current during the diode conduction time is

$$i_{D,\text{avg}} \cong \frac{V_M}{R} \left(1 + \frac{\pi}{2} \sqrt{\frac{2V_M}{V_r}} \right) \quad (2.22)$$

The average diode current over the entire input signal period is

$$i_D(\text{avg}) = \frac{V_M}{R} \left(1 + \frac{\pi}{2} \sqrt{\frac{2V_M}{V_r}} \right) \frac{\Delta t}{T} \quad (2.23)$$

For the full-wave rectifier, we have $1/2T = f$, so

$$\Delta t = \frac{1}{\omega} \sqrt{\frac{2V_r}{V_M}} = \frac{1}{2\pi f} \sqrt{\frac{2V_r}{V_M}} \quad (2.24(a))$$

Then

$$\frac{\Delta t}{T} = \frac{1}{2\pi f} \sqrt{\frac{2V_r}{V_M}} 2f = \frac{1}{\pi} \sqrt{\frac{2V_r}{V_M}} \quad (2.24(b))$$

Then the average current through the diode during the entire cycle for a full-wave rectifier is

$$i_D(\text{avg}) = \frac{1}{\pi} \sqrt{\frac{2V_r}{V_M}} \frac{V_M}{R} \left(1 + \frac{\pi}{2} \sqrt{\frac{2V_M}{V_r}} \right) \quad (2.25)$$

DESIGN EXAMPLE 2.4

Objective: Design a full-wave rectifier to meet particular specifications.

A full-wave rectifier is to be designed to produce a peak output voltage of 12 V, deliver 120 mA to the load, and produce an output with a ripple of not more than 5 percent. An input line voltage of 120 V (rms), 60 Hz is available.

Solution: A full-wave bridge rectifier will be used, because of the advantages previously discussed. The effective load resistance is

$$R = \frac{V_O}{I_L} = \frac{12}{0.12} = 100 \Omega$$

Assuming a diode cut-in voltage of 0.7 V, the peak value of v_S is

$$v_S(\text{max}) = v_O(\text{max}) + 2V_\gamma = 12 + 2(0.7) = 13.4 \text{ V}$$

For a sinusoidal signal, this produces an rms voltage value of

$$v_{S,\text{rms}} = \frac{13.4}{\sqrt{2}} = 9.48 \text{ V}$$

The transformer turns ratio is then

$$\frac{N_1}{N_2} = \frac{120}{9.48} = 12.7$$

For a 5 percent ripple, the ripple voltage is

$$V_r = (0.05)V_M = (0.05)(12) = 0.6 \text{ V}$$

The required filter capacitor is found to be

$$C = \frac{V_M}{2fRV_r} = \frac{12}{2(60)(100)(0.6)} \Rightarrow 1667 \mu\text{F}$$

The peak diode current, from Equation (2.21), is

$$i_{D,\text{peak}} = \frac{12}{100} \left[1 + \pi \sqrt{\frac{2(12)}{0.6}} \right] = 2.50 \text{ A}$$

and the average diode current over the entire signal period, from Equation (2.25), is

$$i_D(\text{avg}) = \frac{1}{\pi} \sqrt{\frac{2(0.6)}{12}} \left(\frac{12}{100} \right) \left(1 + \frac{\pi}{2} \sqrt{\frac{2(12)}{0.6}} \right) \Rightarrow 132 \text{ mA}$$

Finally, the peak inverse voltage that each diode must sustain is

$$\text{PIV} = v_R(\text{max}) = v_S(\text{max}) - V_\gamma = 13.4 - 0.7 = 12.7 \text{ V}$$

Comment: The minimum specifications for the diodes in this full-wave rectifier circuit are: a peak current of 2.50 A, an average current of 132 mA, and a peak inverse voltage of 12.7 V. In order to meet the desired ripple specification, the required filter capacitance must be large, since the effective load resistance is small.

Design Pointer: (1) A particular turns ratio was determined for the transformer. However, this particular transformer design is probably not commercially available. This means an expensive custom transformer design would be required, or if a standard transformer is used, then additional circuit design is required to meet the output voltage specification. (2) A constant 120 V (rms) input voltage is assumed to be available. However, this voltage can fluctuate, so the output voltage will also fluctuate.

We will see later how more sophisticated designs will solve these two problems.

Computer Verification: Since we simply used an assumed cut-in voltage for the diode and used approximations in the development of the ripple voltage equations, we can use PSpice to give us a more accurate evaluation of the circuit. The PSpice circuit schematic and the steady-state output voltage are shown in Figure 2.12. We see that the peak output voltage is 11.6 V, which is close to the desired 12 V. One reason for the slight discrepancy is that the diode voltage drop for the maximum input voltage is slightly greater than 0.8 V rather than the assumed 0.7 V. The ripple voltage is approximately 0.5 V, which is within the 0.6 V specification.

Discussion: In the PSpice simulation, a standard diode, 1N4002, was used. In order for the computer simulation to be valid, the diode used in the simulation and in the actual circuit must match. In this example, to reduce the diode voltage and increase the peak output voltage, a diode with a larger cross-sectional area should be used.

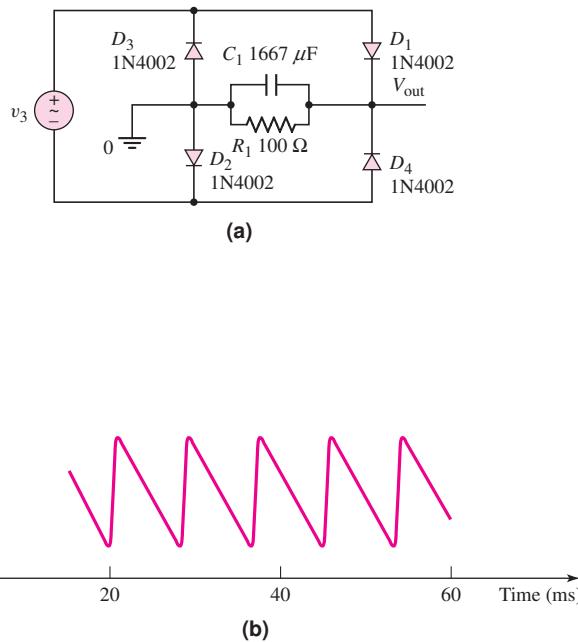


Figure 2.12 (a) PSpice circuit schematic of diode bridge circuit with an RC filter; (b) Steady-state output voltage of PSpice analysis of diode bridge circuit for a 60 Hz input sine wave with a peak value of 13.4 V

EXERCISE PROBLEM

Ex 2.4: The input voltage to the half-wave rectifier in Figure 2.8(a) is $v_S = 75 \sin[2\pi(60)t]$ V. Assume a diode cut-in voltage of $V_\gamma = 0$. The ripple voltage is to be no more than $V_r = 4$ V. If the filter capacitor is $50 \mu\text{F}$, determine the minimum load resistance that can be connected to the output. (Ans. $R = 6.25 \text{ k}\Omega$)

2.1.4 Detectors

One of the first applications of semiconductor diodes was as a detector for amplitude-modulated (AM) radio signals. An amplitude-modulated signal consists of a radio-frequency carrier wave whose amplitude varies with an audio frequency as shown in Figure 2.13(a). The detector circuit is shown in Figure 2.13(b) and is a half-wave rec-

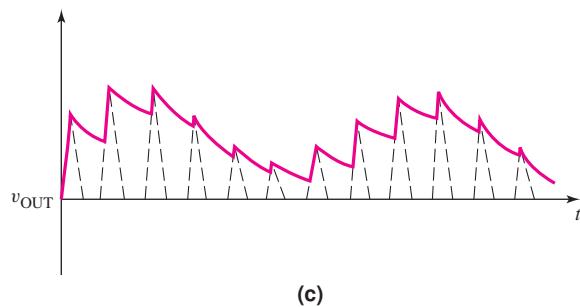
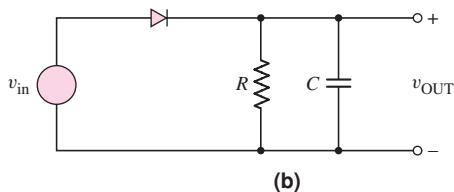
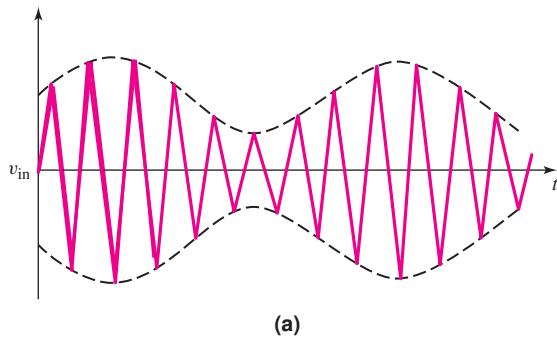


Figure 2.13 The signals and circuit for demodulation of an amplitude-modulated signal. (a) The amplitude-modulated input signal. (b) The detector circuit. (c) The demodulated output signal.

tifier circuit with an RC filter on the output. For this application, the RC time constant should be approximately equal to the period of the carrier signal, so that the output voltage can follow each peak value of the carrier signal. If the time constant is too large, the output will not be able to change fast enough and the output will not represent the audio output. The output of the detector is shown in Figure 2.13(c).

The output of the detector circuit is then coupled to an amplifier through a capacitor to remove the dc component of the signal, and the output of the amplifier is then fed to a speaker.

2.1.5 Voltage Doubler Circuit

A **voltage doubler circuit** is very similar to the full-wave rectifier, except that two diodes are replaced by capacitors, and it can produce a voltage equal to approximately twice the peak output of a transformer (Figure 2.14).

Figure 2.15(a) shows the equivalent circuit when the voltage polarity at the “top” of the transformer is negative; Figure 2.15(b) shows the equivalent circuit for the opposite polarity. In the circuit in Figure 2.15(a), the forward diode resistance of D_2 is small; therefore, the capacitor C_1 will charge to almost the peak value of v_s . Terminal 2 on C_1 is positive with respect to terminal 1. As the magnitude of v_s decreases from its peak value, C_1 discharges through R_L and C_2 . We assume that the time constant $R_L C_2$ is very long compared to the period of the input signal.

As the polarity of v_s changes to that shown in Figure 2.15(b), the voltage across C_1 is essentially constant at V_M , with terminal 2 remaining positive. As v_s reaches its maximum value, the voltage across C_2 essentially becomes V_M . By Kirchhoff’s voltage law,

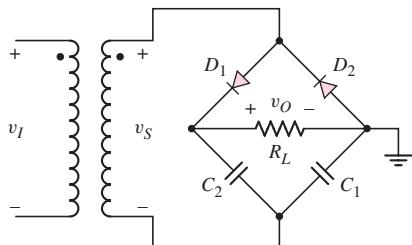


Figure 2.14 A voltage doubler circuit

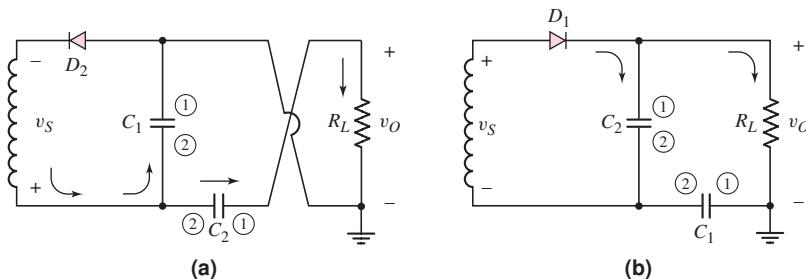


Figure 2.15 Equivalent circuit of the voltage doubler circuit: (a) negative input cycle and (b) positive input cycle

the peak voltage across R_L is now essentially equal to $2V_M$, or twice the peak output of the transformer. The same ripple effect occurs as in the output voltage of the rectifier circuits, but if C_1 and C_2 are relatively large, then the ripple voltage V_r is quite small.

There are also voltage tripler and voltage quadrupler circuits. These circuits provide a means by which multiple dc voltages can be generated from a single ac source and power transformer.

Test Your Understanding

TYU 2.1 Consider the circuit in Figure 2.4. The input voltage is $v_s(t) = 15 \sin \omega t$ (V) and the diode cut-in voltage is $V_y = 0.7$ V. The voltage V_B varies between $4 \leq V_B \leq 8$ V. The peak current is to be limited to $i_D(\text{peak}) = 18$ mA. (a) Determine the minimum value of R . (b) Using the results of part (a), determine the range in peak current and the range in duty cycle. (Ans. (a) $R = 572 \Omega$; (b) $11 \leq i_D(\text{peak}) \leq 18$ mA, $30.3 \leq \text{duty cycle} \leq 39.9\%$).

TYU 2.2 The circuit in Figure 2.5(a) is used to rectify a sinusoidal input signal with a peak voltage of 120 V and a frequency of 60 Hz. A filter capacitor is connected in parallel with R . If the output voltage cannot drop below 100 V, determine the required value of the capacitance C . The transformer has a turns ratio of $N_1 : N_2 = 1 : 1$, where N_2 is the number of turns on each of the secondary windings. Assume the diode cut-in voltage is 0.7 V and the output resistance is $2.5 \text{ k}\Omega$. (Ans. $C = 20.6 \mu\text{F}$)

TYU 2.3 The secondary transformer voltage of the rectifier circuit shown in Figure 2.6(a) is $v_S = 50 \sin[2\pi(60)t]$ V. Each diode has a cut-in voltage of $V_y = 0.7$ V, and the load resistance is $R = 10 \text{ k}\Omega$. Determine the value of the filter capacitor that must be connected in parallel with R such that the ripple voltage is no greater than $V_r = 2$ V. (Ans. $C = 20.3 \mu\text{F}$)

TYU 2.4 Determine the fraction (percent) of the cycle that each diode is conducting in (a) Exercise EX2.4, (b) Exercise TYU2.2, and (c) Exercise TYU2.3. (Ans. (a) 5.2%, (b) 18.1%, (c) 9.14%)



2.2 ZENER DIODE CIRCUITS

Objective: • Apply the characteristics of the Zener diode to a Zener diode voltage regulator circuit.

In Chapter 1, we saw that the breakdown voltage of a Zener diode was nearly constant over a wide range of reverse-bias currents (Figure 1.21). This makes the Zener diode useful in a **voltage regulator**, or a constant-voltage reference circuit. In this chapter, we will look at an ideal voltage reference circuit, and the effects of including a nonideal **Zener resistance**.

The Zener diode voltage regulator circuits designed in this section will then be added to the design of the electronic power supply in Figure 2.1. We should note that in actual power supply designs, the voltage regulator will be a more sophisticated integrated circuit rather than the simpler Zener diode design that will be developed

here. One reason is that a standard Zener diode with a particular desired breakdown voltage may not be available. However, this section will provide the basic concept of a voltage regulator.

2.2.1 Ideal Voltage Reference Circuit

Figure 2.16 shows a Zener voltage regulator circuit. For this circuit, the output voltage should remain constant, even when the output load resistance varies over a fairly wide range, and when the input voltage varies over a specific range. The variation in V_{PS} may be the ripple voltage from a rectifier circuit.

We determine, initially, the proper input resistance R_i . The resistance R_i limits the current through the Zener diode and drops the “excess” voltage between V_{PS} and V_Z . We can write

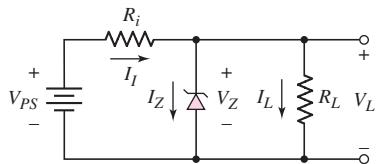


Figure 2.16 A Zener diode voltage regulator circuit

$$R_i = \frac{V_{PS} - V_Z}{I_i} = \frac{V_{PS} - V_Z}{I_Z + I_L} \quad (2.26)$$

which assumes that the Zener resistance is zero for the ideal diode. Solving this equation for the diode current, I_Z , we get

$$I_Z = \frac{V_{PS} - V_Z}{R_i} - I_L \quad (2.27)$$

where $I_L = V_Z/R_L$, and the variables are the input voltage source V_{PS} and the load current I_L .

For proper operation of this circuit, the diode must remain in the breakdown region and the power dissipation in the diode must not exceed its rated value. In other words:

1. The current in the diode is a minimum, $I_Z(\min)$, when the load current is a maximum, $I_L(\max)$, and the source voltage is a minimum, $V_{PS}(\min)$.
2. The current in the diode is a maximum, $I_Z(\max)$, when the load current is a minimum, $I_L(\min)$, and the source voltage is a maximum, $V_{PS}(\max)$.

Inserting these two specifications into Equation (2.26), we obtain

$$R_i = \frac{V_{PS}(\min) - V_Z}{I_Z(\min) + I_L(\max)} \quad (2.28(a))$$

and

$$R_i = \frac{V_{PS}(\max) - V_Z}{I_Z(\max) + I_L(\min)} \quad (2.28(b))$$

Equating these two expressions, we then obtain

$$\begin{aligned} [V_{PS}(\min) - V_Z] \cdot [I_Z(\max) + I_L(\min)] \\ = [V_{PS}(\max) - V_Z] \cdot [I_Z(\min) + I_L(\max)] \end{aligned} \quad (2.29)$$

Reasonably, we can assume that we know the range of input voltage, the range of output load current, and the Zener voltage. Equation (2.29) then contains two unknowns, $I_Z(\min)$ and $I_Z(\max)$. Further, as a minimum requirement, we can set the minimum Zener current to be one-tenth the maximum Zener current, or $I_Z(\min) = 0.1I_Z(\max)$. (More stringent design requirements may require the minimum Zener current to be 20 to 30 percent of the maximum value.) We can then solve for $I_Z(\max)$, using Equation (2.29), as follows:

$$I_Z(\max) = \frac{I_L(\max) \cdot [V_{PS}(\max) - V_Z] - I_L(\min) \cdot [V_{PS}(\min) - V_Z]}{V_{PS}(\min) - 0.9V_Z - 0.1V_{PS}(\max)} \quad (2.30)$$

Using the maximum current thus obtained from Equation (2.30), we can determine the maximum required power rating of the Zener diode. Then, combining Equation (2.30) with either Equation (2.28(a)) or (2.28(b)), we can determine the required value of the input resistance R_i .

DESIGN EXAMPLE 2.5

Objective: Design a voltage regulator using the circuit in Figure 2.16.

The voltage regulator is to power a car radio at $V_L = 9$ V from an automobile battery whose voltage may vary between 11 and 13.6 V. The current in the radio will vary between 0 (off) to 100 mA (full volume).

The equivalent circuit is shown in Figure 2.17.

Solution: The maximum Zener diode current can be determined from Equation (2.30) as

$$I_Z(\max) = \frac{(100)[13.6 - 9] - 0}{11 - (0.9)(9) - (0.1)(13.6)} \cong 300 \text{ mA}$$

The maximum power dissipated in the Zener diode is then

$$P_Z(\max) = I_Z(\max) \cdot V_Z = (300)(9) \Rightarrow 2.7 \text{ W}$$

The value of the current-limiting resistor R_i , from Equation (2.28(b)), is

$$R_i = \frac{13.6 - 9}{0.3 + 0} = 15.3 \Omega$$

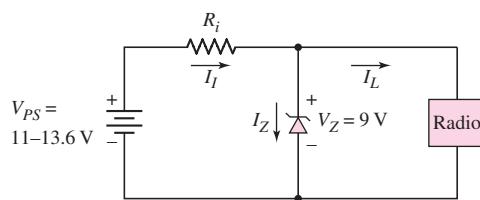


Figure 2.17 Circuit for Design Example 2.5

The maximum power dissipated in this resistor is

$$P_{Ri}(\max) = \frac{(V_{PS}(\max) - V_Z)^2}{R_i} = \frac{(13.6 - 9)^2}{15.3} \cong 1.4 \text{ W}$$

We find

$$I_Z(\min) = \frac{11 - 9}{15.3} - 0.10 \Rightarrow 30.7 \text{ mA}$$

Comment: From this design, we see that the minimum power ratings of the Zener diode and input resistor are 2.7 W and 1.4 W, respectively. The minimum Zener diode current occurs for $V_{PS}(\min)$ and $I_L(\max)$. We find $I_Z(\min) = 30.7$ mA, which is approximately 10 percent of $I_Z(\max)$ as specified by the design equations.

Design Pointer: (1) The variable input in this example was due to a variable battery voltage. However, referring back to Example 2.4, the variable input could also be a function of using a standard transformer with a given turns ratio as opposed to a custom design with a particular turns ratio and/or having a 120 V (rms) input voltage that is not exactly constant.

(2) The 9 V output is a result of using a 9 V Zener diode. However, a Zener diode with exactly a 9 V breakdown voltage may also not be available. We will again see later how more sophisticated designs can solve this problem.

EXERCISE PROBLEM

Ex 2.5: The Zener diode regulator circuit shown in Figure 2.16 has an input voltage that varies between 10 and 14 V, and a load resistance that varies between $R_L = 20$ and 100Ω . Assume a 5.6 Zener diode is used, and assume $I_Z(\min) = 0.1I_Z(\max)$. Find the value of R_i required and the minimum power rating of the diode. (Ans. $P_Z = 3.31 \text{ W}$, $R_i \cong 13 \Omega$)

The operation of the Zener diode circuit shown in Figure 2.17 can be visualized by using load lines. Summing currents at the Zener diode, we have

$$\frac{v_{PS} - V_Z}{R_i} = I_Z + \frac{V_Z}{R_L} \quad (2.31)$$

Solving for V_Z , we obtain

$$V_Z = v_{PS} \left(\frac{R_L}{R_i + R_L} \right) - I_Z \left(\frac{R_i R_L}{R_i + R_L} \right) \quad (2.32)$$

which is the load line equation. Using the parameters of Example 2.5, the load resistance varies from $R_L = \infty$ ($I_L = 0$) to $R_L = 9/0.1 = 90 \Omega$ ($I_L = 100 \text{ mA}$). The current limiting resistor is $R_i = 15 \Omega$ and the input voltage varies over the range $11 \leq v_{PS} \leq 13.6 \text{ V}$.

We may write load line equations for the various circuit conditions.

- A: $v_{PS} = 11 \text{ V}$, $R_L = \infty$; $V_Z = 11 - I_Z(15)$
- B: $v_{PS} = 11 \text{ V}$, $R_L = 90 \Omega$; $V_Z = 9.43 - I_Z(12.9)$
- C: $v_{PS} = 13.6 \text{ V}$, $R_L = \infty$; $V_Z = 13.6 - I_Z(15)$
- D: $v_{PS} = 13.6 \text{ V}$, $R_L = 90 \Omega$; $V_Z = 11.7 - I_Z(12.9)$

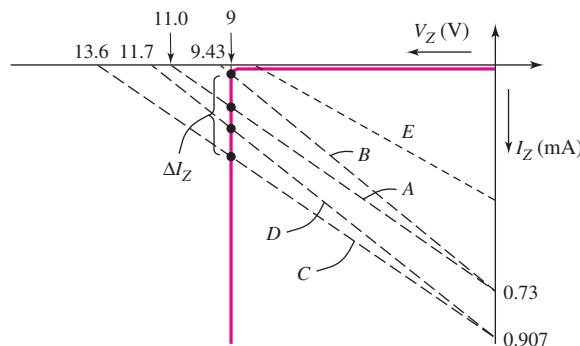


Figure 2.18 Zener diode I - V characteristics with various load lines superimposed

Figure 2.18 shows the Zener diode I - V characteristics. Superimposed on the figure are the four load lines designated as A, B, C, and D. Each load line intersects the diode characteristics in the breakdown region, which is the required condition for proper diode operation. The variation in Zener diode current ΔI_Z for the various combinations of input voltage and load resistance is shown on the figure.

If we were to choose the input resistance to be $R_i = 25 \Omega$ and let $v_{PS} = 11 \text{ V}$ and $R_L = 90 \Omega$, the load line equation (Equation (2.32)) becomes

$$V_Z = 8.61 - I_Z(19.6) \quad (2.33)$$

This load line is plotted as curve E on Figure 2.18. We see that this load line does not intersect the diode characteristics in the breakdown region. For this condition, the output voltage will not equal the breakdown voltage of $V_Z = 9 \text{ V}$; the circuit does not operate “properly.”

2.2.2

Zener Resistance and Percent Regulation

In the ideal Zener diode, the Zener resistance is zero. In actual Zener diodes, however, this is not the case. The result is that the output voltage will fluctuate slightly with a fluctuation in the input voltage, and will fluctuate with changes in the output load resistance.

Figure 2.19 shows the equivalent circuit of the voltage regulator including the Zener resistance. Because of the Zener resistance, the output voltage will change with a change in the Zener diode current.

Two figures of merit can be defined for a voltage regulator. The first is the **source regulation** and is a measure of the change in output voltage with a change in source

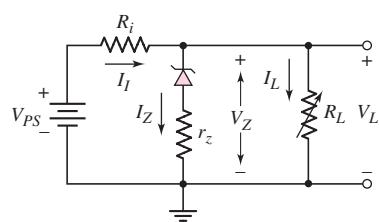


Figure 2.19 A Zener diode voltage regulator circuit with a nonzero Zener resistance

voltage. The second is the **load regulation** and is a measure of the change in output voltage with a change in load current.

The source regulation is defined as

$$\text{Source regulation} = \frac{\Delta v_L}{\Delta v_{PS}} \times 100\% \quad (2.34)$$

where Δv_L is the change in output voltage with a change of Δv_{PS} in the input voltage.

The load regulation is defined as

$$\text{Load regulation} = \frac{v_{L,\text{no load}} - v_{L,\text{full load}}}{v_{L,\text{full load}}} \times 100\% \quad (2.35)$$

where $v_{L,\text{no load}}$ is the output voltage for zero load current and $v_{L,\text{full load}}$ is the output voltage for the maximum rated output current.

The circuit approaches that of an ideal voltage regulator as the source and load regulation factors approach zero.

EXAMPLE 2.6

Objective: Determine the source regulation and load regulation of a voltage regulator circuit.

Consider the circuit described in Example 2.5 and assume a Zener resistance of $r_z = 2 \Omega$.

Solution: Consider the effect of a change in input voltage for a no-load condition ($R_L = \infty$). For $v_{PS} = 13.6 \text{ V}$, we find

$$I_Z = \frac{13.6 - 9}{15.3 + 2} = 0.2659 \text{ A}$$

Then

$$v_{L,\text{max}} = 9 + (2)(0.2659) = 9.532 \text{ V}$$

For $v_{PS} = 11 \text{ V}$, we find

$$I_Z = \frac{11 - 9}{15.3 + 2} = 0.1156 \text{ A}$$

Then

$$v_{L,\text{min}} = 9 + (2)(0.1156) = 9.231 \text{ V}$$

We obtain

$$\text{Source regulation} = \frac{\Delta v_L}{\Delta v_{PS}} \times 100\% = \frac{9.532 - 9.231}{13.6 - 11} \times 100\% = 11.6\%$$

Now consider the effect of a change in load current for $v_{PS} = 13.6 \text{ V}$. For $I_L = 0$, we find

$$I_Z = \frac{13.6 - 9}{15.3 + 2} = 0.2659 \text{ A}$$

and

$$v_{L,\text{no load}} = 9 + (2)(0.2659) = 9.532 \text{ V}$$

For a load current of $I_L = 100 \text{ mA}$, we find

$$I_Z = \frac{13.6 - [9 + I_Z(2)]}{15.3} - 0.10$$

which yields

$$I_Z = 0.1775 \text{ A}$$

Then

$$v_{L,\text{full load}} = 9 + (2)(0.1775) = 9.355 \text{ V}$$

We now obtain

$$\begin{aligned}\text{Load regulation} &= \frac{v_{L,\text{no load}} - v_{L,\text{full load}}}{v_{L,\text{full load}}} \times 100\% \\ &= \frac{9.532 - 9.355}{9.355} \times 100\% = 1.89\%\end{aligned}$$

Comment: The ripple voltage on the input of 2.6 V is reduced by approximately a factor of 10. The change in output load results in a small percentage change in the output voltage.

EXERCISE PROBLEM

Ex 2.6: Repeat Example 2.6 for $r_z = 4 \Omega$. Assume all other parameters are the same as listed in the example. (Ans. Source regulation = 20.7%, load regulation = 3.29%)

Test Your Understanding

TYU 2.5 Consider the circuit shown in Figure 2.19. Let $V_{PS} = 12 \text{ V}$, $V_{ZO} = 6.2 \text{ V}$, and $r_z = 3 \Omega$. The power rating of the diode is $P = 1 \text{ W}$. (a) Determine $I_Z(\text{max})$ and R_i . (b) If $I_Z(\text{min}) = 0.1I_Z(\text{max})$, determine $R_L(\text{min})$ and the load regulation. (Ans. (a) $I_Z(\text{max}) = 150 \text{ mA}$, $R_i = 35.7 \Omega$; (b) $R_L(\text{min}) = 42.7 \Omega$, 6.09 %).

TYU 2.6 Suppose the current-limiting resistor in Example 2.5 is replaced by one whose value is $R_i = 20 \Omega$. Determine the minimum and maximum Zener diode current. Does the circuit operate “properly”? (Ans. $I_Z(\text{min}) = 0$, $I_Z(\text{max}) = 230 \text{ mA}$).

TYU 2.7 Suppose the power supply voltage in the circuit shown in Figure 2.17 drops to $V_{PS} = 10 \text{ V}$. Let $R_i = 15.3 \Omega$. What is the maximum load current in the radio if the minimum Zener diode current is to be maintained at $I_Z(\text{min}) = 30 \text{ mA}$? (Ans. $I_L(\text{max}) = 35.4 \text{ mA}$).



2.3

CLIPPER AND CLAMPER CIRCUITS

Objective: • Apply the nonlinear characteristics of diodes to create waveshaping circuits known as clippers and clampers.

In this section, we continue our discussion of nonlinear circuit applications of diodes. Diodes can be used in waveshaping circuits that either limit or “clip” portions of a signal, or shift the dc voltage level. The circuits are called **clippers** and **clampers**, respectively.

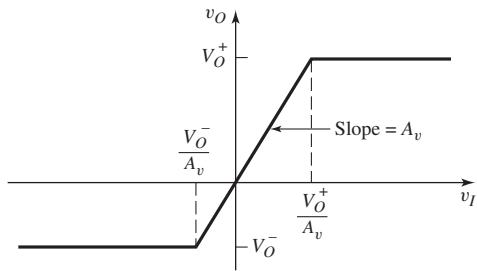


Figure 2.20 General voltage transfer characteristics of a limiter circuit

2.3.1 Clippers

Clipper circuits, also called **limiter circuits**, are used to eliminate portions of a signal that are above or below a specified level. For example, the half-wave rectifier is a clipper circuit, since all voltages below zero are eliminated. A simple application of a clipper is to limit the voltage at the input to an electronic circuit so as to prevent breakdown of the transistors in the circuit. The circuit may be used to measure the frequency of the signal, if the amplitude is not an important part of the signal.

Figure 2.20 shows the general voltage transfer characteristics of a limiter circuit. The limiter is a linear circuit if the input signal is in the range $V_O^-/A_v \leq v_I \leq V_O^+/A_v$, where A_v is the slope of the transfer curve. If $A_v \leq 1$, as in diode circuits, the circuit is a **passive limiter**. If $v_I > V_O^+/A_v$, the output is limited to a maximum value of V_O^+ . Similarly, if $v_I < V_O^-/A_v$, the output is limited to a minimum value of V_O^- . Figure 2.20 shows the general transfer curve of a double limiter, in which both the positive and negative peak values of the input signal are clipped.

Various combinations of V_O^+ and V_O^- are possible. Both parameters may be positive, both negative, or one may be positive while the other negative, as indicated in the figure. If either V_O^- approaches minus infinity or V_O^+ approaches plus infinity, then the circuit reverts to a single limiter.

Figure 2.21(a) is a single-diode clipper circuit. The diode D_1 is off as long as $v_I < V_B + V_\gamma$. With D_1 off, the current is approximately zero, the voltage drop across R is essentially zero, and the output voltage follows the input voltage. When $v_I > V_B + V_\gamma$, the diode turns on, the output voltage is clipped, and v_O equals $V_B + V_\gamma$. The output signal is shown in Figure 2.21(b). In this circuit, the output is clipped above $V_B + V_\gamma$.

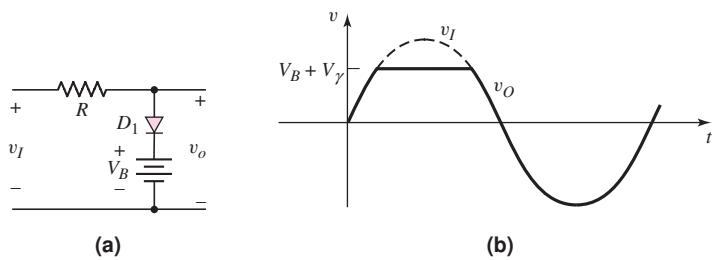


Figure 2.21 Single-diode clipper: (a) circuit and (b) output response

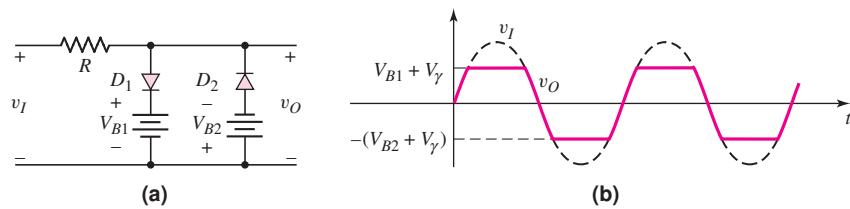


Figure 2.22 A parallel-based diode clipper circuit and its output response

The resistor R in Figure 2.21 is selected to be large enough so that the forward diode current is limited to be within reasonable values (usually in the milliampere range), but small enough so that the reverse diode current produces a negligible voltage drop. Normally, a wide range of resistor values will result in satisfactory performance of a given circuit.

Other clipping circuits can be constructed by reversing the diode, the polarity of the voltage source, or both.

Positive and negative clipping can be performed simultaneously by using a double limiter or a **parallel-based clipper**, such as the circuit shown in Figure 2.22. The input and output signals are also shown in the figure. The parallel-based clipper is designed with two diodes and two voltage sources oriented in opposite directions.

EXAMPLE 2.7

Objective: Find the output of the parallel-based clipper in Figure 2.23(a).

For simplicity, assume that $V_\gamma = 0$ and $r_f = 0$ for both diodes.

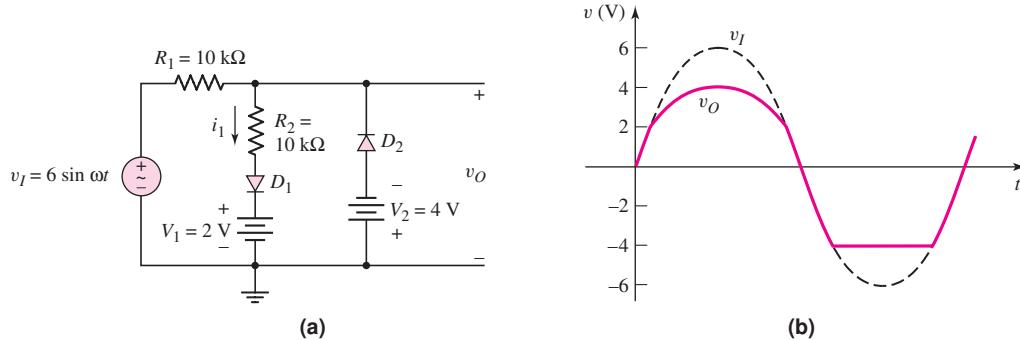


Figure 2.23 Figure for Example 2.7

Solution: For $t = 0$, we see that $v_I = 0$ and both D_1 and D_2 are reverse biased. For $0 < v_I \leq 2$ V, D_1 and D_2 remain off; therefore, $v_O = v_I$. For $v_I > 2$ V, D_1 turns on and

$$i_1 = \frac{v_I - 2}{10 + 10}$$

Also,

$$v_O = i_1 R_2 + 2 = \frac{1}{2}(v_I - 2) + 2 = \frac{1}{2}v_I + 1$$

If $v_I = 6 \text{ V}$, then $v_O = 4 \text{ V}$.

For $-4 < v_I < 0 \text{ V}$, both D_1 and D_2 are off and $v_O = v_I$. For $v_I \leq -4 \text{ V}$, D_2 turns on and the output is constant at $v_O = -4 \text{ V}$. The input and output waveforms are plotted in Figure 2.23(b).

Comment: If we assume that $V_\gamma \neq 0$, the output will be very similar to the results calculated here. The only difference will be the points at which the diodes turn on.

EXERCISE PROBLEM

Ex 2.7: Design a parallel-based clipper that will yield the voltage transfer function shown in Figure 2.24. Assume diode cut-in voltages of $V_\gamma = 0.7 \text{ V}$. (Ans. For Figure 2.23(b), $V_2 = 4.3$, $V_1 = 1.8 \text{ V}$, and $R_1 = 2R_2$)

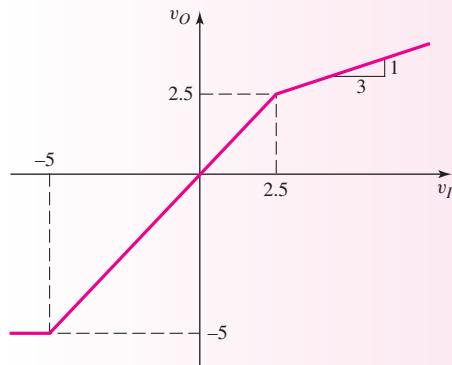


Figure 2.24 Figure for Exercise Ex 2.7

Diode clipper circuits can also be designed such that the dc power supply is in series with the input signal. Figure 2.25 shows one example. The battery in series with the input signal causes the input signal to be superimposed on the V_B dc voltage. The resulting conditioned input signal and corresponding output signal is also shown in Figure 2.25.

In all of the clipper circuits considered, we have included batteries that basically set the limits of the output voltage. However, batteries need periodic replacement, so that these circuits are not practical. Zener diodes, operated in the reverse breakdown region, provide essentially a constant voltage drop. We can replace the batteries by Zener diodes.

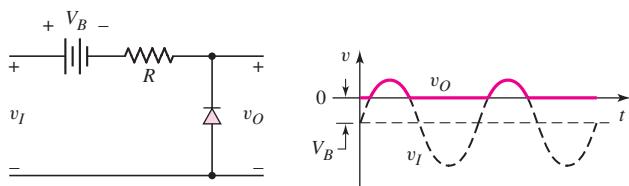


Figure 2.25 Series-based diode clipper circuit and resulting output response

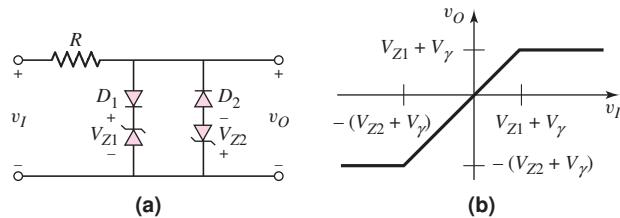


Figure 2.26 (a) Parallel-based clipper circuit using Zener diodes; (b) voltage transfer characteristics

Figure 2.26(a) shows a parallel based clipper circuit using Zener diodes. The voltage transfer characteristics are shown in Figure 2.26(b). The performance of the circuit in Figure 2.26(a) is essentially the same as that shown in Figure 2.22.

2.3.2 Clampers

Clamping shifts the entire signal voltage by a dc level. In steady state, the output waveform is an exact replica of the input waveform, but the output signal is shifted by a dc value that depends on the circuit. The distinguishing feature of a clamper is that it adjusts the dc level without needing to know the exact waveform.

An example of clamping is shown in Figure 2.27(a). The sinusoidal input voltage signal is shown in Figure 2.27(b). Assume that the capacitor is initially uncharged. During the first 90 degrees of the input waveform, the voltage across the capacitor follows the input, and $v_C = v_I$ (assuming that $r_f = 0$ and $V_\gamma = 0$). After v_I and v_C reach their peak values, v_I begins to decrease and the diode becomes reverse biased. Ideally, the capacitor cannot discharge, so the voltage across the capacitor remains constant at $v_C = V_M$. By Kirchhoff's voltage law

$$v_O = -v_C + v_I = -V_M + V_M \sin \omega t \quad (2.36(a))$$

or

$$v_O = V_M(\sin \omega t - 1) \quad (2.36(b))$$

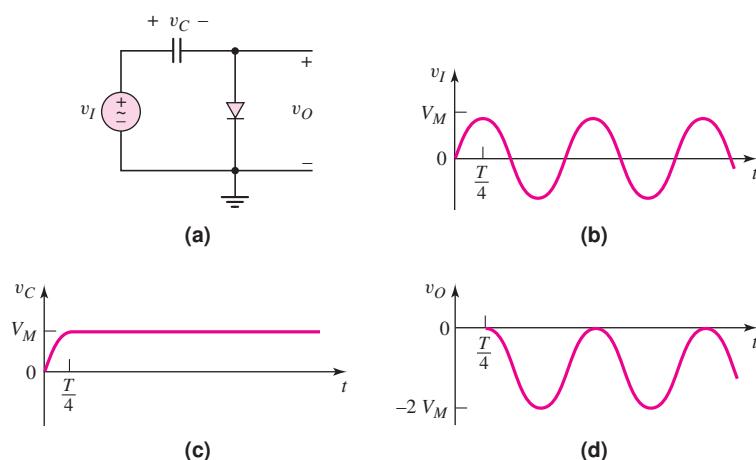


Figure 2.27 Action of a diode clamper circuit: (a) a typical diode clamper circuit, (b) the sinusoidal input signal, (c) the capacitor voltage, and (d) the output voltage

The capacitor and output voltages are shown in Figures 2.27(c) and (d). The output voltage is “clamped” at zero volts, that is, $v_O \leq 0$. In steady state, the waveshapes of the input and output signals are the same, and the output signal is shifted by a certain dc level compared to the input signal.

A clamping circuit that includes an independent voltage source V_B is shown in Figure 2.28(a). In this circuit, the $R_L C$ time constant is assumed to be large, where R_L is the load resistance connected to the output. If we assume, for simplicity, that $r_f = 0$ and $V_\gamma = 0$, then the output is clamped at V_B . Figure 2.28(b) shows an example of a sinusoidal input signal and the resulting output voltage signal. When the polarity of V_B is as shown, the output is shifted in a negative voltage direction. Similarly, Figure 2.28(c) shows a square-wave input signal and the resulting output voltage signal. For the square-wave signal, we have neglected the diode capacitance effects and assume the voltage can change instantaneously.

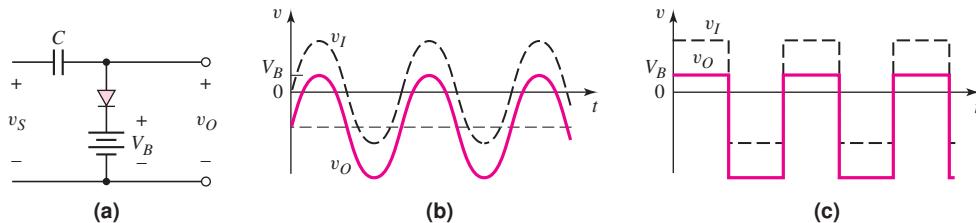


Figure 2.28 Action of a diode clamper circuit with a voltage source assuming an ideal diode ($V_r = 0$): (a) the circuit, (b) steady-state sinusoidal input and output signals, and (c) steady-state square-wave input and output signals

Electronic signals tend to lose their dc levels during signal transmission. For example, the dc level of a TV signal may be lost during transmission, so that the dc level must be restored at the TV receiver. The following example illustrates this effect.

EXAMPLE 2.8

Objective: Find the steady-state output of the diode-clamper circuit shown in Figure 2.29(a).

The input v_I is assumed to be a sinusoidal signal whose dc level has been shifted with respect to a receiver ground by a value V_B during transmission. Assume $V_\gamma = 0$ and $r_f = 0$ for the diode.

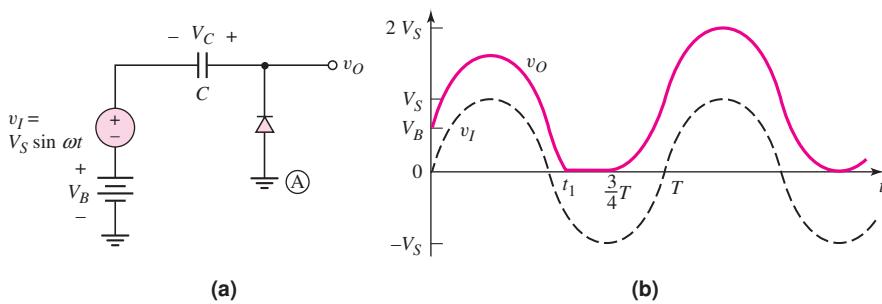


Figure 2.29 (a) Circuit for Example 2.8; (b) input and output waveforms

Solution: Figure 2.29(b) shows the sinusoidal input signal. If the capacitor is initially uncharged, then the output voltage is $v_O = V_B$ at $t = 0$ (diode reverse-biased). For $0 \leq t \leq t_1$, the effective RC time constant is infinite, the voltage across the capacitor does not change, and $v_O = v_I + V_B$.

At $t = t_1$, the diode becomes forward biased; the output cannot go negative, so the voltage across the capacitor changes (the $r_f C$ time constant is zero).

At $t = (\frac{3}{4})T$, the input signal begins increasing and the diode becomes reverse biased, so the voltage across the capacitor now remains constant at $V_C = V_S - V_B$ with the polarity shown. The output voltage is now given by

$$v_O = (V_S - V_B) + v_I + V_B = (V_S - V_B) + V_S \sin \omega t + V_B$$

or

$$v_O = V_S(1 + \sin \omega t)$$

Comment: For $t > (\frac{3}{4})T$, steady state is reached. The output signal waveform is an exact replica of the input signal waveform and is now measured with respect to the reference ground at terminal A.

EXERCISE PROBLEM

Ex 2.8: Sketch the steady-state output voltage for the input signal given for the circuit shown in Figure 2.30. Assume $V_\gamma = r_f = 0$. (Ans. Square wave between $+2$ V and -8 V)

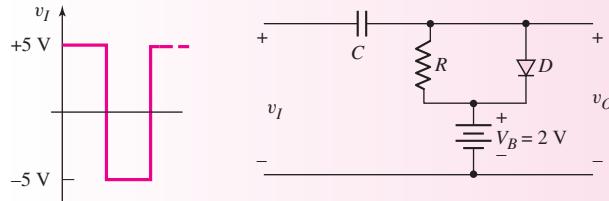


Figure 2.30 Figure for Exercise Ex 2.8

Test Your Understanding

TYU 2.8 Consider the circuit in Figure 2.23(a). Let $R_1 = 5\text{ k}\Omega$, $R_2 = 2\text{ k}\Omega$, $V_1 = 1$ V, and $V_2 = 3$ V. Let $V_\gamma = 0.7$ V for each diode. Plot the voltage transfer characteristics (v_O versus v_I) for $-5 \leq v_I \leq 5$ V. (Ans. For $v_I \leq -3.7$ V, $v_O = -3.7$ V; for $-3.7 \leq v_I \leq 1.7$ V, $v_O = v_I$; for $v_I \geq 1.7$ V, $v_O = 0.286 v_I + 1.21$)

TYU 2.9 Determine the steady-state output voltage v_O for the circuit in Figure 2.31(a), if the input is as shown in Figure 2.31(b). Assume the diode cut-in voltage is $V_\gamma = 0$. (Ans. Output is a square wave between $+5$ V and $+35$ V)

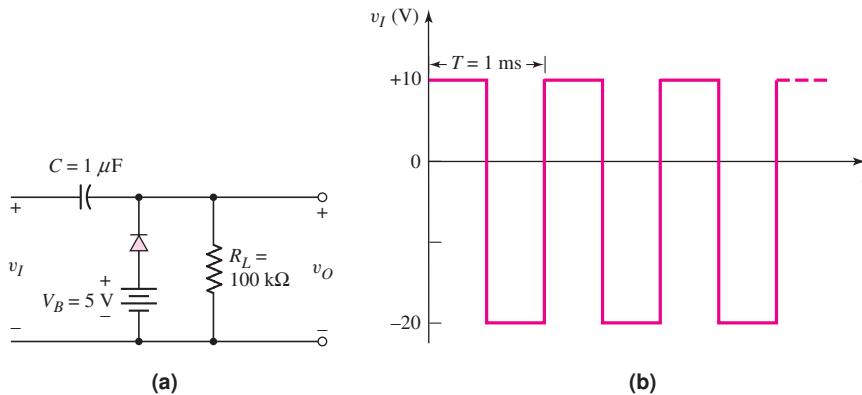


Figure 2.31 Figure for Exercise TYU 2.9: (a) the circuit and (b) input signal

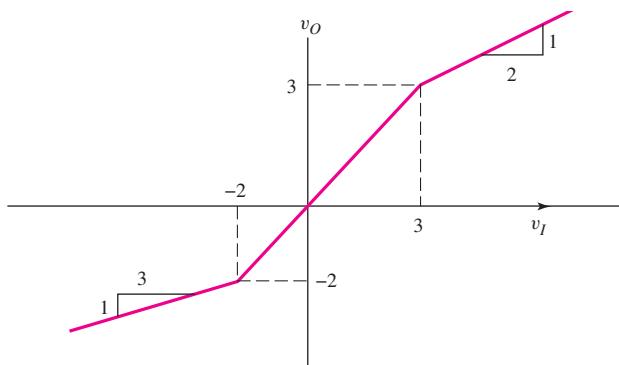


Figure 2.32 Figure for Exercise TYU 2.10

TYU 2.10 Design a parallel-based clipper circuit that will yield the voltage transfer characteristics shown in Figure 2.32. Assume a diode cut-in voltage of $V_V = 0.7 \text{ V}$. (Ans. From Figure 2.23(a), $V_1 = 2.3 \text{ V}$, $V_2 = 1.3 \text{ V}$, $R_1 = R_2$, include R_3 in series with D_2 , where $R_3 = 0.5R_1$)



2.4

MULTIPLE-DIODE CIRCUITS

Objective: • Examine the techniques used to analyze circuits that contain more than one diode.

Since a diode is a nonlinear device, part of the analysis of a diode circuit involves determining whether the diode is on or off. If a circuit contains more than one diode, the analysis is complicated by the various possible combinations of on and off.

In this section, we will look at several multiple-diode circuits. We will see, for example, how diode circuits can be used to perform logic functions. This section serves as an introduction to digital logic circuits that will be considered in detail in Chapters 16 and 17.

2.4.1 Example Diode Circuits

As a brief introduction, consider two single-diode circuits. Figure 2.33(a) shows a diode in series with a resistor. A plot of voltage transfer characteristics, v_O versus v_I , shows the piecewise linear nature of this circuit (Figure 2.33(b)). The diode does not begin to conduct until $v_I = V_\gamma$. Consequently, for $v_I \leq V_\gamma$, the output voltage is zero; for $v_I > V_\gamma$, the output voltage is $v_O = v_I - V_\gamma$.

Figure 2.34(a) shows a similar diode circuit, but with the input voltage source explicitly included to show that there is a path for the diode current. The voltage transfer characteristic is shown in Figure 2.34(b). In this circuit, the diode remains conducting for $v_I < V_S - V_\gamma$, and the output voltage is $v_O = v_I + V_\gamma$. When $v_I > V_S - V_\gamma$, the diode turns off and the current through the resistor is zero; therefore, the output remains constant at V_S .

These two examples demonstrate the piecewise linear nature of the diode and the diode circuit. They also demonstrate that there are regions where the diode is “on,” or conducting, and regions where the diode is “off,” or nonconducting.

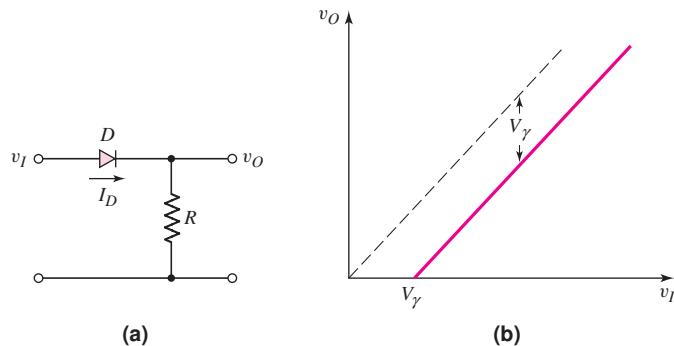


Figure 2.33 Diode and resistor in series: (a) circuit and (b) voltage transfer characteristics

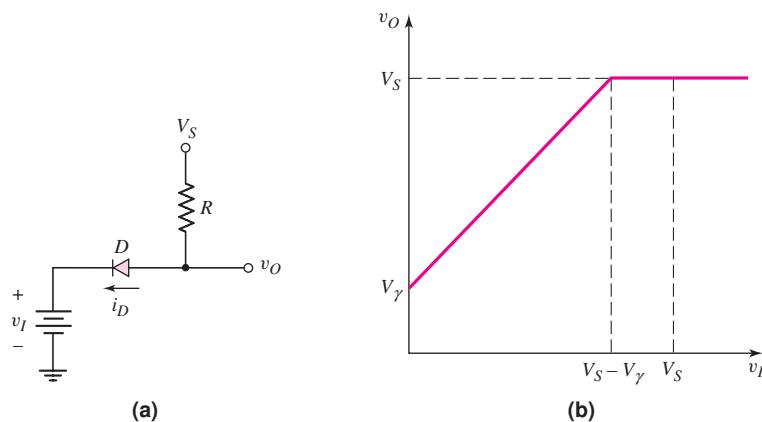


Figure 2.34 Diode with input voltage source: (a) circuit and (b) voltage transfer characteristics

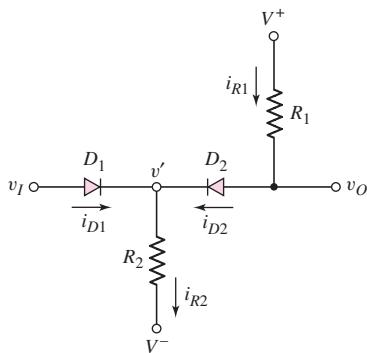


Figure 2.35 A two-diode circuit

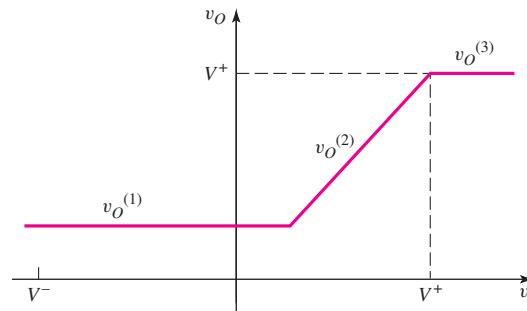


Figure 2.36 Voltage transfer characteristics for the two-diode circuit in Figure 2.35

In multidiode circuits, each diode may be either on or off. Consider the two-diode circuit in Figure 2.35. Since each diode may be either on or off, the circuit has four possible states. However, some of these states may not be feasible because of diode directions and voltage polarities.

If we assume that $V^+ > V^-$ and that $V^+ - V^- > V_\gamma$, there is at least a possibility that D_2 can be turned on. First, v' cannot be less than V^- . Then, for $v_I = V^-$, diode D_1 must be off. In this case, D_2 is on, $i_{R1} = i_{D2} = i_{R2}$, and

$$v_O = V^+ - i_{R1} R_1 \quad (2.37)$$

where

$$i_{R1} = \frac{V^+ - V_\gamma - V^-}{R_1 + R_2} \quad (2.38)$$

Voltage v' is one diode drop below v_O , and D_1 remains off as long as v_I is less than the output voltage. As v_I increases and becomes equal to v_O , both D_1 and D_2 turn on. This condition or state is valid as long as $v_I < V^+$. When $v_I = V^+$, $i_{R1} = i_{D2} = 0$, at which point D_2 turns off and v_O cannot increase any further.

Figure 2.36 shows the resulting plot of v_O versus v_I . Three distinct regions, $v_O^{(1)}$, $v_O^{(2)}$, and $v_O^{(3)}$, correspond to the various conducting states of D_1 and D_2 . The fourth possible state, corresponding to both D_1 and D_2 being off, is not feasible in this circuit.

EXAMPLE 2.9

Objective: Determine the output voltage and diode currents for the circuit shown in Figure 2.35, for two values of input voltage.

Assume the circuit parameters are $R_1 = 5 \text{ k}\Omega$, $R_2 = 10 \text{ k}\Omega$, $V_\gamma = 0.7 \text{ V}$, $V^+ = +5 \text{ V}$, and $V^- = -5 \text{ V}$. Determine v_O , i_{D1} , and i_{D2} for $v_I = 0$ and $v_I = 4 \text{ V}$.

Solution: For $v_I = 0$, assume initially that D_1 is off. The currents are then

$$i_{R1} = i_{D2} = i_{R2} = \frac{V^+ - V_\gamma - V^-}{R_1 + R_2} = \frac{5 - 0.7 - (-5)}{5 + 10} = 0.62 \text{ mA}$$

The output voltage is

$$v_O = V^+ - i_{R1} R_1 = 5 - (0.62)(5) = 1.9 \text{ V}$$

and v' is

$$v' = v_O - V_\gamma = 1.9 - 0.7 = 1.2 \text{ V}$$

From these results, we see that diode D_1 is indeed cut off, $i_{D1} = 0$, and our analysis is valid.

For $v_I = 4$ V, we see from Figure 2.36 that $v_O = v_I$; therefore, $v_O = v_I = 4$ V. In this region, both D_1 and D_2 are on, and

$$i_{R1} = i_{D2} = \frac{V^+ - v_O}{R_1} = \frac{5 - 4}{5} = 0.2 \text{ mA}$$

Note that $v' = v_O - V_\gamma = 4 - 0.7 = 3.3$ V. Thus,

$$i_{R2} = \frac{v' - V^-}{R_2} = \frac{3.3 - (-5)}{10} = 0.83 \text{ mA}$$

The current through D_1 is found from $i_{D1} + i_{D2} = i_{R2}$ or

$$i_{D1} = i_{R2} - i_{D2} = 0.83 - 0.2 = 0.63 \text{ mA}$$

Comment: For $v_I = 0$, we see that $v_O = 1.9$ V and $v' = 1.2$ V. This means that D_1 is reverse biased, or off, as we initially assumed. For $v_I = 4$ V, we have $i_{D1} > 0$ and $i_{D2} > 0$, indicating that both D_1 and D_2 are forward biased, as we assumed.

Computer analysis: For multidiode circuits, a PSpice analysis may be useful in determining the conditions under which the various diodes are conducting or not conducting. This avoids guessing the conducting state of each diode in a hand analysis. Figure 2.37 is the PSpice circuit schematic of the diode circuit in Figure 2.35.

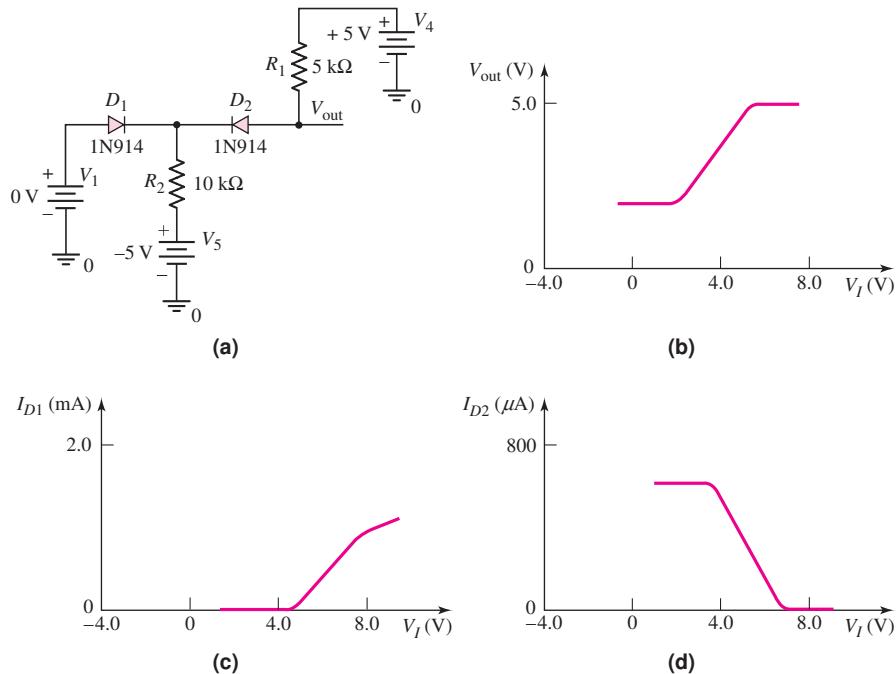


Figure 2.37 (a) PSpice circuit schematic; (b) output voltage; (c) current in diode 1, and (d) current in diode 2 for the diode circuit in Example 2.9

Figure 2.37 also shows the output voltage and the two diode currents as the input is varied between -1 V and $+7$ V. From these curves, we can determine when the diodes turn on and off.

Comment: The hand analysis results, based on the piecewise linear model for the diode, agree very well with the computer simulation results. This gives us confidence in the piecewise linear model when quick hand calculations are made.

EXERCISE PROBLEM

Ex 2.9: Consider the circuit shown in Figure 2.38, in which the diode cut-in voltages are $V_y = 0.6$ V. Plot v_o versus v_I for $0 \leq v_I \leq 10$ V. (Ans. For $0 \leq v_I \leq 3.5$ V, $v_o = 4.4$ V; for $v_I > 3.5$ V, D_2 turns off; and for $v_I \geq 9.4$ V, $v_o = 10$ V)

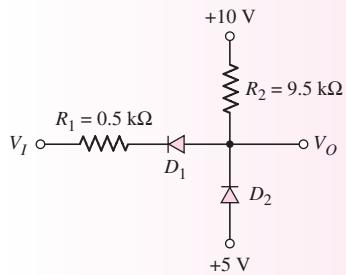


Figure 2.38 Figure for Exercise Ex 2.9

Problem-Solving Technique: Multiple Diode Circuits

Analyzing multidiode circuits requires determining if the individual devices are “on” or “off.” In many cases, the choice is not obvious, so we must initially guess the state of each device, then analyze the circuit to determine if we have a solution consistent with our initial guess. To do this, we can:

1. Assume the state of a diode. If a diode is assumed on, the voltage across the diode is assumed to be V_y . If a diode is assumed to be off, the current through the diode is assumed to be zero.
2. Analyze the “linear” circuit with the assumed diode states.
3. Evaluate the resulting state of each diode. If the initial assumption were that a diode is “off” and the analysis shows that $I_D = 0$ and $V_D \leq V_y$, then the assumption is correct. If, however, the analysis actually shows that $I_D > 0$ and/or $V_D > V_y$, then the initial assumption is incorrect. Similarly, if the initial assumption were that a diode is “on” and the analysis shows that $I_D \geq 0$ and $V_D = V_y$, then the initial assumption is correct. If, however, the analysis shows that $I_D < 0$ and/or $V_D < V_y$, then the initial assumption is incorrect.
4. If any initial assumption is proven incorrect, then a new assumption must be made and the new “linear” circuit must be analyzed. Step 3 must then be repeated.

EXAMPLE 2.10

Objective: Demonstrate how inconsistencies develop in a solution with incorrect assumptions.

For the circuit shown in Figure 2.35, assume that parameters are the same as those given in Example 2.9. Determine v_O , i_{D1} , i_{D2} , and i_{R2} for $v_I = 0$.

Solution: Assume initially that both D_1 and D_2 are conducting (i.e., on). Then, $v' = -0.7$ V and $v_O = 0$. The two currents are

$$i_{R1} = i_{D2} = \frac{V^+ - v_O}{R_1} = \frac{5 - 0}{5} = 1.0 \text{ mA}$$

and

$$i_{R2} = \frac{v' - V^-}{R_2} = \frac{-0.7 - (-5)}{10} = 0.43 \text{ mA}$$

Summing the currents at the v' node, we find that

$$i_{D1} = i_{R2} - i_{D2} = 0.43 - 1.0 = -0.57 \text{ mA}$$

Since this analysis shows the D_1 current to be negative, which is an impossible or inconsistent solution, our initial assumption must be incorrect. If we go back to Example 2.9, we will see that the correct solution is D_1 off and D_2 on when $v_I = 0$.

Comment: We can perform linear analyses on diode circuits, using the piecewise linear model. However, we must first determine if each diode in the circuit is operating in the “on” linear region or the “off” linear region.

EXERCISE PROBLEM

Ex 2.10: Consider the circuit shown in Figure 2.39. The cut-in voltage of each diode is $V_\gamma = 0.7$ V. (a) Let $v_I = 5$ V. Assume both diodes are conducting. Is this a correct assumption? Why or why not? Determine I_{R1} , I_{D1} , I_{D2} , and v_O . (b) Repeat part (a) for $v_I = 10$ V. (Ans. (a) D_1 is off, $I_{D1} = 0$, $I_{R1} = I_{D2} = 0.754$ mA, $v_O = 3.72$ V; (b) $I_{D1} = 0.9$ mA, $I_{D2} = 1.9$ mA, $I_{R1} = 1.0$ mA, $v_O = 8.3$ V)

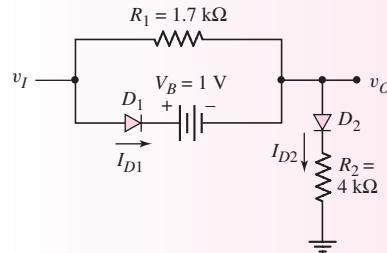


Figure 2.39 Figure for Exercise Ex 2.10

EXAMPLE 2.11

Objective: Determine the current in each diode and the voltages V_A and V_B in the multidiode circuit shown in Figure 2.40. Let $V_\gamma = 0.7$ V for each diode.

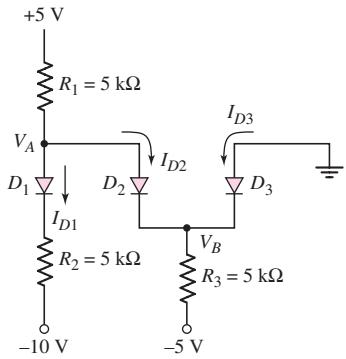


Figure 2.40 Diode circuit for Example 2.11

Solution: Initially assume each diode is in its conducting state. Starting with D_3 and considering the voltages, we see that

$$V_B = -0.7 \text{ V} \quad \text{and} \quad V_A = 0$$

Summing currents at the V_A node, we find

$$\frac{5 - V_A}{5} = I_{D2} + \frac{(V_A - 0.7) - (-10)}{5}$$

Since $V_A = 0$, we obtain

$$\frac{5}{5} = I_{D2} + \frac{9.3}{5} \Rightarrow I_{D2} = -0.86 \text{ mA}$$

which is inconsistent with the assumption that all diodes are “on” (an “on” diode would have a positive diode current).

Now assume that D_1 and D_3 are on and D_2 is off. We see that

$$I_{D1} = \frac{5 - 0.7 - (-10)}{5 + 5} = 1.43 \text{ mA}$$

and

$$I_{D3} = \frac{(0 - 0.7) - (-5)}{5} = 0.86 \text{ mA}$$

We find the voltages as

$$V_B = -0.7 \text{ V}$$

and

$$V_A = 5 - (1.43)(5) = -2.15 \text{ V}$$

From the values of V_A and V_B , the diode D_2 is indeed reverse biased and off, so $I_{D2} = 0$.

Comment: With more diodes in a circuit, the number of combinations of diodes being either on or off increases, which may increase the number of times a circuit must be analyzed before a correct solution is obtained. In the case of multiple diode circuits, a computer simulation might save time.

EXERCISE PROBLEM

Ex 2.11: Repeat Example 2.11 for the case when $R_1 = 8 \text{ k}\Omega$, $R_2 = 4 \text{ k}\Omega$, and $R_3 = 2\text{k}\Omega$. (Ans. $V_B = -0.7 \text{ V}$, $I_{D3} = 2.15 \text{ mA}$, $I_{D2} = 0$, $I_{D1} = 1.19 \text{ mA}$, $V_A = -4.53 \text{ V}$)

2.4.2 Diode Logic Circuits

Diodes in conjunction with other circuit elements can perform certain **logic functions**, such as AND and OR. The circuit in Figure 2.41 is an example of a diode logic circuit. The four conditions of operation of this circuit depend on various combinations of input voltages. If $V_1 = V_2 = 0$, there is no excitation to the circuit so both diodes are off and $V_O = 0$. If at least one input goes to 5 V, for example, at least one diode turns on and $V_O = 4.3 \text{ V}$, assuming $V_\gamma = 0.7 \text{ V}$.

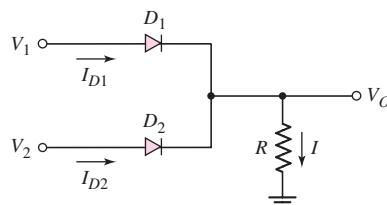


Figure 2.41 A two-input diode OR logic circuit

These results are shown in Table 2.1. By definition, in a positive logic system, a voltage near zero corresponds to a logic 0 and a voltage close to the supply voltage of 5 V corresponds to a logic 1. The results shown in Table 2.1 indicate that this circuit performs the OR logic function. The circuit of Figure 2.41, then, is a two-input diode OR logic circuit.

Next, consider the circuit in Figure 2.42. Assume a diode cut-in voltage of $V_\gamma = 0.7 \text{ V}$. Again, there are four possible states, depending on the combination of input voltages. If at least one input is at zero volts, then at least one diode is conducting and $V_O = 0.7 \text{ V}$. If both $V_1 = V_2 = 5 \text{ V}$, there is no potential difference between the supply voltage and the input voltage. All currents are zero and $V_O = 5 \text{ V}$.

These results are shown in Table 2.2. This circuit performs the AND logic function. The circuit of Figure 2.42 is a two-input diode AND logic circuit.

Table 2.1 Two-diode OR logic circuit response

$V_1 \text{ (V)}$	$V_2 \text{ (V)}$	$V_o \text{ (V)}$
0	0	0
5	0	4.3
0	5	4.3
5	5	4.3

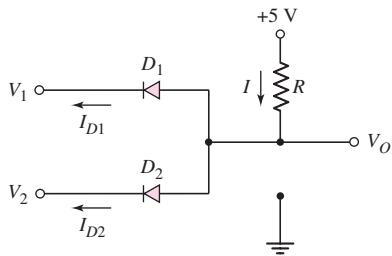


Figure 2.42 A two-input diode AND logic circuit

If we examine Tables 2.1 and 2.2, we see that the input “low” and “high” voltages may not be the same as the output “low” and “high” voltages. As an example, for the AND circuit (Table 2.2), the input “low” is 0 V, but the output “low” is 0.7 V. This can create a problem because the output of one logic gate is often the input to another logic gate. Another problem occurs when diode logic circuits are connected in cascade; that is, the output of one OR gate is connected to the input of a second OR gate. The logic 1 levels of the two OR gates are not the same (see Problems 2.61 and 2.62). The logic 1 level degrades or decreases as additional logic gates are connected. However, these problems may be overcome with the use of amplifying devices (transistors) in digital logic systems.

Table 2.2 Two-diode AND logic circuit response

V_1 (V)	V_2 (V)	V_o (V)
0	0	0.7
5	0	0.7
0	5	0.7
5	5	5

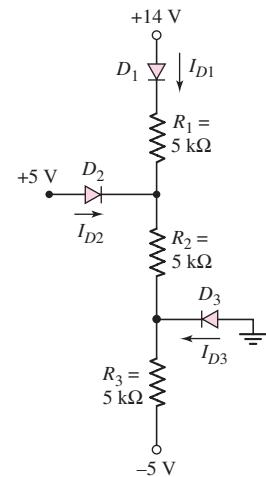


Figure 2.43 Figure for Exercise TYU 2.11

Test Your Understanding

TYU 2.11 The cut-in voltage of each diode in the circuit shown in Figure 2.43 is $V_y = 0.7$ V. Determine I_{D1} , I_{D2} , I_{D3} , V_A , and V_B . (Ans. $I_{D1} = 1.22$ mA, $I_{D2} = I_{D3} = 0$, $V_A = 7.2$ V, $V_B = 1.1$ V)

TYU 2.12 Repeat Exercise TYU 2.11 for $R_1 = 8\text{ k}\Omega$, $R_2 = 12\text{ k}\Omega$, and $R_3 = 2.5\text{ k}\Omega$. (Ans. $I_{D1} = 0.7$ mA, $I_{D2} = 0$, $I_{D3} = 1.02$ mA, $V_A = 7.7$ V, $V_B = -0.7$ V)

TYU 2.13 Consider the OR logic circuit shown in Figure 2.41. Assume a diode cut-in voltage of $V_y = 0.6$ V. (a) Plot V_O versus V_1 for $0 \leq V_1 \leq 5$ V, if $V_2 = 0$. (b) Repeat part (a) if $V_2 = 3$ V. (Ans. (a) $V_O = 0$ for $V_1 \leq 0.6$ V, $V_O = V_1 - 0.6$ for $0.6 \leq V_1 \leq 5$ V; (b) $V_O = 2.4$ V for $0 \leq V_1 \leq 3$ V, $V_O = V_1 - 0.6$ for $3 \leq V_1 \leq 5$ V)

TYU 2.14 Consider the AND logic circuit shown in Figure 2.42. Assume a diode cut-in voltage of $V_y = 0.6$ V. (a) Plot V_O versus V_1 for $0 \leq V_1 \leq 5$ V, if $V_2 = 0$. (b) Repeat part (a) if $V_2 = 3$ V. (Ans. (a) $V_O = 0.6$ V for all V_1 , (b) $V_O = V_1 + 0.6$ for $0 \leq V_1 \leq 3$ V, $V_O = 3.6$ V for $V_1 \geq 3$ V)



2.5 PHOTODIODE AND LED CIRCUITS

Objective: • Understand the operation and characteristics of specialized photodiode and light-emitting diode circuits.

A photodiode converts an optical signal into an electrical current, and a light-emitting diode (LED) transforms an electrical current into an optical signal.

2.5.1 Photodiode Circuit

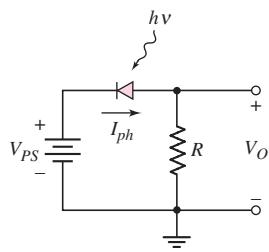


Figure 2.44 A photodiode circuit. The diode is reverse biased

Figure 2.44 shows a typical photodiode circuit in which a reverse-bias voltage is applied to the photodiode. If the photon intensity is zero, the only current through the diode is the reverse-saturation current, which is normally very small. Photons striking the diode create excess electrons and holes in the space-charge region. The electric field quickly separates these excess carriers and sweeps them out of the space-charge region, thus creating a **photocurrent** in the reverse-bias direction. The photocurrent is

$$I_{ph} = \eta e \Phi A \quad (2.39)$$

where η is the quantum efficiency, e is the electronic charge, Φ is the photon flux density (#/ $\text{cm}^2\text{-s}$), and A is the junction area. This linear relationship between photocurrent and photon flux is based on the assumption that the reverse-bias voltage across the diode is constant. This in turn means that the voltage drop across R induced by the photocurrent must be small, or that the resistance R is small.

EXAMPLE 2.12

Objective: Calculate the photocurrent generated in a photodiode.

For the photodiode shown in Figure 2.44 assume the quantum efficiency is 1, the junction area is 10^{-2} cm^2 , and the incident photon flux is $5 \times 10^{17} \text{ cm}^{-2} - \text{s}^{-1}$.

Solution: From Equation (2.39), the photocurrent is

$$I_{ph} = \eta e \Phi A = (1)(1.6 \times 10^{-19})(5 \times 10^{17})(10^{-2}) \Rightarrow 0.8 \text{ mA}$$

Comment: The incident photon flux is normally given in terms of light intensity, in lumens, foot-candles, or W/cm^2 . The light intensity includes the energy of the photons, as well as the photon flux.

EXERCISE PROBLEM

Ex 2.12: (a) Photons with an energy of $h\nu = 2 \text{ eV}$ are incident on the photodiode shown in Figure 2.44. The junction area is $A = 0.5 \text{ cm}^2$, the quantum efficiency is $\eta = 0.8$, and the light intensity is $6.4 \times 10^{-2} \text{ W}/\text{cm}^2$. Determine the photocurrent I_{ph} . (b) If $R = 1 \text{ k}\Omega$, determine the minimum power supply voltage V_{PS} needed to ensure that the diode is reverse biased. (Ans. (a) $I_{ph} = 12.8 \text{ mA}$, (b) $V_{PS}(\min) = 12.8 \text{ V}$)

2.5.2 LED Circuit

A light-emitting diode (LED) is the inverse of a photodiode; that is, a current is converted into an optical signal. If the diode is forward biased, electrons and holes are injected across the space-charge region, where they become excess minority carriers. These excess minority carriers diffuse into the neutral n- and p-regions, where they recombine with majority carriers, and the recombination can result in the emission of a photon.

LEDs are fabricated from compound semiconductor materials, such as gallium arsenide or gallium arsenide phosphide. These materials are direct-bandgap semiconductors. Because these materials have higher bandgap energies than silicon, the forward-bias junction voltage is larger than that in silicon-based diodes.

It is common practice to use a seven-segment LED for the numeric readout of digital instruments, such as a digital voltmeter. The **seven-segment display** is sketched in Figure 2.45. Each segment is an LED normally controlled by IC logic gates.

Figure 2.46 shows one possible circuit connection, known as a common-anode display. In this circuit, the anodes of all LEDs are connected to a 5 V source and the inputs are controlled by logic gates. If V_{I1} is “high,” for example, D_1 is off and there is no light output. When V_{I1} goes “low,” D_1 becomes forward biased and produces a light output.

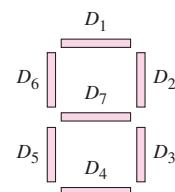


Figure 2.45 Seven-segment LED display

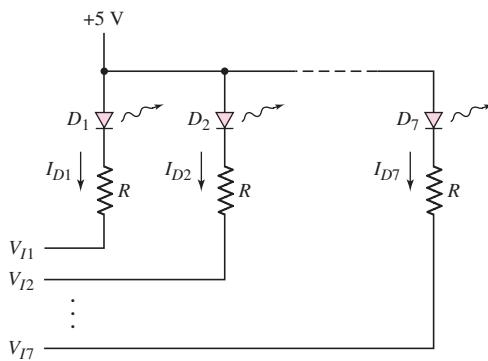


Figure 2.46 Control circuit for the seven-segment LED display

EXAMPLE 2.13

Objective: Determine the value of R required to limit the current in the circuit in Figure 2.46 when the input is in the low state.

Assume that a diode current of 10 mA produces the desired light output, and that the corresponding forward-bias voltage drop is 1.7 V.

Solution: If $V_I = 0.2$ V in the “low” state, then the diode current is

$$I = \frac{5 - V_y - V_I}{R}$$

The resistance R is then determined as

$$R = \frac{5 - V_\gamma - V_I}{I} = \frac{5 - 1.7 - 0.2}{10} \Rightarrow 310 \Omega$$

Comment: Typical LED current-limiting resistor values are in the range 300 to 350 Ω .

EXERCISE PROBLEM

Ex 2.13: Determine the value of resistance R required to limit the current in the circuit shown in Figure 2.46 to $I = 15$ mA. Assume $V_\gamma = 1.7$ V, $r_f = 15 \Omega$, and $V_I = 0.2$ V in the “low” state. (Ans. $R = 192 \Omega$)

One application of LEDs and photodiodes is in **optoisolators**, in which the input signal is electrically decoupled from the output (Figure 2.47). An input signal applied to the LED generates light, which is subsequently detected by the photodiode. The photodiode then converts the light back to an electrical signal. There is no electrical feedback or interaction between the output and input portions of the circuit.

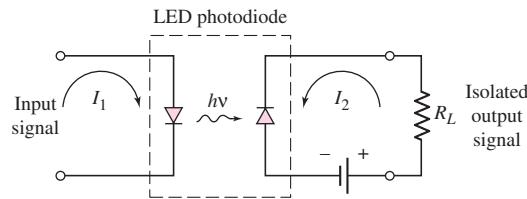


Figure 2.47 Optoisolator using an LED and a photodiode

2.6 DESIGN APPLICATION: DC POWER SUPPLY

Objective: • Design a dc power supply to meet a set of specifications.

Specifications: The output load current is to vary between 25 and 50 mA while the output voltage is to remain in the range $12 \leq v_O \leq 12.2$ V.

Design Approach: The circuit configuration to be designed is shown in Figure 2.48. A diode bridge circuit with an RC filter will be used and a Zener diode will be in parallel with the output load.

Choices: An ac input voltage with an rms value in the range $110 \leq v_I \leq 120$ V and at 60 Hz is available. A Zener diode with a Zener voltage of $V_{ZD} = 12$ V and a Zener resistance of 2Ω that can operate over a current range of $10 \leq I_Z \leq 100$ mA is available. Also, a transformer with an 8:1 turns ratio is available.

Solution: With an 8:1 transformer turns ratio, the peak value of v_S is in the range $19.4 \leq v_S \leq 21.2$ V. Assuming diode turn-on voltages of $V_\gamma = 0.7$ V, the peak value of v_{O1} is in the range $18.0 \leq v_{O1} \leq 19.8$ V.

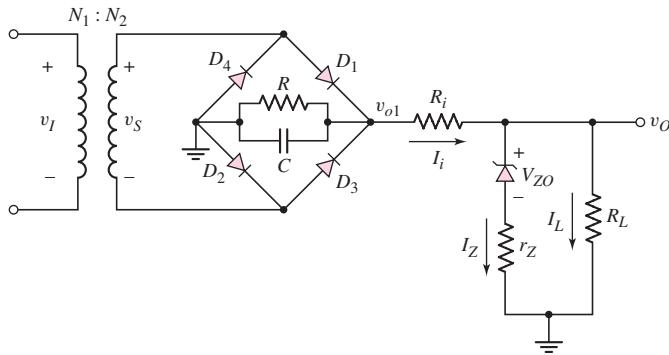


Figure 2.48 DC power supply circuit for design application

For $v_{o1}(\text{max})$ and minimum load current, let $I_Z = 90 \text{ mA}$. Then

$$v_O = V_{ZO} + I_Z r_z = 12 + (0.090)(2) = 12.18 \text{ V}$$

The input current is

$$I_i = I_Z + I_L = 90 + 25 = 115 \text{ mA}$$

The input resistance R_i must then be

$$R_i = \frac{v_{o1} - v_O}{I_i} = \frac{19.8 - 12.18}{0.115} = 66.3 \Omega$$

The minimum Zener current occurs for $I_L(\text{max})$ and $v_{o1}(\text{min})$. The voltage $v_{o1}(\text{min})$ occurs for $v_S(\text{min})$ and must also take into account the ripple voltage. Let $I_Z(\text{min}) = 20 \text{ mA}$. Then the output voltage is

$$v_O = V_{ZO} + I_Z r_z = 12 + (0.020)(2) = 12.04 \text{ V}$$

The output voltage is within the specified range of output voltage.

We now find

$$I_i = I_Z + I_L = 20 + 50 = 70 \text{ mA}$$

and

$$v_{o1}(\text{min}) = I_i R_i + v_O = (0.070)(66.3) + 12.04$$

or

$$v_{o1}(\text{min}) = 16.68 \text{ V}$$

The minimum ripple voltage of the filter is then

$$V_r = v_S(\text{min}) - 1.4 - v_{o1}(\text{min}) = 19.4 - 1.4 - 16.68$$

or

$$V_r = 1.32 \text{ V}$$

Now, let $R_1 = 500 \Omega$. The effective resistance to ground from v_{o1} is $R_1 \parallel R_{i,\text{eff}}$ where $R_{i,\text{eff}}$ is the effective resistance to ground through R_i and the other circuit elements. We can approximate

$$R_{i,\text{eff}} \approx \frac{v_S(\text{avg}) - 1.4}{I_i(\text{max})} = \frac{20.3 - 1.4}{0.115} = 164 \Omega$$

Then $R_1 \parallel R_{i,\text{eff}} = 500 \parallel 164 = 123.5 \Omega$. The required filter capacitance is found from

$$C = \frac{V_M}{2fRV_r} = \frac{19.8}{2(60)(123.5)(1.32)} \Rightarrow 1012 \mu\text{F}$$

Comments: To obtain the proper output voltage in this design, an appropriate Zener diode must be available. We will see in Chapter 9 how an op-amp can be incorporated to provide a more flexible design.

2.7 SUMMARY

- Diode circuits, taking advantage of the nonlinear $i-v$ characteristics of the pn junction, were analyzed and designed in this chapter.
- Half-wave and full-wave rectifier circuits convert a sinusoidal (i.e., ac) signal to an approximate dc signal. A dc power supply, which is used to bias electronic circuits and systems, utilizes these types of circuits. An RC filter can be connected to the output of the rectifier circuit to reduce the ripple effect.
- Zener diodes operate in the reverse-breakdown region and are used in voltage reference or regulator circuits. The percent regulation, a figure of merit for regulator circuits, was defined and determined for various regulator circuits.
- Techniques used to analyze multidiode circuits were developed. The technique requires making assumptions as to whether a diode is conducting (on) or not conducting (off). After analyzing the circuit using these assumptions, we must go back and verify that the assumptions made were valid.
- Diode circuits can be designed to perform basic digital logic functions, such as the AND and OR function. However, there are some inconsistencies between input and output logic values as well as some loading effects, which will severely limit the use of diode logic gates as stand-alone circuits.
- The LED converts electrical current to light and is used extensively in such applications as the seven-segment alphanumeric display. Conversely, the photodiode detects an incident light signal and transforms it into an electrical current.
- As an application, a simple dc power supply was designed using a rectifier circuit in conjunction with a Zener diode.

CHECKPOINT

After studying this chapter, the reader should have the ability to:

- ✓ In general, apply the diode piecewise linear model in the analysis of diode circuits.
- ✓ Analyze diode rectifier circuits, including the calculation of ripple voltage.
- ✓ Analyze Zener diode circuits, including the effect of a Zener resistance.
- ✓ Determine the output signal for a given input signal of diode clipper and clammer circuits.
- ✓ Analyze circuits with multiple diodes by making initial assumptions and then verifying these initial assumptions.

REVIEW QUESTIONS

- What characteristic of a diode is used in the design of diode signal processing circuits?
- Describe a simple half-wave diode rectifier circuit and sketch the output voltage versus time.
- Describe a simple full-wave diode rectifier circuit and sketch the output voltage versus time.
- What is the advantage of connecting an RC filter to the output of a diode rectifier circuit?
- Define ripple voltage. How can the magnitude of the ripple voltage be reduced?
- Describe a simple Zener diode voltage reference circuit.
- What effect does the Zener diode resistance have on the voltage reference circuit operation? Define load regulation.
- What are the general characteristics of diode clipper circuits?
- Describe a simple diode clipper circuit that limits the negative portion of a sinusoidal input voltage to a specified value.
- What are the general characteristics of diode clamper circuits?
- What one circuit element, besides a diode, is present in all diode clamper circuits?
- Describe the procedure used in the analysis of a circuit containing two diodes. How many initial assumptions concerning the state of the circuit are possible?
- Describe a diode OR logic circuit. Compare a logic 1 value at the output compared to a logic 1 value at the input. Are they the same value?
- Describe a diode AND logic circuit. Compare a logic 0 value at the output compared to a logic 0 value at the input. Are they the same value?
- Describe a simple circuit that can be used to turn an LED on or off with a high or low input voltage.

PROBLEMS

[Note: In the following problems, assume $r_f = 0$ unless otherwise specified.]

Section 2.1 Rectifier Circuits

- Consider the circuit shown in Figure P2.1. Let $R = 1\text{ k}\Omega$, $V_y = 0.6\text{ V}$, and $r_f = 20\text{ }\Omega$. (a) Plot the voltage transfer characteristics v_o versus v_I over the range $-10 \leq v_I \leq 10\text{ V}$. (b) Assume $v_I = 10 \sin \omega t\text{ (V)}$. (i) Sketch v_o versus time for the sinusoidal input. (ii) Find the average value of v_o . (iii) Determine the peak diode current. (iv) What is the PIV of the diode?
- For the circuit shown in Figure P2.1, show that for $v_I \geq 0$, the output voltage is approximately given by

$$v_o = v_I - V_T \ln\left(\frac{v_o}{I_S R}\right)$$

- A half-wave rectifier such as shown in Figure 2.2(a) has a $2\text{ k}\Omega$ load. The input is a 120 V (rms), 60 Hz signal and the transformer is a $10:1$ step-down transformer. The diode has a cut-in voltage of $V_y = 0.7\text{ V}$ ($r_f = 0$). (a) What is the peak output voltage? (b) Determine the peak diode current.

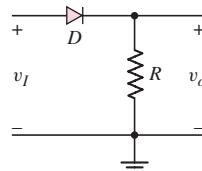


Figure P2.1

- (c) What is the fraction (percent) of a cycle that $v_O > 0$. (d) Determine the average output voltage. (e) Find the average current in the load.
- 2.4 Consider the battery charging circuit shown in Figure 2.4(a). Assume that $V_B = 9$ V, $V_S = 15$ V, and $\omega = 2\pi(60)$. (a) Determine the value of R such that the average battery charging current is $i_D = 0.8$ A. (b) Find the fraction of time that the diode is conducting.
- 2.5 Figure P2.5 shows a simple full-wave battery charging circuit. Assume $V_B = 9$ V, $V_Y = 0.7$ V, and $v_S = 15 \sin[2\pi(60)t]$ (V). (a) Determine R such that the peak battery charging current is 1.2 A. (b) Determine the average battery charging current. (c) Determine the fraction of time that each diode is conducting.

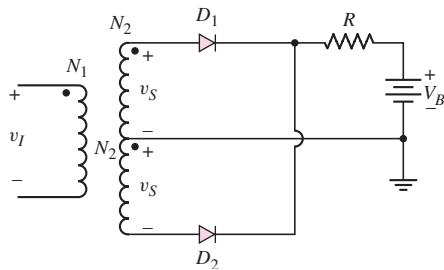


Figure P2.5

- 2.6 The full-wave rectifier circuit shown in Figure 2.5(a) in the text is to deliver 0.2 A and 12 V (peak values) to a load. The ripple voltage is to be limited to 0.25 V. The input signal is 120 V (rms) at 60 Hz. Assume diode cut-in voltages of 0.7 V. (a) Determine the required turns ratio of the transformer. (b) Find the required value of the capacitor. (c) What is the PIV rating of the diode?
- 2.7 The input signal voltage to the full-wave rectifier circuit in Figure 2.6(a) in the text is $v_I = 160 \sin[2\pi(60)t]$ V. Assume $V_Y = 0.7$ V for each diode. Determine the required turns ratio of the transformer to produce a peak output voltage of (a) 25 V, and (b) 100 V. (c) What must be the diode PIV rating for each case?
- 2.8 The output resistance of the full-wave rectifier in Figure 2.6(a) in the text is $R = 150 \Omega$. A filter capacitor is connected in parallel with R . Assume $V_Y = 0.7$ V. The peak output voltage is to be 12 V and the ripple voltage is to be no more than 0.3 V. The input frequency is 60 Hz. (a) Determine the required rms value of v_S . (b) Determine the required filter capacitance value. (c) Determine the peak current through each diode.
- 2.9 Repeat Problem 2.8 for the half-wave rectifier in Figure 2.2(a).
- 2.10 Consider the half-wave rectifier circuit shown in Figure 2.8(a) in the text. Assume $v_S = 10 \sin[2\pi(60)t]$ (V), $V_Y = 0.7$ V, and $R = 500 \Omega$. (a) What is the peak output voltage? (b) Determine the value of capacitance required such that the ripple voltage is no more than $V_r = 0.5$ V. (c) What is the PIV rating of the diode?
- 2.11 The parameters of the half-wave rectifier circuit in Figure 2.8(a) in the text are $R = 1 \text{ k}\Omega$, $C = 350 \mu\text{F}$, and $V_Y = 0.7$ V. Assume $v_S(t) = V_S \sin[2\pi(60)t]$ (V) where V_S is in the range of $11 \leq V_S \leq 13$ V. (a) What is the range in output voltage? (b) Determine the range in ripple voltage. (c) If the ripple voltage is to be limited to $V_r = 0.4$ V, determine the minimum value of capacitance required.

- 2.12 The full-wave rectifier circuit shown in Figure P2.12 has an input signal whose frequency is 60 Hz. The rms value of $v_s = 8.5$ V. Assume each diode cut-in voltage is $V_y = 0.7$ V. (a) What is the maximum value of V_o ? (b) If $R = 10 \Omega$, determine the value of C such that the ripple voltage is no larger than 0.25 V. (c) What must be the PIV rating of each diode?

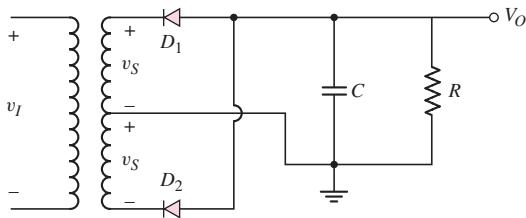


Figure P2.12

- 2.13 Consider the full-wave rectifier circuit in Figure 2.7 of the text. The output resistance is $R_L = 125 \Omega$, each diode cut-in voltage is $V_y = 0.7$ V, and the frequency of the input signal is 60 Hz. A filter capacitor is connected in parallel with R_L . The magnitude of the peak output voltage is to be 15 V and the ripple voltage is to be no more than 0.35 V. (a) Determine the rms value of v_s and (b) the required value of the capacitor.
- 2.14 The circuit in Figure P2.14 is a complementary output rectifier. If $v_s = 26 \sin [2\pi(60)t]$ V, sketch the output waveforms v_o^+ and v_o^- versus time, assuming $V_y = 0.6$ V for each diode.

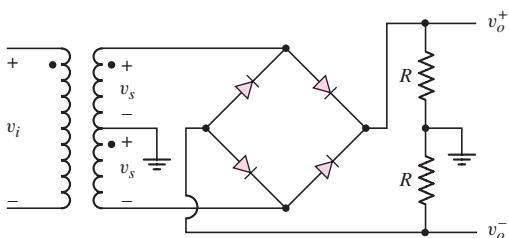


Figure P2.14

- 2.15 A full-wave rectifier is to be designed using the center-tapped transformer configuration. The peak output voltage is to be 12 V, the nominal load current is to be 0.5 A, and the ripple voltage is to be limited to 3 percent. Assume $V_y = 0.8$ V and let $v_i = 120\sqrt{2} \sin[2\pi(60)t]$ V. (a) What is the transformer turns ratio? (b) What is the minimum value of C required? (c) What is the peak diode current? (d) Determine the average diode current. (e) What is the PIV rating of the diodes.
- 2.16 A full-wave rectifier is to be designed using the bridge circuit configuration. The peak output voltage is to be 9 V, the nominal load current is to be 100 mA, and the ripple voltage is to be limited to $V_r = 0.2$ V. Assume $V_y = 0.8$ V and let $v_i = 120\sqrt{2} \sin[2\pi(60)t]$ (V). (a) What is the transformer turns ratio? (b) What is the minimum value of C required? (c) What

is the peak diode current? (d) Determine the average diode current. (e) What is the PIV rating of the diodes.

- *2.17 Sketch v_o versus time for the circuit in Figure P2.17 with the input shown. Assume $V_\gamma = 0$.

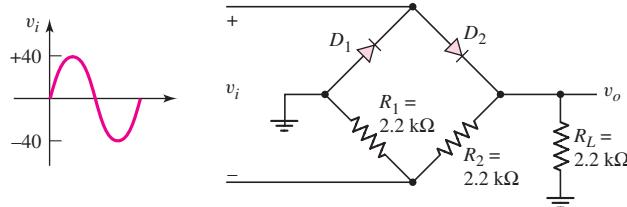


Figure P2.17

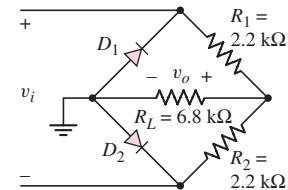


Figure P2.18

- *2.18 (a) Sketch v_o versus time for the circuit in Figure P2.18. The input is a sine wave given by $v_i = 10 \sin \omega t$ V. Assume $V_\gamma = 0$. (b) Determine the rms value of the output voltage.

Section 2.2 Zener Diode Circuits

- 2.19 Consider the circuit shown in Figure P2.19. The Zener diode voltage is $V_Z = 3.9$ V and the Zener diode incremental resistance is $r_z = 0$. (a) Determine I_Z , I_L , and the power dissipated in the diode. (b) Repeat part (a) if the 4 kΩ load resistor is increased to 10 kΩ.

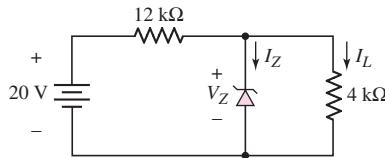


Figure P2.19

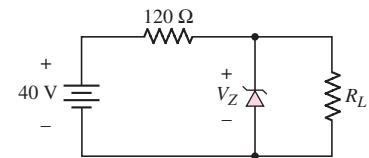


Figure P2.20

- 2.20 Consider the Zener diode circuit shown in Figure P2.20. Assume $V_Z = 12$ V and $r_z = 0$. (a) Calculate the Zener diode current and the power dissipated in the Zener diode for $R_L = \infty$. (b) What is the value of R_L such that the current in the Zener diode is one-tenth of the current supplied by the 40 V source? (c) Determine the power dissipated in the Zener diode for the conditions of part (b).

- 2.21 Consider the Zener diode circuit shown in Figure P2.21. Let $V_I = 60$ V, $R_I = 150 \Omega$, and $V_{ZO} = 15.4$ V. Assume $r_z = 0$. The power rating of the diode is 4 W and the minimum diode current is to be 15 mA. (a) Determine the range of diode currents. (b) Determine the range of load resistance.

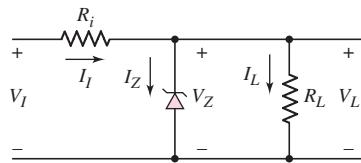


Figure P2.21

- *2.22 In the voltage regulator circuit in Figure P2.21, $V_I = 20$ V, $V_Z = 10$ V, $R_i = 222 \Omega$, and $P_Z(\text{max}) = 400$ mW. (a) Determine I_L , I_Z , and I_I , if $R_L = 380 \Omega$. (b) Determine the value of R_L that will establish $P_Z(\text{max})$ in the diode. (c) Repeat part (b) if $R_i = 175 \Omega$.
- 2.23 A Zener diode is connected in a voltage regulator circuit as shown in Figure P2.21. The Zener voltage is $V_Z = 10$ V and the Zener resistance is assumed to be $r_z = 0$. (a) Determine the value of R_i such that the Zener diode remains in breakdown if the load current varies from $I_L = 50$ to 500 mA and if the input voltage varies from $V_I = 15$ to 20 V. Assume $I_Z(\text{min}) = 0.1I_Z(\text{max})$. (b) Determine the power rating required for the Zener diode and the load resistor.
- 2.24 Consider the Zener diode circuit in Figure 2.19 in the text. Assume parameter values of $V_{ZO} = 5.6$ V (diode voltage when $I_Z \approx 0$), $r_z = 3 \Omega$, and $R_i = 50 \Omega$. Determine V_L , I_Z , I_L , and the power dissipated in the diode for (a) $V_{PS} = 10$ V, $R_L = \infty$; (b) $V_{PS} = 10$ V, $R_L = 200 \Omega$; (c) $V_{PS} = 12$ V, $R_L = \infty$; and (d) $V_{PS} = 12$ V, $R_L = 200 \Omega$.
- D2.25 Design a voltage regulator circuit such as shown in Figure P2.21 so that $V_L = 7.5$ V. The Zener diode voltage is $V_Z = 7.5$ V at $I_Z = 10$ mA. The incremental diode resistance is $r_z = 12 \Omega$. The nominal supply voltage is $V_I = 12$ V and the nominal load resistance is $R_L = 1 \text{k}\Omega$. (a) Determine R_i . (b) If V_I varies by ± 10 percent, calculate the source regulation. What is the variation in output voltage? (c) If R_L varies over the range of $1 \text{k}\Omega \leq R_L \leq \infty$, what is the variation in output voltage? Determine the load regulation.
- 2.26 The percent regulation of the Zener diode regulator shown in Figure 2.16 is 5 percent. The Zener voltage is $V_{ZO} = 6$ V and the Zener resistance is $r_z = 3 \Omega$. Also, the load resistance varies between 500 and 1000 Ω , the input resistance is $R_i = 280 \Omega$, and the minimum power supply voltage is $V_{PS}(\text{min}) = 15$ V. Determine the maximum power supply voltage allowed.
- *2.27 A voltage regulator is to have a nominal output voltage of 10 V. The specified Zener diode has a rating of 1 W, has a 10 V drop at $I_Z = 25$ mA, and has a Zener resistance of $r_z = 5 \Omega$. The input power supply has a nominal value of $V_{PS} = 20$ V and can vary by ± 25 percent. The output load current is to vary between $I_L = 0$ and 20 mA. (a) If the minimum Zener current is to be $I_Z = 5$ mA, determine the required R_i . (b) Determine the maximum variation in output voltage. (c) Determine the percent regulation.
- *2.28 Consider the circuit in Figure P2.28. Let $V_Y = 0$. The secondary voltage is given by $v_s = V_s \sin \omega t$, where $V_s = 24$ V. The Zener diode has parameters $V_Z = 16$ V at $I_Z = 40$ mA and $r_z = 2 \Omega$. Determine R_i such that the load current can vary over the range $40 \leq I_L \leq 400$ mA with $I_Z(\text{min}) = 40$ mA and find C such that the ripple voltage is no larger than 1 V.

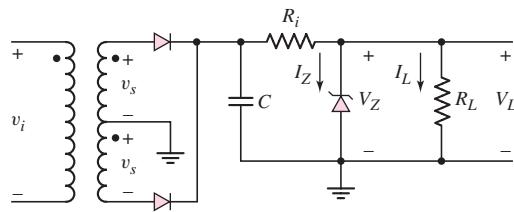


Figure P2.28

- *2.29 The secondary voltage in the circuit in Figure P2.28 is $v_s = 12 \sin \omega t$ V. The Zener diode has parameters $V_Z = 8$ V at $I_Z = 100$ mA and $r_z = 0.5 \Omega$. Let $V_\gamma = 0$ and $R_i = 3 \Omega$. Determine the percent regulation for load currents between $I_L = 0.2$ and 1 A. Find C such that the ripple voltage is no larger than 0.8 V.

Section 2.3 Clipper and Clamper Circuits

- 2.30 The parameters in the circuit shown in Figure P2.30 are $V_\gamma = 0.7$ V, $V_{Z1} = 2.3$ V, and $V_{Z2} = 5.6$ V. Plot v_O versus v_I over the range of $-10 \leq v_I \leq +10$ V.

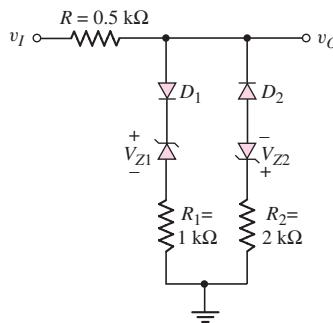


Figure P2.30

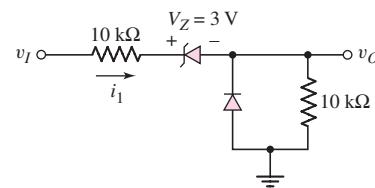


Figure P2.31

- 2.31 Consider the circuit in Figure P2.31. Let $V_\gamma = 0$. (a) Plot v_O versus v_I over the range $-10 \leq v_I \leq +10$ V. (b) Plot i_1 over the same input voltage range as part (a).
 2.32 For the circuit in Figure P2.32, (a) plot v_O versus v_I for $0 \leq v_I \leq 15$ V. Assume $V_\gamma = 0.7$ V. Indicate all breakpoints. (b) Plot i_D over the same range of input voltage. (c) Compare the results of parts (a) and (b) with a computer simulation.

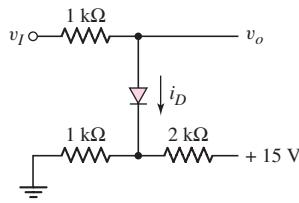


Figure P2.32

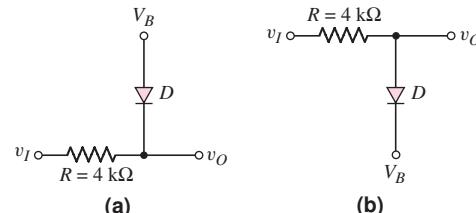


Figure P2.33

- 2.33 Each diode cut-in voltage is 0.7 V for the circuits shown in Figure P2.33. (a) Plot v_O versus v_I over the range $-5 \leq v_I \leq +5$ V for the circuit in Figure P2.33(a) for (i) $V_B = 1.8$ V and (ii) $V_B = -1.8$ V. (b) Repeat part (a) for the circuit shown in Figure P2.33(b).
 *2.34 The diode in the circuit of Figure P2.34(a) has piecewise linear parameters $V_\gamma = 0.7$ V and $r_f = 10 \Omega$. (a) Plot v_O versus v_I for $-30 \leq v_I \leq 30$ V. (b) If the triangular wave, shown in Figure P2.34(b), is applied, plot the output versus time.

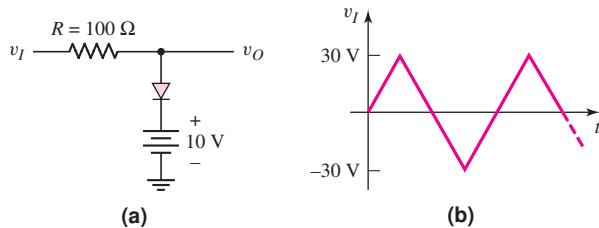


Figure P2.34

- 2.35 Consider the circuits shown in Figure P2.35. Each diode cut-in voltage is $V_\gamma = 0.7$ V. (a) Plot v_O versus v_I over the range $-10 \leq v_I \leq +10$ V for the circuit in Figure P2.35(a) for (i) $V_B = 5$ V and (ii) $V_B = -5$ V. (b) Repeat part (a) for the circuit in Figure P2.35(b).

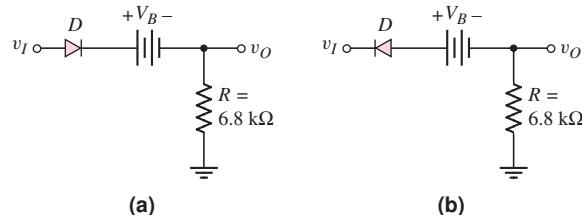


Figure P2.35

- 2.36 Plot v_O for each circuit in Figure P2.36 for the input shown. Assume
 (a) $V_Y = 0$ and (b) $V_Y = 0.6$ V.

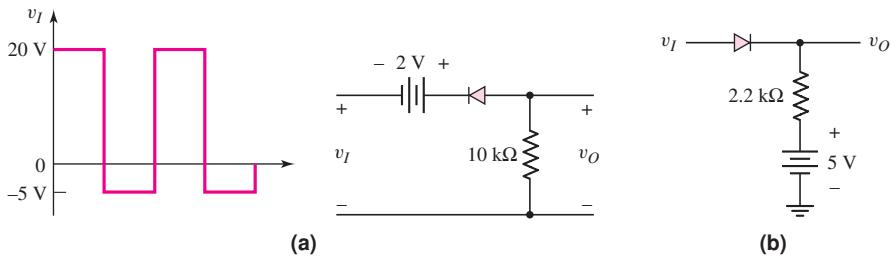


Figure P2.36

- 2.37 Consider the parallel clipper circuit in Figure 2.26 in the text. Assume $V_{Z1} = 6$ V, $V_{Z2} = 4$ V, and $V_y = 0.7$ V for all diodes. For $v_I = 10 \sin \omega t$, sketch v_O versus time over two periods of the input signal.

*2.38 A car's radio may be subjected to voltage spikes induced by coupling from the ignition system. Pulses on the order of ± 250 V and lasting for $120 \mu s$ may exist. Design a clipper circuit using resistors, diodes, and Zener diodes to limit the input voltage between $+14$ V and -0.7 V. Specify power ratings of the components.

- 2.39 Sketch the steady-state output voltage v_O versus time for each circuit in Figure P2.39 with the input voltage given in Figure P2.39(a). Assume $V_\gamma = 0$ and assume the RC time constant is large.

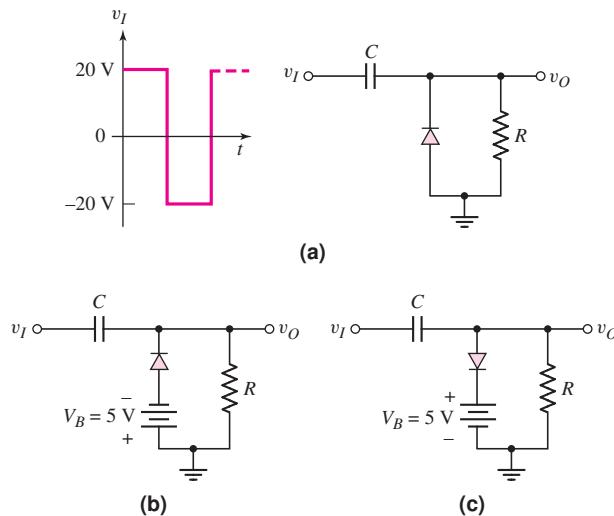


Figure P2.39

- D2.40 Design a diode clamper to generate a steady-state output voltage v_O from the input voltage v_I shown in Figure P2.40 if (a) $V_\gamma = 0$ and (b) $V_\gamma = 0.7$ V.

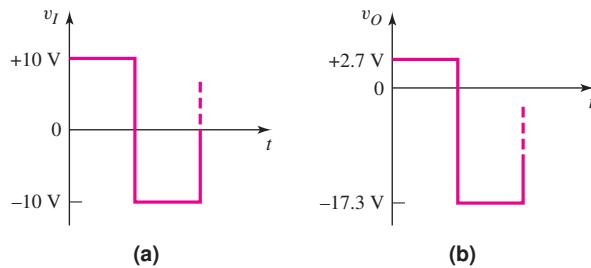


Figure P2.40

- D2.41 Design a diode clamper to generate a steady-state output voltage v_O from the input voltage v_I in Figure P2.41 if $V_\gamma = 0$.

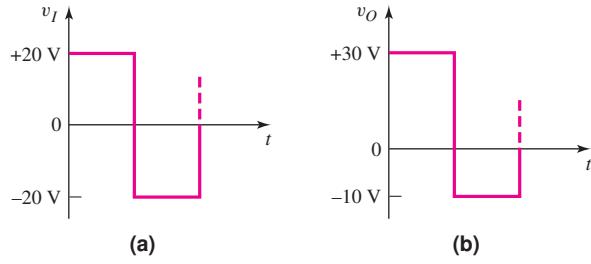


Figure P2.41

- 2.42 For the circuit in Figure P2.39(b), let $V_Y = 0$ and $v_I = 10 \sin \omega t$ (V). Plot v_O versus time over three cycles of input voltage. Assume the initial voltage across the capacitor is zero and assume the RC time constant is very large.
- 2.43 Repeat Problem 2.42 for the circuit in Figure P2.39(c) for (i) $V_B = 5$ V and (ii) $V_B = -5$ V.

Section 2.4 Multiple Diode Circuits

- 2.44 The diodes in the circuit in Figure P2.44 have piecewise linear parameters of $V_Y = 0.6$ V and $r_f = 0$. Determine the output voltage V_O and the diode currents I_{D1} and I_{D2} for the following input conditions: (a) $V_1 = 10$ V, $V_2 = 0$; (b) $V_1 = 5$ V, $V_2 = 0$; (c) $V_1 = 10$ V, $V_2 = 5$ V; and (d) $V_1 = V_2 = 10$ V. (e) Compare the results of parts (a) through (d) with a computer simulation analysis.
- 2.45 In the circuit in Figure P2.45 the diodes have the same piecewise linear parameters as described in Problem 2.44. Calculate the output voltage V_O and the currents I_{D1} , I_{D2} , and I for the following input conditions: (a) $V_1 = V_2 = 10$ V; (b) $V_1 = 10$ V, $V_2 = 0$; (c) $V_1 = 10$ V, $V_2 = 5$ V; and (d) $V_1 = V_2 = 0$.

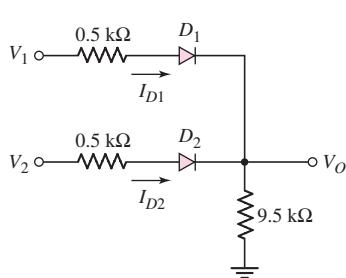


Figure P2.44

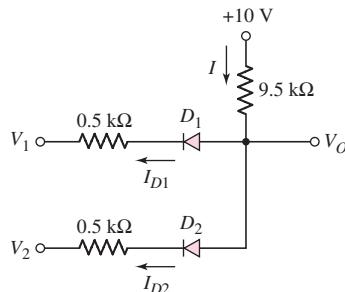


Figure P2.45

- 2.46 The diodes in the circuit in Figure P2.46 have the same piecewise linear parameters as described in Problem 2.44. Determine the output voltage V_O and the currents I_{D1} , I_{D2} , I_{D3} , and I for the following input conditions: (a) $V_1 = V_2 = 0$; (b) $V_1 = V_2 = 5$ V; (c) $V_1 = 5$ V, $V_2 = 0$; and (d) $V_1 = 5$ V, $V_2 = 2$ V.

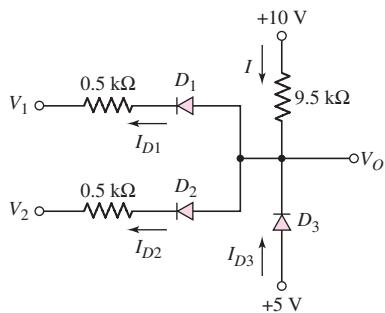
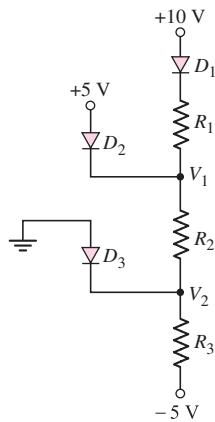


Figure P2.46



- 2.47 Consider the circuit shown in Figure P2.47. Assume each diode cut-in voltage is $V_\gamma = 0.6$ V. (a) Determine R_1 , R_2 , and R_3 such that $I_{D1} = 0.2$ mA, $I_{D2} = 0.3$ mA, and $I_{D3} = 0.5$ mA. (b) Find V_1 , V_2 , and each diode current for $R_1 = 10$ k Ω , $R_2 = 4$ k Ω , and $R_3 = 2.2$ k Ω . (c) Repeat part (b) for $R_1 = 3$ k Ω , $R_2 = 6$ k Ω , and $R_3 = 2.5$ k Ω . (d) Repeat part (b) for $R_1 = 6$ k Ω , $R_2 = 3$ k Ω , and $R_3 = 6$ k Ω .

- 2.48 The diode cut-in voltage for each diode in the circuit shown in Figure P2.48 is 0.7 V. Determine the value of R such that (a) $I_{D1} = I_{D2}$, (b) $I_{D1} = 0.2I_{D2}$, and (c) $I_{D1} = 5I_{D2}$.

- 2.49 Consider the circuit in Figure P2.49. Each diode cut-in voltage is $V_\gamma = 0.7$ V. (a) For $R_2 = 1.1$ k Ω , determine I_{D1} , I_{D2} , and V_A . (b) Repeat part (a) for $R_2 = 2.5$ k Ω . (c) Find R_2 such that $V_A = 0$. What are the values of I_{D1} and I_{D2} ?

Figure P2.47

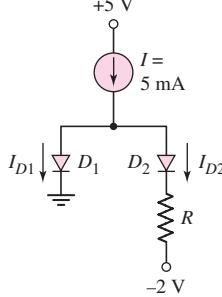


Figure P2.48

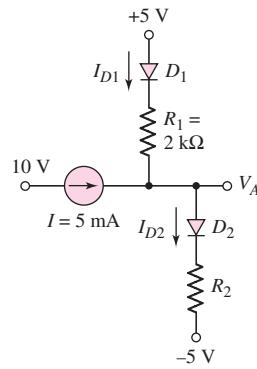
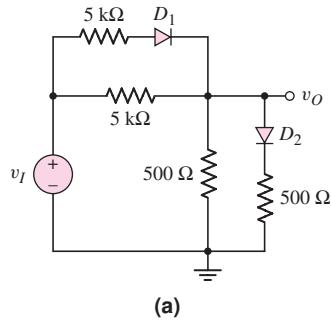
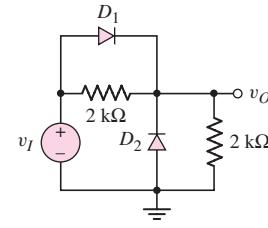


Figure P2.49

- 2.50 In each circuit shown in Figure P2.50, the diode cut-in voltage is $V_\gamma = 0.6$ V. (a) For the circuit in Figure P2.50(a), determine v_O for (i) $v_I = +5$ V and (ii) $v_I = -5$ V. (b) Repeat part (a) for the circuit in Figure P2.50(b). (c) Plot the voltage transfer characteristics, v_O versus v_I , of each circuit over the range $-5 \leq v_I \leq +5$ V.



(a)



(b)

Figure P2.50

- *2.51 Assume $V_\gamma = 0.7$ V for each diode in the circuit in Figure P2.51. Plot v_O versus v_I for $-10 \leq v_I \leq +10$ V.

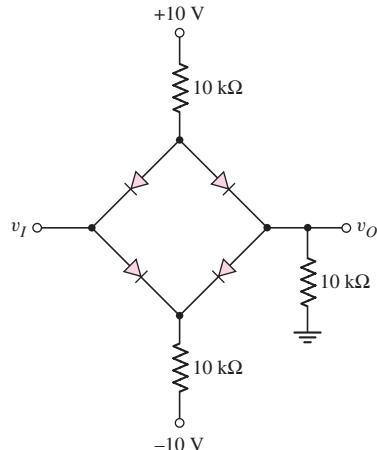


Figure P2.51

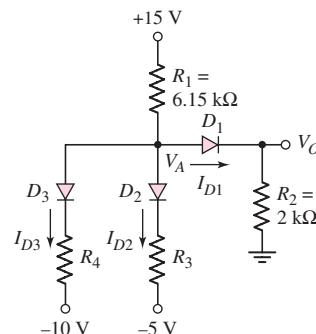


Figure P2.52

- 2.52 The cut-in voltage of each diode in the circuit shown in Figure P2.52 is $V_\gamma = 0.7$ V. Determine I_{D1} , I_{D2} , I_{D3} , and V_A for (a) $R_3 = 14$ k Ω , $R_4 = 24$ k Ω ; (b) $R_3 = 3.3$ k Ω , $R_4 = 5.2$ k Ω ; and (c) $R_3 = 3.3$ k Ω , $R_4 = 1.32$ k Ω .
- 2.53 Let $V_\gamma = 0.7$ V for each diode in the circuit in Figure P2.53. (a) Find I_{D1} and V_O for $R_1 = 5$ k Ω and $R_2 = 10$ k Ω . (b) Repeat part (a) for $R_1 = 10$ k Ω and $R_2 = 5$ k Ω .

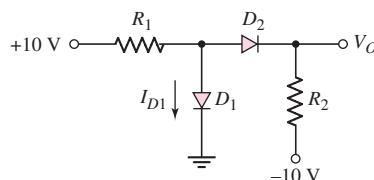


Figure P2.53

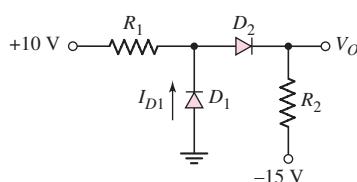


Figure P2.54

- 2.54 For the circuit shown in Figure P2.54, let $V_\gamma = 0.7$ V for each diode. Calculate I_{D1} and V_O for (a) $R_1 = 10$ k Ω , $R_2 = 5$ k Ω and for (b) $R_1 = 5$ k Ω , $R_2 = 10$ k Ω .
- 2.55 Assume each diode cut-in voltage is $V_\gamma = 0.7$ V for the circuit in Figure P2.55. Determine I_{D1} and V_O for (a) $R_1 = 10$ k Ω , $R_2 = 5$ k Ω and (b) $R_1 = 5$ k Ω , $R_2 = 10$ k Ω .

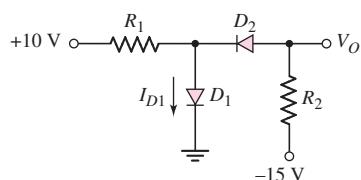


Figure P2.55

- 2.56 If $V_\gamma = 0.7$ V for the diode in the circuit in Figure P2.56 determine I_D and V_O .

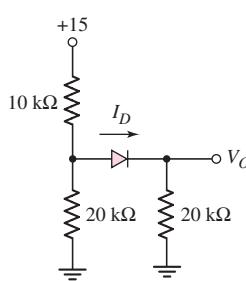


Figure P2.56

- 2.57 Let $V_\gamma = 0.7$ V for the diode in the circuit in Figure P2.57. Determine I_D , V_D , V_A , and V_B for (a) $V_1 = V_2 = 6$ V; (b) $V_1 = 2$ V, $V_2 = 5$ V; (c) $V_1 = 5$ V, $V_2 = 4$ V; and (d) $V_1 = 2$ V, $V_2 = 8$ V.

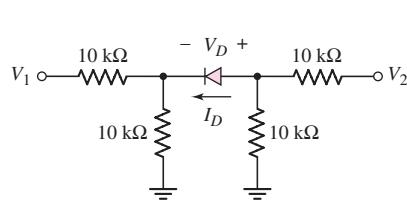


Figure P2.57

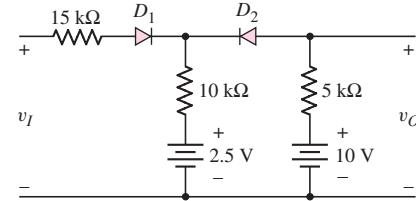


Figure P2.58

- 2.58 (a) Each diode in the circuit in Figure P2.58 has piecewise linear parameters of $V_\gamma = 0$ and $r_f = 0$. Plot v_O versus v_I for $0 \leq v_I \leq 30$ V. Indicate the breakpoints and give the state of each diode in the various regions of the plot.
 (b) Compare the results of part (a) with a computer simulation analysis.
- 2.59 Each diode cut-in voltage in the circuit in Figure P2.59 is 0.7 V. Determine I_{D1} , I_{D2} , I_{D3} , and v_O for (a) $v_I = 0.5$ V, (b) $v_I = 1.5$ V, (c) $v_I = 3.0$ V, and (d) $v_I = 5.0$ V.

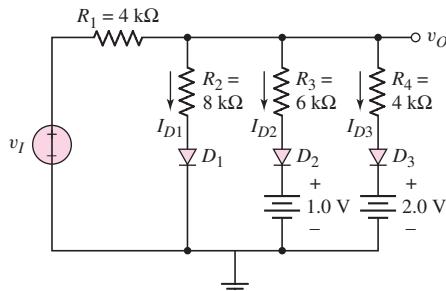


Figure P2.59

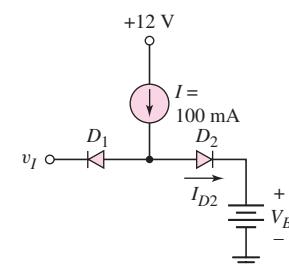


Figure P2.60

- 2.60 Let $V_\gamma = 0.7$ V for each diode in the circuit shown in Figure P2.60. Plot I_{D2} versus v_I over the range $0 \leq v_I \leq 12$ V for (a) $V_B = 4.5$ V and (b) $V_B = 9$ V.
- 2.61 Consider the circuit in Figure P2.61. The output of a diode OR logic gate is connected to the input of a second diode OR logic gate. Assume $V_\gamma = 0.6$ V for each diode. Determine the outputs V_{O1} and V_{O2} for: (a) $V_1 = V_2 = 0$; (b) $V_1 = 5$ V, $V_2 = 0$; and (c) $V_1 = V_2 = 5$ V. What can be said about the relative values of V_{O1} and V_{O2} in their “high” state?

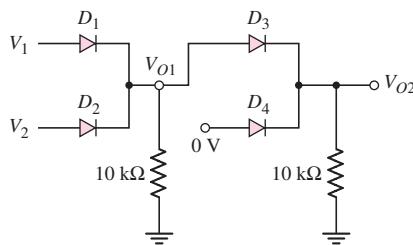


Figure P2.61

- 2.62 Consider the circuit in Figure P2.62. The output of a diode AND logic gate is connected to the input of a second diode AND logic gate. Assume $V_y = 0.6$ V for each diode. Determine the outputs V_{O1} and V_{O2} for: (a) $V_1 = V_2 = 5$ V; (b) $V_1 = 0$, $V_2 = 5$ V; and (c) $V_1 = V_2 = 0$. What can be said about the relative values of V_{O1} and V_{O2} in their “low” state?

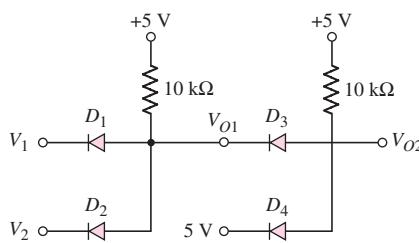


Figure P2.62

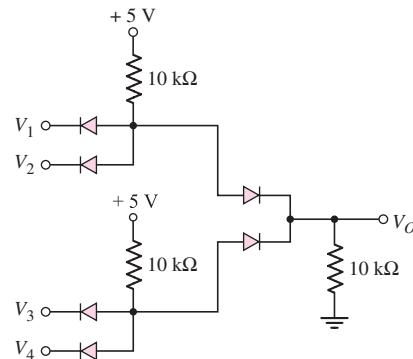


Figure P2.63

- 2.63 Determine the Boolean expression for V_O in terms of the four input voltages for the circuit in Figure P2.63 (Hint: A truth table might be helpful.)

Section 2.5 LED and Photodiode Circuits

- 2.64 Consider the circuit shown in Figure P2.64. The forward-bias cut-in voltage of the diode is 1.5 V and the forward-bias resistance is $r_f = 10 \Omega$. Determine the value of R required to limit the current to $I = 12$ mA when $V_I = 0.2$ V.
- 2.65 The light-emitting diode in the circuit shown in Figure P2.64 has parameters $V_y = 1.7$ V and $r_f = 0$. Light will first be detected when the current is $I = 8$ mA. If $R = 750 \Omega$, determine the value of V_I at which light will first be detected.
- 2.66 The parameters of D_1 and D_2 in the circuit shown in Figure P2.66 are $V_y = 1.7$ V and $r_f = 20 \Omega$. The current in each diode is to be limited to $I_D = 15$ mA for $V_I = \pm 5$ V. Determine the required value of R .
- 2.67 If the resistor in Example 2.12 is $R = 2 \text{ k}\Omega$ and the diode is to be reverse biased by at least 1 V, determine the minimum power supply voltage required.
- 2.68 Consider the photodiode circuit shown in Figure 2.44. Assume the quantum efficiency is 1. A photocurrent of 0.6 mA is required for an incident photon flux of $\Phi = 10^{17} \text{ cm}^{-2}\text{s}^{-1}$. Determine the required cross-sectional area of the diode.

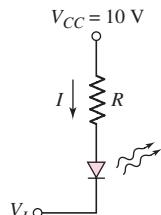


Figure P2.64

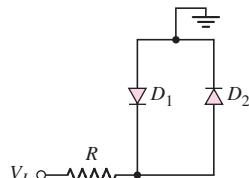


Figure P2.66

COMPUTER SIMULATION PROBLEMS

- 2.69 Consider the voltage doubler circuit in Figure 2.14. Assume a 60 Hz, 120 V (rms) signal is applied at the input of the transformer with a 20:1 turns ratio. Let $R = 10 \text{ k}\Omega$ and $C_1 = C_2 = 200 \mu\text{F}$. Using a computer simulation, plot the output voltage over four cycles of input voltage.

- 2.70 Consider the parameters and results of Example 2.2. Use a computer simulation to plot the output voltage of each rectifier over four cycles of input voltage. Also determine the PIV of each diode. How do the computer results compare with the results of the example?
- 2.71 (a) Using a computer simulation, verify the results of Exercise TYU2.3.
 (b) Determine the ripple voltage if a filter capacitance of $C = 50 \mu\text{F}$ is connected in parallel with the load resistance.
- 2.72 (a) Using a computer simulation, determine each diode current and voltage in the circuit shown in Figure 2.40. (b) Repeat part (a) using the circuit parameters given in Exercise 2.11.

DESIGN PROBLEMS

- *D2.73 Consider the full-wave bridge rectifier circuit. The input signal is 120 V (rms) at 60 Hz. The load resistance is $R_L = 250 \Omega$. The peak output voltage is to be 9 V and the ripple voltage is to be no more than 5 percent. Determine the required turns ratio and the required value of filter capacitance.
- *D2.74 Design a simple dc voltage source using a 120 V (rms), 60 Hz input signal to a nominal 10 V output signal. A Zener diode with parameters $V_{Z0} = 10 \text{ V}$ and $r_z = 3 \Omega$ is available. The rated power of the Zener diode is 5 W. The source regulation is to be limited to 2 percent.
- *D2.75 A clipper is to be designed such that $v_O = 2.5 \text{ V}$ for $v_I \geq 2.5 \text{ V}$ and $v_O = -1.25 \text{ V}$ for $v_I \leq -1.25 \text{ V}$.
- *D2.76 Design a circuit to provide the voltage transfer characteristics shown in Figure P2.76. Use diodes and Zener diodes with appropriate breakdown voltages in the design. The maximum current in the circuit is to be limited to 1mA.

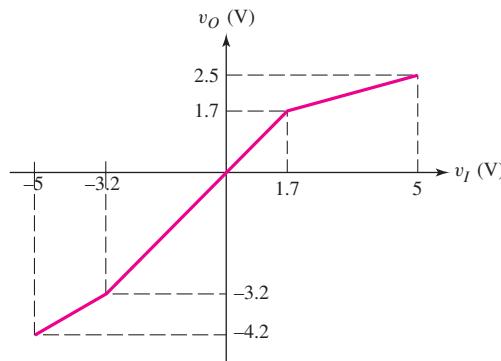


Figure P2.76

The Field-Effect Transistor

In this chapter, we introduce a major type of transistor, the metal-oxide-semiconductor field-effect transistor (MOSFET). The MOSFET led to the electronics revolution of the 1970s and 1980s, in which the microprocessor made possible powerful desktop computers, laptop computers, sophisticated handheld calculators, iPods, and a plethora of other electronic systems. The MOSFET can be made very small, so high-density very large scale integration (VLSI) circuits and high-density memories are possible.

Two complementary devices, the n-channel MOSFET (NMOS) and the p-channel MOSFET (PMOS), exist. Each device is equally important and allows a high degree of flexibility in electronic circuit design. The $i-v$ characteristics of these devices are introduced, and the dc analysis and design techniques of MOSFET circuits are developed.

Another type of field-effect transistor is the junction FET. There are two general categories of junction field-effect transistors (JFETs)—the pn junction FET (pn JFET) and the metal-semiconductor field-effect transistor (MESFET), which is fabricated with a Schottky barrier junction. JFETs were developed before MOSFETs, but the applications and uses of MOSFETs have far surpassed those of the JFET. However, we will consider a few JFET circuits in this chapter.

PREVIEW

In this chapter, we will:

- Study and understand the structure, operation, and characteristics of the various types of MOSFETs.
- Understand and become familiar with the dc analysis and design techniques of MOSFET circuits.
- Examine three applications of MOSFET circuits.
- Investigate current source biasing of MOSFET circuits, such as those used in integrated circuits.
- Analyze the dc biasing of multistage or multitransistor circuits.
- Understand the operation and characteristics of the junction field-effect transistor, and analyze the dc response of JFET circuits.
- As an application, incorporate a MOS transistor in a circuit design that enhances the simple diode electronic thermometer discussed in Chapter 1.



3.1

MOS FIELD-EFFECT TRANSISTOR

Objective: • Understand the operation and characteristics of the various types of metal-oxide semiconductor field-effect transistors (MOSFETs).

The **metal-oxide-semiconductor field-effect transistor (MOSFET)** became a practical reality in the 1970s. The MOSFET, compared to BJTs, can be made very small (that is, it occupies a very small area on an IC chip). Since digital circuits can be designed using only MOSFETs, with essentially no resistors or diodes required, high-density VLSI circuits, including microprocessors and memories, can be fabricated. The MOSFET has made possible the handheld calculator, the powerful personal computer, and the laptop computer. MOSFETs can also be used in analog circuits, as we will see in the next chapter.

In the MOSFET, the current is controlled by an electric field applied perpendicular to both the semiconductor surface and to the direction of current. The phenomenon used to modulate the conductance of a semiconductor, or control the current in a semiconductor, by applying an electric field perpendicular to the surface is called the **field effect**. The basic transistor principle is that the voltage between two terminals controls the current through the third terminal.

In the following two sections, we will discuss the various types of MOSFETs, develop the $i-v$ characteristics, and then consider the dc biasing of various MOSFET circuit configurations. After studying these sections, you should be familiar and comfortable with the MOSFET and MOSFET circuits.

3.1.1

Two-Terminal MOS Structure

The heart of the MOSFET is the metal-oxide-semiconductor capacitor shown in Figure 3.1. The metal may be aluminum or some other type of metal. In most cases, the metal is replaced by a high-conductivity polycrystalline silicon layer deposited on the oxide. However, the term metal is usually still used in referring to MOSFETs. In the figure, the parameter t_{ox} is the thickness of the oxide and ϵ_{ox} is the oxide permittivity.

The physics of the MOS structure can be explained with the aid of a simple parallel-plate capacitor.¹ Figure 3.2(a) shows a parallel-plate capacitor with the top plate at a negative voltage with respect to the bottom plate. An insulator material separates the two plates. With this bias, a negative charge exists on the top plate, a positive charge exists on the bottom plate, and an electric field is induced between the two plates, as shown.

A MOS capacitor with a p-type semiconductor substrate is shown in Figure 3.2(b). The top metal terminal, also called the **gate**, is at a negative voltage with respect to the semiconductor substrate. From the example of the parallel-plate capacitor, we can see that a negative charge will exist on the top metal plate and an electric field will be induced in the direction shown in the figure. If the electric field penetrates the

¹The capacitance of a parallel plate capacitor, neglecting fringing fields, is $C = \epsilon A/d$, where A is the area of one plate, d is the distance between plates, and ϵ is the permittivity of the medium between the plates.

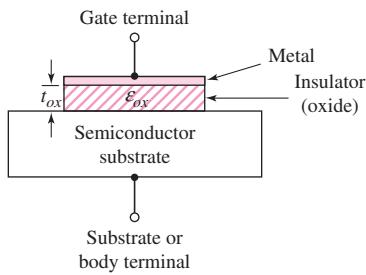


Figure 3.1 The basic MOS capacitor structure

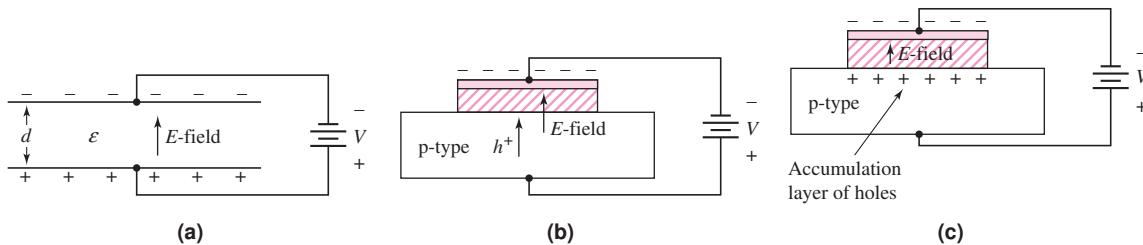


Figure 3.2 (a) A parallel-plate capacitor, showing the electric field and conductor charges, (b) a corresponding MOS capacitor with a negative gate bias, showing the electric field and charge flow, and (c) the MOS capacitor with an accumulation layer of holes

semiconductor, the holes in the p-type semiconductor will experience a force toward the oxide-semiconductor interface. The equilibrium distribution of charge in the MOS capacitor with this particular applied voltage is shown in Figure 3.2(c). An accumulation layer of positively charged holes at the oxide-semiconductor interface corresponds to the positive charge on the bottom “plate” of the MOS capacitor.

Figure 3.3(a) shows the same MOS capacitor, but with the polarity of the applied voltage reversed. A positive charge now exists on the top metal plate and the induced electric field is in the opposite direction, as shown. In this case, if the electric field penetrates the semiconductor, holes in the p-type material will experience a force away from the oxide-semiconductor interface. As the holes are pushed away from the interface, a negative space-charge region is created, because of the fixed acceptor impurity atoms. The negative charge in the induced depletion region corresponds to the negative charge on the bottom “plate” of the MOS capacitor. Figure 3.3(b) shows the equilibrium distribution of charge in the MOS capacitor with this applied voltage.

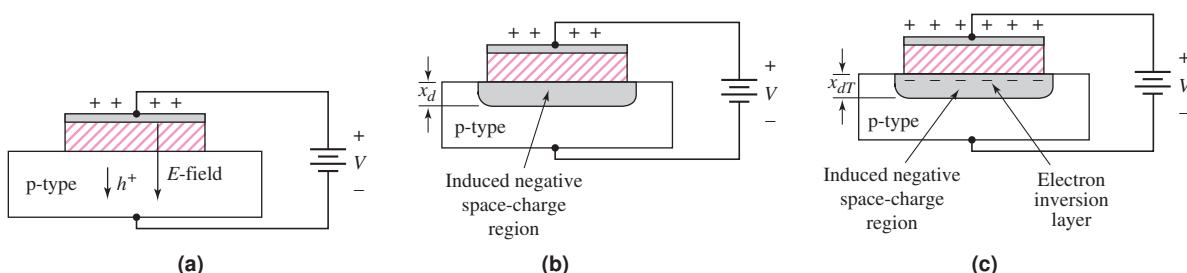


Figure 3.3 The MOS capacitor with p-type substrate: (a) effect of positive gate bias, showing the electric field and charge flow, (b) the MOS capacitor with an induced space-charge region due to a moderate positive gate bias, and (c) the MOS capacitor with an induced space-charge region and electron inversion layer due to a larger positive gate bias

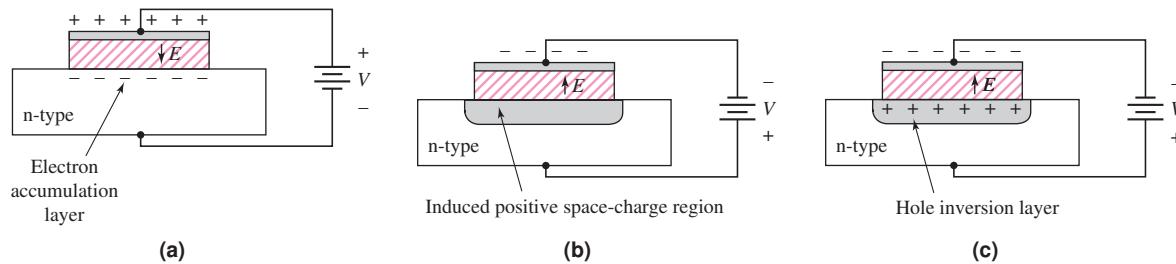


Figure 3.4 The MOS capacitor with n-type substrate: (a) effect of a positive gate bias and the formation of an electron accumulation layer, (b) the MOS capacitor with an induced space-charge region due to a moderate negative gate bias, and (c) the MOS capacitor with an induced space-charge region and hole inversion layer due to a larger negative gate bias

When a larger positive voltage is applied to the gate, the magnitude of the induced electric field increases. Minority carrier electrons are attracted to the oxide–semiconductor interface, as shown in Figure 3.3(c). This region of minority carrier electrons is called an **electron inversion layer**. The magnitude of the charge in the inversion layer is a function of the applied gate voltage.

The same basic charge distributions can be obtained in a MOS capacitor with an n-type semiconductor substrate. Figure 3.4(a) shows this MOS capacitor structure, with a positive voltage applied to the top gate terminal. A positive charge is created on the top gate and an electric field is induced in the direction shown. In this situation, an accumulation layer of electrons is induced in the n-type semiconductor.

Figure 3.4(b) shows the case when a negative voltage is applied to the gate terminal. A positive space-charge region is induced in the n-type substrate by the induced electric field. When a larger negative voltage is applied, a region of positive charge is created at the oxide–semiconductor interface, as shown in Figure 3.4(c). This region of minority carrier holes is called a **hole inversion layer**. The magnitude of the positive charge in the inversion layer is a function of the applied gate voltage.

The term **enhancement mode** means that a voltage must be applied to the gate to create an inversion layer. For the MOS capacitor with a p-type substrate, a positive gate voltage must be applied to create the electron inversion layer; for the MOS capacitor with an n-type substrate, a negative gate voltage must be applied to create the hole inversion layer.

3.1.2

n-Channel Enhancement-Mode MOSFET

We will now apply the concepts of an inversion layer charge in a MOS capacitor to create a transistor.

Transistor Structure

Figure 3.5(a) shows a simplified cross section of a MOS field-effect transistor. The gate, oxide, and p-type substrate regions are the same as those of a MOS capacitor. In addition, we now have two n-regions, called the **source terminal** and **drain terminal**. The current in a MOSFET is the result of the flow of charge in the inversion layer, also called the **channel region**, adjacent to the oxide–semiconductor interface.

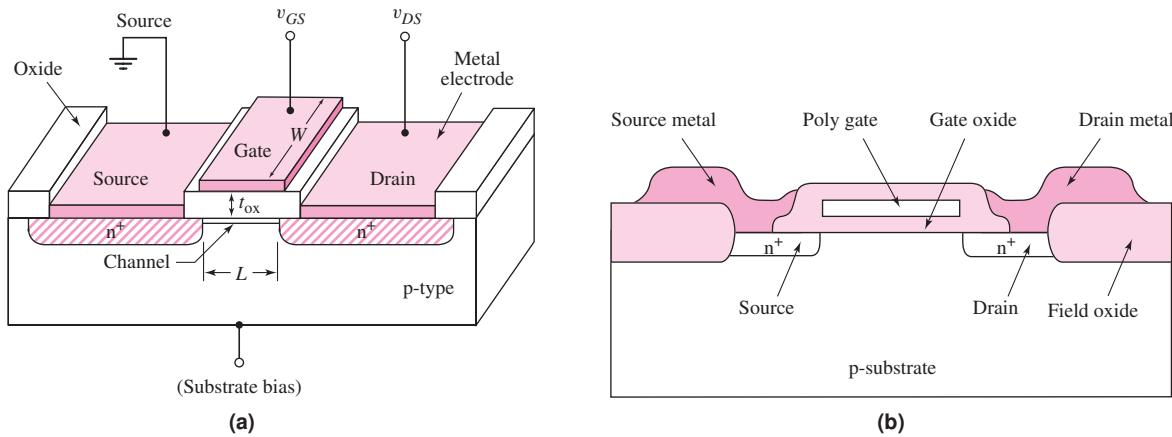


Figure 3.5 (a) Schematic diagram of an n-channel enhancement-mode MOSFET and (b) an n-channel MOSFET, showing the field oxide and polysilicon gate

The channel length L and channel width W are defined on the figure. The channel length of a typical integrated circuit MOSFET is less than $1 \mu\text{m}$ (10^{-6} m), which means that MOSFETs are small devices. The oxide thickness t_{ox} is typically on the order of 400 angstroms, or less.

The diagram in Figure 3.5(a) is a simplified sketch of the basic structure of the transistor. Figure 3.5(b) shows a more detailed cross section of a MOSFET fabricated into an integrated circuit configuration. A thick oxide, called the **field oxide**, is deposited outside the area in which the metal interconnect lines are formed. The gate material is usually heavily doped polysilicon. Even though the actual structure of a MOSFET may be fairly complex, the simplified diagram may be used to develop the basic transistor characteristics.

Basic Transistor Operation

With zero bias applied to the gate, the source and drain terminals are separated by the p-region, as shown in Figure 3.6(a). This is equivalent to two back-to-back diodes, as shown in Figure 3.6(b). The current in this case is essentially zero. If a large enough positive gate voltage is applied, an electron inversion layer is created at the oxide–semiconductor interface and this layer “connects” the n-source to the n-drain,

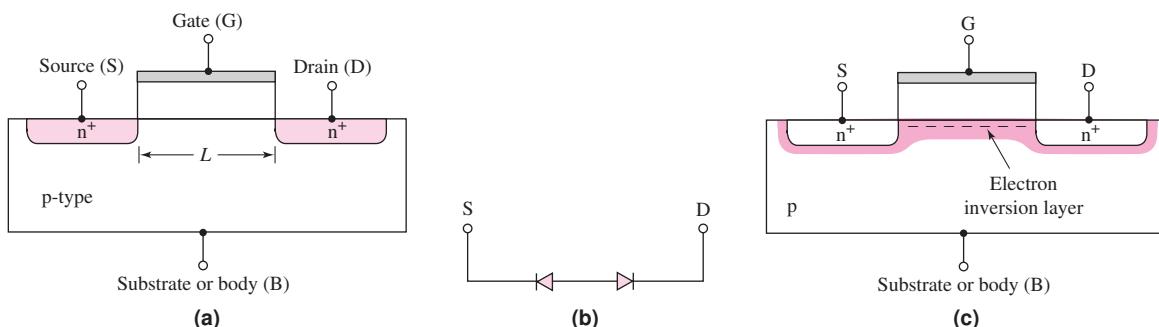


Figure 3.6 (a) Cross section of the n-channel MOSFET prior to the formation of an electron inversion layer, (b) equivalent back-to-back diodes between source and drain when the transistor is in cutoff, and (c) cross section after the formation of an electron inversion layer

as shown in Figure 3.6(c). A current can then be generated between the source and drain terminals. Since a voltage must be applied to the gate to create the inversion charge, this transistor is called an **enhancement-mode MOSFET**. Also, since the carriers in the inversion layer are electrons, this device is also called an **n-channel MOSFET (NMOS)**.

The source terminal supplies carriers that flow through the channel, and the drain terminal allows the carriers to drain from the channel. For the n-channel MOSFET, electrons flow from the source to the drain with an applied drain-to-source voltage, which means the conventional current enters the drain and leaves the source. The magnitude of the current is a function of the amount of charge in the inversion layer, which in turn is a function of the applied gate voltage. Since the gate terminal is separated from the channel by an oxide or insulator, there is no gate current. Similarly, since the channel and substrate are separated by a space-charge region, there is essentially no current through the substrate.

3.1.3 Ideal MOSFET Current-Voltage Characteristics—NMOS Device

The **threshold voltage** of the n-channel MOSFET, denoted as V_{TN} , is defined² as the applied gate voltage needed to create an inversion charge in which the density is equal to the concentration of majority carriers in the semiconductor substrate. In simple terms, we can think of the threshold voltage as the gate voltage required to “turn on” the transistor.

For the n-channel enhancement-mode MOSFET, the threshold voltage is positive because a positive gate voltage is required to create the inversion charge. If the gate voltage is less than the threshold voltage, the current in the device is essentially zero. If the gate voltage is greater than the threshold voltage, a drain-to-source current is generated as the drain-to-source voltage is applied. The gate and drain voltages are measured with respect to the source.

Figure 3.7(a) shows an n-channel enhancement-mode MOSFET with the source and substrate terminals connected to ground. The gate-to-source voltage is less than the threshold voltage, and there is a small drain-to-source voltage. With this bias configuration, there is no electron inversion layer, the drain-to-substrate pn junction is reverse biased, and the drain current is zero (neglecting pn junction leakage currents).

Figure 3.7(b) shows the same MOSFET with an applied gate voltage greater than the threshold voltage. In this situation, an electron inversion layer is created and, when a small drain voltage is applied, electrons in the inversion layer flow from the source to the positive drain terminal. The conventional current enters the drain terminal and leaves the source terminal. Note that a positive drain voltage creates a reverse-biased drain-to-substrate pn junction, so current flows through the channel region and not through a pn junction.

The i_D versus v_{DS} characteristics³ for small values of v_{DS} are shown in Figure 3.8. When $v_{GS} < V_{TN}$, the drain current is zero. When v_{GS} is greater than V_{TN} ,

²The usual notation for threshold voltage is V_T . However, since we have defined the thermal voltage as $V_T = kT/q$, we will use V_{TN} for the threshold voltage of the n-channel device.

³The voltage notation v_{DS} and v_{GS} , with the dual subscript, denotes the voltage between the drain (D) and source (S) and between the gate (G) and source (S), respectively. Implicit in the notation is that the first subscript is positive with respect to the second subscript.

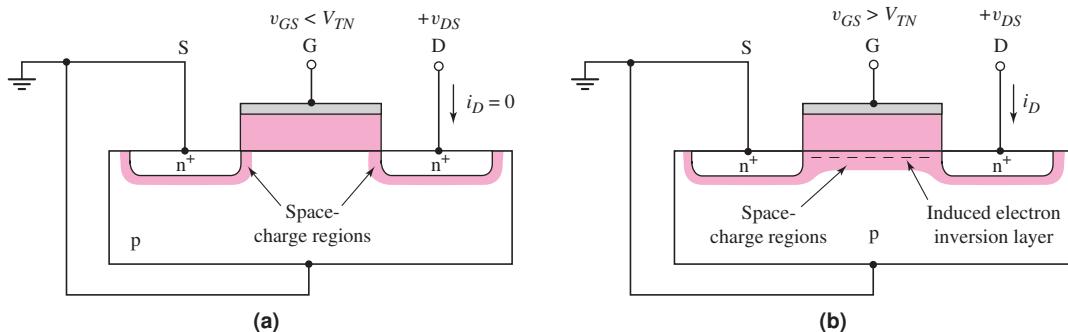


Figure 3.7 The n-channel enhancement-mode MOSFET (a) with an applied gate voltage $v_{GS} < V_{TN}$, and (b) with an applied gate voltage $v_{GS} > V_{TN}$

the channel inversion charge is formed and the drain current increases with v_{DS} . Then, with a larger gate voltage, a larger inversion charge density is created, and the drain current is greater for a given value of v_{DS} .

Figure 3.9(a) shows the basic MOS structure for $v_{GS} > V_{TN}$ and a small applied v_{DS} . In the figure, the thickness of the inversion channel layer qualitatively indicates the relative charge density, which for this case is essentially constant along the entire channel length. The corresponding i_D versus v_{DS} curve is also shown in the figure.

Figure 3.9(b) shows the situation when v_{DS} increases. As the drain voltage increases, the voltage drop across the oxide near the drain terminal decreases, which means that the induced inversion charge density near the drain also decreases. The incremental conductance of the channel at the drain then decreases, which causes the slope of the i_D versus v_{DS} curve to decrease. This effect is shown in the i_D versus v_{DS} curve in the figure.

As v_{DS} increases to the point where the potential difference, $v_{GS} - v_{DS}$, across the oxide at the drain terminal is equal to V_{TN} , the induced inversion charge density at the drain terminal is zero. This effect is shown schematically in Figure 3.9(c). For this condition, the incremental channel conductance at the drain is zero, which means that the slope of the i_D versus v_{DS} curve is zero. We can write

$$v_{GS} - v_{DS(\text{sat})} = V_{TN} \quad (3.1\text{(a)})$$

or

$$v_{DS(\text{sat})} = v_{GS} - V_{TN} \quad (3.1\text{(b)})$$

where $v_{DS(\text{sat})}$ is the drain-to-source voltage that produces zero inversion charge density at the drain terminal.

When v_{DS} becomes larger than $v_{DS(\text{sat})}$, the point in the channel at which the inversion charge is just zero moves toward the source terminal. In this case, electrons enter the channel at the source, travel through the channel toward the drain, and then, at the point where the charge goes to zero, are injected into the space-charge region, where they are swept by the E -field to the drain contact. In the ideal MOSFET, the drain current is constant for $v_{DS} > v_{DS(\text{sat})}$. This region of the i_D versus v_{DS} characteristic is referred to as the **saturation region**, which is shown in Figure 3.9(d).

As the applied gate-to-source voltage changes, the i_D versus v_{DS} curve changes. In Figure 3.8, we saw that the initial slope of i_D versus v_{DS} increases as v_{GS} increases. Also, Equation (3.1(b)) shows that $v_{DS(\text{sat})}$ is a function of v_{GS} . Therefore, we can

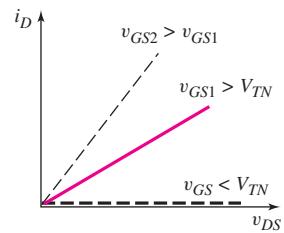


Figure 3.8 Plot of i_D versus v_{DS} characteristic for small values of v_{DS} at three v_{GS} voltages

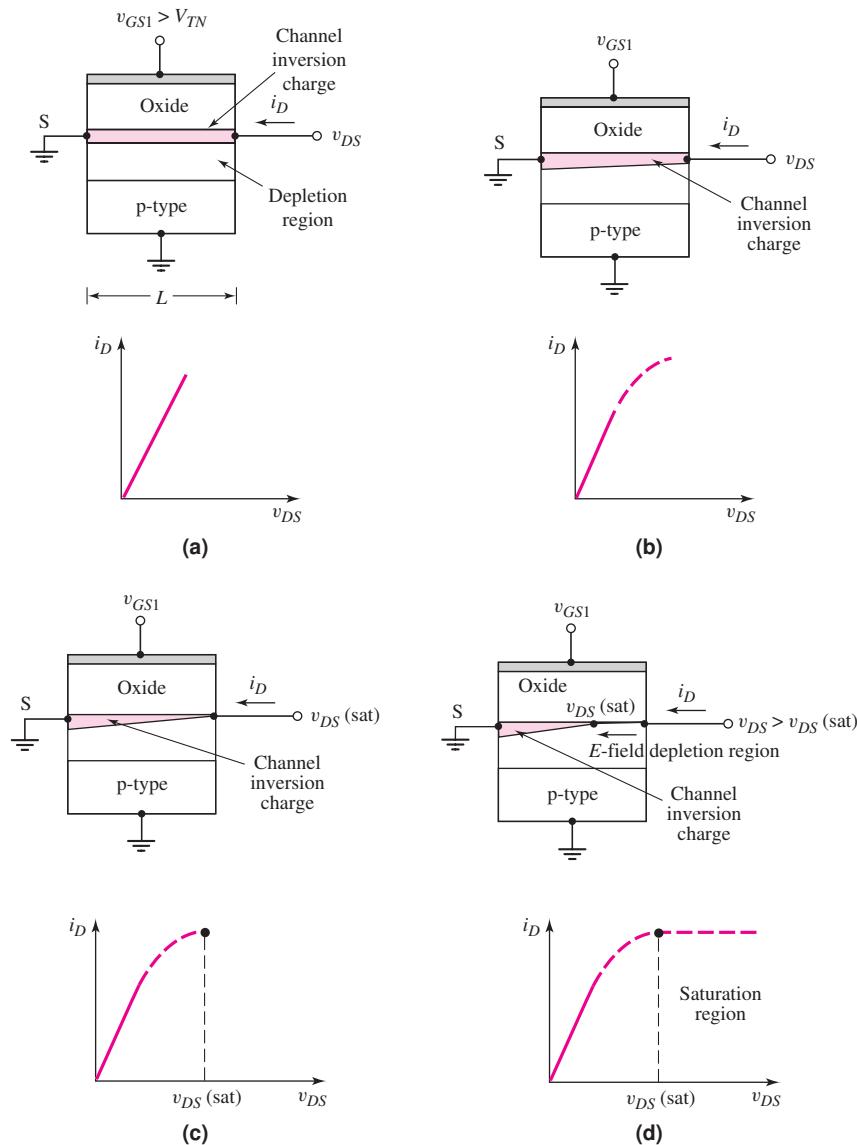


Figure 3.9 Cross section and i_D versus v_{DS} curve for an n-channel enhancement-mode MOSFET when $v_{GS} > V_{TN}$ for (a) a small v_{DS} value, (b) a larger v_{DS} value but for $v_{DS} < v_{DS}(\text{sat})$, (c) $v_{DS} = v_{DS}(\text{sat})$, and (d) $v_{DS} > v_{DS}(\text{sat})$

generate the family of curves for this n-channel enhancement mode MOSFET as shown in Figure 3.10.

Although the derivation of the current–voltage characteristics of the MOSFET is beyond the scope of this text, we can define the relationships. The region for which $v_{DS} < v_{DS}(\text{sat})$ is known as the **nonsaturation** or **triode region**. The ideal current–voltage characteristics in this region are described by the equation

$$i_D = K_n [2(v_{GS} - V_{TN})v_{DS} - v_{DS}^2] \quad (3.2(a))$$

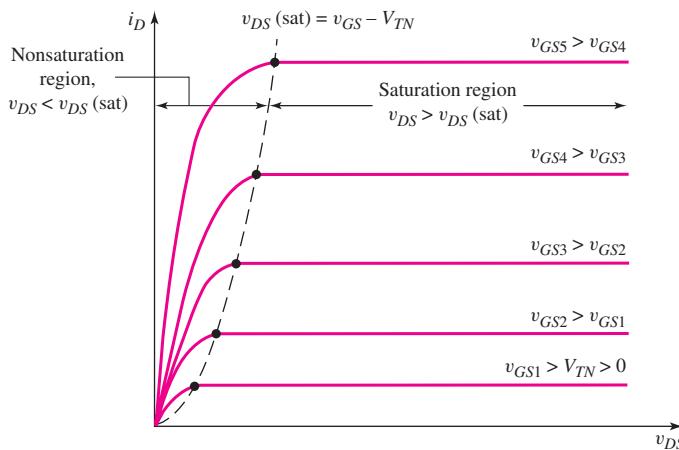


Figure 3.10 Family of i_D versus v_{DS} curves for an n-channel enhancement-mode MOSFET.

Note that the $v_{DS}(\text{sat})$ voltage is a single point on each of the curves. This point denotes the transition between the nonsaturation region and the saturation region

In the saturation region, the ideal current–voltage characteristics for $v_{GS} > V_{TN}$ are described by the equation

$$i_D = K_n(v_{GS} - V_{TN})^2 \quad (3.2\text{(b)})$$

In the saturation region, since the ideal drain current is independent of the drain-to-source voltage, the incremental or small-signal resistance is infinite. We see that

$$r_0 = \Delta v_{DS}/\Delta i_D|_{v_{GS}=\text{const.}} = \infty$$

The parameter K_n is sometimes called the transconductance parameter for the n-channel device. However, this term is not to be confused with the small-signal transconductance parameter introduced in the next chapter. For simplicity, we will refer to this parameter as the **conduction parameter**, which for an n-channel device is given by

$$K_n = \frac{W\mu_n C_{\text{ox}}}{2L} \quad (3.3\text{(a)})$$

where C_{ox} is the oxide capacitance per unit area. The capacitance is given by

$$C_{\text{ox}} = \epsilon_{\text{ox}}/t_{\text{ox}}$$

where t_{ox} is the oxide thickness and ϵ_{ox} is the oxide permittivity. For silicon devices, $\epsilon_{\text{ox}} = (3.9)(8.85 \times 10^{-14}) \text{ F/cm}$. The parameter μ_n is the mobility of the electrons in the inversion layer. The channel width W and channel length L were shown in Figure 3.5(a).

As Equation (3.3(a)) indicates, the conduction parameter is a function of both electrical and geometric parameters. The oxide capacitance and carrier mobility are essentially constants for a given fabrication technology. However, the geometry, or width-to-length ratio W/L , is a variable in the design of MOSFETs that is used to produce specific current–voltage characteristics in MOSFET circuits.

We can rewrite the conduction parameter in the form

$$K_n = \frac{k'_n}{2} \cdot \frac{W}{L} \quad (3.3\text{(b)})$$

where $k'_n = \mu_n C_{\text{ox}}$ and is called the **process conduction parameter**. Normally, k'_n is considered to be a constant for a given fabrication technology, so Equation (3.3(b)) indicates that the width-to-length ratio W/L is the transistor design variable.

EXAMPLE 3.1

Objective: Calculate the current in an n-channel MOSFET.

Consider an n-channel enhancement-mode MOSFET with the following parameters: $V_{TN} = 0.4$ V, $W = 20$ μm , $L = 0.8$ μm , $\mu_n = 650$ $\text{cm}^2/\text{V}\cdot\text{s}$, $t_{\text{ox}} = 200$ Å, and $\epsilon_{\text{ox}} = (3.9)(8.85 \times 10^{-14})$ F/cm. Determine the current when the transistor is biased in the saturation region for (a) $v_{GS} = 0.8$ V and (b) $v_{GS} = 1.6$ V.

Solution: The conduction parameter is determined by Equation (3.3(a)). First, consider the units involved in this equation, as follows:

$$K_n = \frac{W(\text{cm}) \cdot \mu_n \left(\frac{\text{cm}^2}{\text{V}\cdot\text{s}} \right) \epsilon_{\text{ox}} \left(\frac{\text{F}}{\text{cm}} \right)}{2L(\text{cm}) \cdot t_{\text{ox}}(\text{cm})} = \frac{\text{F}}{\text{V}\cdot\text{s}} = \frac{(\text{C}/\text{V})}{\text{V}\cdot\text{s}} = \frac{\text{A}}{\text{V}^2}$$

The value of the conduction parameter is therefore

$$K_n = \frac{W\mu_n\epsilon_{\text{ox}}}{2Lt_{\text{ox}}} = \frac{(20 \times 10^{-4})(650)(3.9)(8.85 \times 10^{-14})}{2(0.8 \times 10^{-4})(200 \times 10^{-8})}$$

or

$$K_n = 1.40 \text{ mA/V}^2$$

From Equation (3.2(b)), we find:

(a) For $v_{GS} = 0.8$ V,

$$i_D = K_n(v_{GS} - V_{TN})^2 = (1.40)(0.8 - 0.4)^2 = 0.224 \text{ mA}$$

(b) For $v_{GS} = 1.6$ V,

$$i_D = (1.40)(1.6 - 0.4)^2 = 2.02 \text{ mA}$$

Comment: The current capability of a transistor can be increased by increasing the conduction parameter. For a given fabrication technology, K_n is adjusted by varying the transistor width W .

EXERCISE PROBLEM

Ex 3.1: An NMOS transistor with $V_{TN} = 1$ V has a drain current $i_D = 0.8$ mA when $v_{GS} = 3$ V and $v_{DS} = 4.5$ V. Calculate the drain current when: (a) $v_{GS} = 2$ V, $v_{DS} = 4.5$ V; and (b) $v_{GS} = 3$ V, $v_{DS} = 1$ V. (Ans. (a) 0.2 mA (b) 0.6 mA)

3.1.4 p-Channel Enhancement-Mode MOSFET

The complementary device of the n-channel enhancement-mode MOSFET is the p-channel enhancement-mode MOSFET.

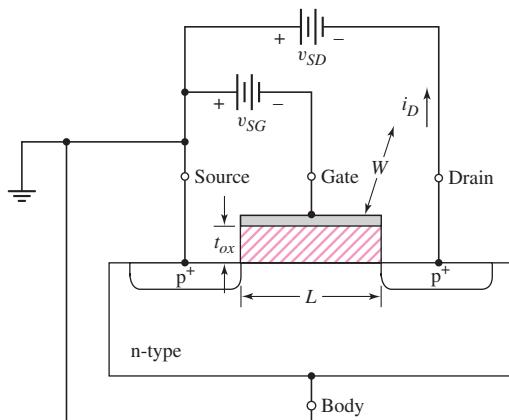


Figure 3.11 Cross section of p-channel enhancement-mode MOSFET. The device is cut off for $v_{SG} = 0$. The dimension W extends into the plane of the page.

Transistor Structure

Figure 3.11 shows a simplified cross section of the p-channel enhancement-mode transistor. The substrate is now n-type and the source and drain areas are p-type. The channel length, channel width, and oxide thickness parameter definitions are the same as those for the NMOS device shown in Figure 3.5(a).

Basic Transistor Operation

The operation of the p-channel device is the same as that of the n-channel device, except the hole is the charge carrier rather than the electron. A negative gate bias is required to induce an inversion layer of holes in the channel region directly under the oxide. The threshold voltage for the p-channel device is denoted as V_{TP} .⁴ Since the threshold voltage is defined as the gate voltage required to induce the inversion layer, then $V_{TP} < 0$ for the p-channel enhancement-mode device.

Once the inversion layer has been created, the p-type source region is the source of the charge carrier so that holes flow from the source to the drain. A negative drain voltage is therefore required to induce an electric field in the channel forcing the holes to move from the source to the drain. The conventional current direction, then, for the PMOS transistor is into the source and out of the drain. The conventional current direction and voltage polarity for the PMOS device are reversed compared to the NMOS device.

Note in Figure 3.11 the reversal of the voltage subscripts. For $v_{SG} > 0$, the gate voltage is negative with respect to that at the source. Similarly, for $v_{SD} > 0$, the drain voltage is negative with respect to that at the source.

3.1.5 Ideal MOSFET Current–Voltage Characteristics—PMOS Device

The ideal current–voltage characteristics of the p-channel enhancement-mode device are essentially the same as those shown in Figure 3.10, noting that the drain current

⁴Using a different threshold voltage parameter for a PMOS device compared to the NMOS device is for clarity only.

is out of the drain and v_{DS} is replaced by v_{SD} . The saturation point is given by $v_{SD}(\text{sat}) = v_{SG} + V_{TP}$. For the p-channel device biased in the nonsaturation region, the current is given by

$$i_D = K_p [2(v_{SG} + V_{TP})v_{SD} - v_{SD}^2] \quad (3.4(a))$$

In the saturation region, the current is given by

$$i_D = K_p (v_{SG} + V_{TP})^2 \quad (3.4(b))$$

and the drain current exits the drain terminal. The parameter K_p is the conduction parameter for the p-channel device and is given by

$$K_p = \frac{W\mu_p C_{\text{ox}}}{2L} \quad (3.5(a))$$

where W , L , and C_{ox} are the channel width, length, and oxide capacitance per unit area, as previously defined. The parameter μ_p is the mobility of holes in the hole inversion layer. In general, the hole inversion layer mobility is less than the electron inversion layer mobility.

We can also rewrite Equation (3.5(a)) in the form

$$K_p = \frac{k'_p}{2} \cdot \frac{W}{L} \quad (3.5(b))$$

where $k'_p = \mu_p C_{\text{ox}}$.

For a p-channel MOSFET biased in the saturation region, we have

$$v_{SD} > v_{SD}(\text{sat}) = v_{SG} + V_{TP} \quad (3.6)$$

EXAMPLE 3.2

Objective: Determine the source-to-drain voltage required to bias a p-channel enhancement-mode MOSFET in the saturation region.

Consider an enhancement-mode p-channel MOSFET for which $K_p = 0.2 \text{ mA/V}^2$, $V_{TP} = -0.50 \text{ V}$, and $i_D = 0.50 \text{ mA}$.

Solution: In the saturation region, the drain current is given by

$$i_D = K_p (v_{SG} + V_{TP})^2$$

or

$$0.50 = 0.2(v_{SG} - 0.50)^2$$

which yields

$$v_{SG} = 2.08 \text{ V}$$

To bias this p-channel MOSFET in the saturation region, the following must apply:

$$v_{SD} > v_{SD}(\text{sat}) = v_{SG} + V_{TP} = 2.08 - 0.5 = 1.58 \text{ V}$$

Comment: Biasing a transistor in either the saturation or the nonsaturation region depends on both the gate-to-source voltage and the drain-to-source voltage.

EXERCISE PROBLEM

Ex 3.2: A PMOS device with $V_{TP} = -1.2$ V has a drain current $i_D = 0.5$ mA when $v_{SG} = 3$ V and $v_{SD} = 5$ V. Calculate the drain current when (a) $v_{SG} = 2$ V, $v_{SD} = 3$ V; and (b) $v_{SG} = 5$ V, $v_{SD} = 2$ V. (Ans. (a) 0.0986 mA, (b) 1.72 mA)

3.1.6 Circuit Symbols and Conventions

The conventional circuit symbol for the n-channel enhancement-mode MOSFET is shown in Figure 3.12(a). The vertical solid line denotes the gate electrode, the vertical broken line denotes the channel (the broken line indicates the device is enhancement mode), and the separation between the gate line and channel line denotes the oxide that insulates the gate from the channel. The polarity of the pn junction between the substrate and the channel is indicated by the arrowhead on the body or substrate terminal. The direction of the arrowhead indicates the type of transistor, which in this case is an n-channel device. This symbol shows the four-terminal structure of the MOSFET device.

In most applications in this text, we will implicitly assume that the source and substrate terminals are connected together. Explicitly drawing the substrate terminal for each transistor in a circuit becomes redundant and makes the circuits appear more complex. Instead, we will use the circuit symbol for the n-channel MOSFET shown in Figure 3.12(b). In this symbol, the arrowhead is on the source terminal and it indicates the direction of current, which for the n-channel device is out of the source. By including the arrowhead in the symbol, we do not need to explicitly indicate the source and drain terminals. We will use this circuit symbol throughout the text except in specific applications.

In more advanced texts and journal articles, the circuit symbol of the n-channel MOSFET shown in Figure 3.12(c) is generally used. The gate terminal is obvious and it is implicitly understood that the “top” terminal is the drain and the “bottom” terminal is the source. The top terminal, in this case the drain, is usually at a more positive voltage than the bottom terminal. In this introductory text, we will use the symbol shown in Figure 3.12(b) for clarity.

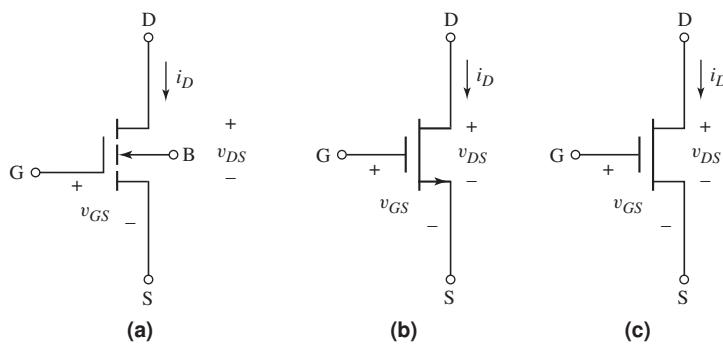


Figure 3.12 The n-channel enhancement-mode MOSFET: (a) conventional circuit symbol, (b) circuit symbol that will be used in this text, and (c) a simplified circuit symbol used in more advanced texts

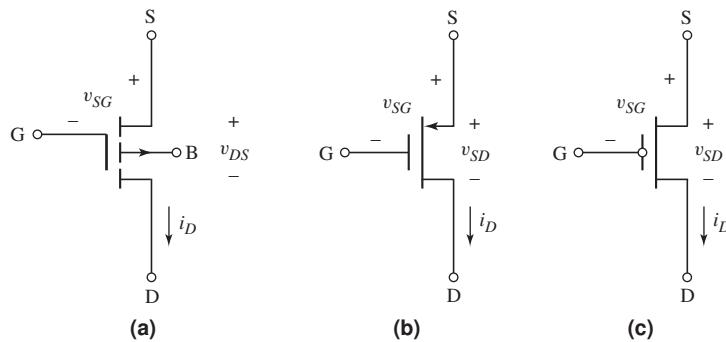


Figure 3.13 The p-channel enhancement-mode MOSFET: (a) conventional circuit symbol, (b) circuit symbol that will be used in this text, and (c) a simplified circuit symbol used in more advanced texts

The conventional circuit symbol for the p-channel enhancement-mode MOSFET appears in Figure 3.13(a). Note that the arrowhead direction on the substrate terminal is reversed from that in the n-channel enhancement-mode device. This circuit symbol again shows the four terminal structure of the MOSFET device.

The circuit symbol for the p-channel enhancement-mode device shown in Figure 3.13(b) will be used in this text. The arrowhead is on the source terminal indicating the direction of the current, which for the p-channel device is into the source terminal.

In more advanced texts and journal articles, the circuit symbol of the p-channel MOSFET shown in Figure 3.13(c) is generally used. Again, the gate terminal is obvious but includes the O symbol to indicate that this is a PMOS device. It is implicitly understood that the “top” terminal is the source and the “bottom” terminal is the drain. The top terminal, in this case the source, is normally at a higher potential than the bottom terminal. Again, in this text, we will use the symbol shown in Figure 3.13(b) for clarity.

3.1.7 Additional MOSFET Structures and Circuit Symbols

Before we start analyzing MOSFET circuits, there are two other MOSFET structures in addition to the n-channel enhancement-mode device and the p-channel enhancement-mode device that need to be considered.

n-Channel Depletion-Mode MOSFET

Figure 3.14(a) shows the cross section of an n-channel **depletion-mode** MOSFET. When zero volts are applied to the gate, an n-channel region or inversion layer exists under the oxide as a result, for example, of impurities introduced during device fabrication. Since an n-region connects the n-source and n-drain, a drain-to-source current may be generated in the channel even with zero gate voltage. The term **depletion mode** means that a channel exists even at zero gate voltage. A negative gate voltage must be applied to the n-channel depletion-mode MOSFET to turn the device off.

Figure 3.14(b) shows the n-channel depletion mode MOSFET with a negative applied gate-to-source voltage. A negative gate voltage induces a space-charge region under the oxide, thereby reducing the thickness of the n-channel region. The reduced thickness decreases the channel conductance, which in turn reduces the drain current. When the gate voltage is equal to the threshold voltage, which is

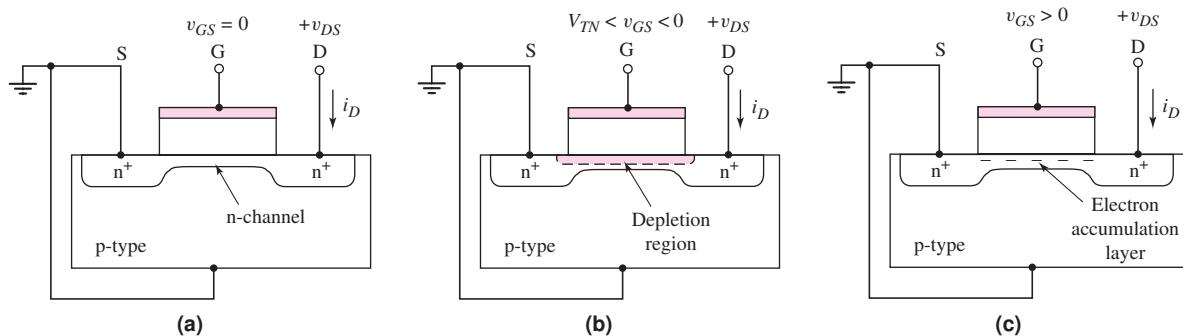


Figure 3.14 Cross section of an n-channel depletion mode MOSFET for: (a) $v_{GS} = 0$, (b) $v_{GS} < 0$, and (c) $v_{GS} > 0$

negative for this device, the induced space-charge region extends completely through the n-channel region, and the current goes to zero. A positive gate voltage creates an electron accumulation layer, as shown in Figure 3.14(c) which increases the drain current. The general i_D versus v_{DS} family of curves for the n-channel depletion-mode MOSFET is shown in Figure 3.15.

The current–voltage characteristics defined by Equations (3.2(a)) and (3.2(b)) apply to both enhancement- and depletion-mode n-channel devices. The only difference is that the threshold voltage V_{TN} is positive for the enhancement-mode MOSFET and negative for the depletion-mode MOSFET. Even though the current–voltage characteristics of enhancement- and depletion-mode devices are described by the same equations, different circuit symbols are used, simply for purposes of clarity.

The conventional circuit symbol for the n-channel depletion-mode MOSFET is shown in Figure 3.16(a). The vertical solid line denoting the channel indicates the device is depletion mode. A comparison of Figures 3.12(a) and 3.16(a) shows that the only difference between the enhancement- and depletion-mode symbols is the broken versus the solid line representing the channel.

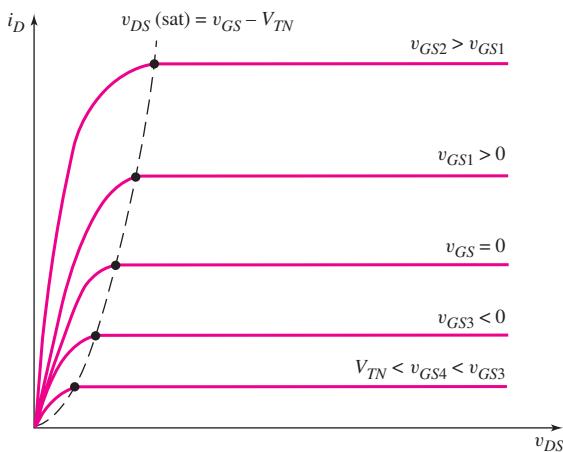


Figure 3.15 Family of i_D versus v_{DS} curves for an n-channel depletion-mode MOSFET. Note again that the $v_{DS(sat)}$ voltage is a single point on each curve.

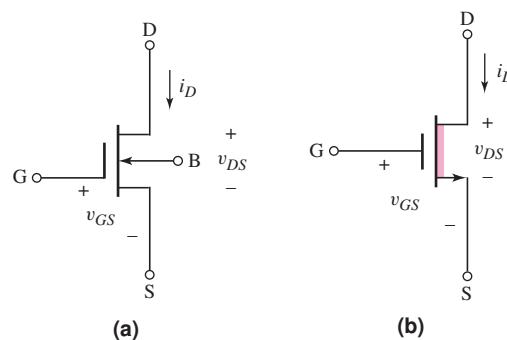


Figure 3.16 The n-channel depletion-mode MOSFET: (a) conventional circuit symbol and (b) simplified circuit symbol

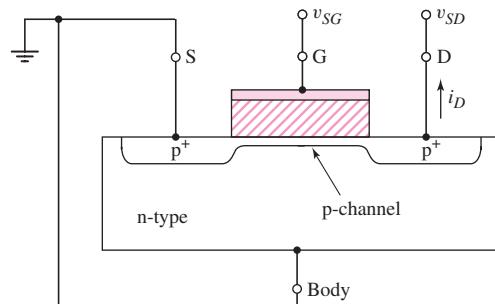


Figure 3.17 Cross section of p-channel depletion-mode MOSFET showing the p-channel under the oxide at zero gate voltage

A simplified symbol for the n-channel depletion-mode MOSFET is shown in Figure 3.16(b). The arrowhead is again on the source terminal and indicates the direction of current, which for the n-channel device is out of the source. The heavy solid line represents the depletion-mode channel region. Again, using a different circuit symbol for the depletion-mode device compared to the enhancement-mode device is simply for clarity in a circuit diagram.

p-Channel Depletion-Mode MOSFET

Figure 3.17 shows the cross section of a p-channel depletion-mode MOSFET, as well as the biasing configuration and current direction. In the depletion-mode device, a channel region of holes already exists under the oxide, even with zero gate voltage. A positive gate voltage is required to turn the device off. Hence the threshold voltage of a p-channel depletion-mode MOSFET is positive ($V_{TP} > 0$).

The conventional and simplified circuit symbols for the p-channel depletion-mode device are shown in Figure 3.18. The heavy solid line in the simplified symbol represents the channel region and denotes the depletion-mode device. The arrowhead is again on the source terminal and it indicates the current direction.

Complementary MOSFs

Complementary MOS (CMOS) technology uses both n-channel and p-channel devices in the same circuit. Figure 3.19 shows the cross section of n-channel and p-channel devices fabricated on the same chip. CMOS circuits, in general, are more complicated to fabricate than circuits using entirely NMOS or PMOS devices. Yet, as we will see in later chapters, CMOS circuits have great advantages over just NMOS or PMOS circuits.

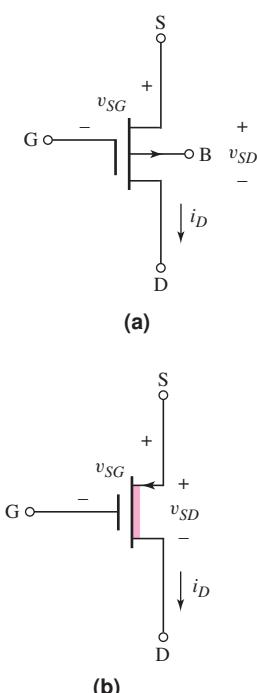


Figure 3.18 The p-channel depletion mode MOSFET:
(a) conventional circuit symbol and (b) simplified circuit symbol

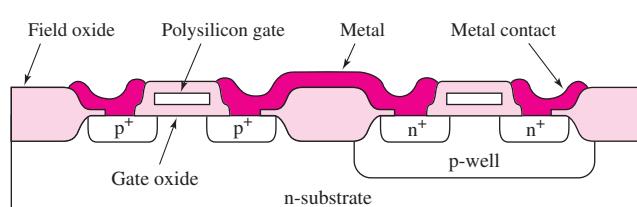


Figure 3.19 Cross sections of n-channel and p-channel transistors fabricated with a p-well CMOS technology

In order to fabricate n-channel and p-channel devices that are electrically equivalent, the magnitude of the threshold voltages must be equal, and the n-channel and p-channel conduction parameters must be equal. Since, in general, μ_n , and μ_p are not equal, the design of equivalent transistors involves adjusting the width-to-length ratios of the transistors.

3.1.8 Summary of Transistor Operation

We have presented a first-order model of the operation of the MOS transistor. For an n-channel enhancement-mode MOSFET, a positive gate-to-source voltage, greater than the threshold voltage V_{TN} , must be applied to induce an electron inversion layer. For $v_{GS} > V_{TN}$, the device is turned on. For an n-channel depletion-mode device, a channel between the source and drain exists even for $v_{GS} = 0$. The threshold voltage is negative, so that a negative value of v_{GS} is required to turn the device off.

For a p-channel device, all voltage polarities and current directions are reversed compared to the NMOS device. For the p-channel enhancement-mode transistor, $V_{TP} < 0$ and for the depletion-mode PMOS transistor, $V_{TP} > 0$.

Table 3.1 lists the first-order equations that describe the $i-v$ relationships in MOS devices. We note that K_n and K_p are positive values and that the drain current i_D is positive into the drain for the NMOS device and positive out of the drain for the PMOS device.

Table 3.1 Summary of the MOSFET current–voltage relationships

NMOS	PMOS
Nonsaturation region ($v_{DS} < v_{DS}(\text{sat})$)	Nonsaturation region ($v_{SD} < v_{SD}(\text{sat})$)
$i_D = K_n[2(v_{GS} - V_{TN})v_{DS} - v_{DS}^2]$	$i_D = K_p[2(v_{SG} + V_{TP})v_{SD} - v_{SD}^2]$
Saturation region ($v_{DS} > v_{DS}(\text{sat})$)	Saturation region ($v_{SD} > v_{SD}(\text{sat})$)
$i_D = K_n(v_{GS} - V_{TN})^2$	$i_D = K_p(v_{SG} + V_{TP})^2$
Transition point	Transition point
$v_{DS}(\text{sat}) = v_{GS} - V_{TN}$	$v_{SD}(\text{sat}) = v_{SG} + V_{TP}$
Enhancement mode	Enhancement mode
$V_{TN} > 0$	$V_{TP} < 0$
Depletion mode	Depletion mode
$V_{TN} < 0$	$V_{TP} > 0$

3.1.9 Short-Channel Effects

The current–voltage relations given by Equations (3.2(a)) and (3.2(b)) for the n-channel device and Equations (3.4(a)) and (3.4(b)) for the p-channel device are the ideal relations for long-channel devices. A long-channel device is generally one whose channel length is greater than $2 \mu\text{m}$. In this device, the horizontal electric field in the channel induced by the drain voltage and the vertical electric field induced by the gate voltage can be treated independently. However, the channel length of present-day devices is on the order of $0.2 \mu\text{m}$ or less.

There are several effects in these short-channel devices that influence and change the long-channel current–voltage characteristics. One such effect is a variation

in threshold voltage. The value of threshold voltage is a function of the channel length. This variation must be considered in the design and fabrication of these devices. The threshold voltage also becomes a function of the drain voltage. As the drain voltage increases, the effective threshold voltage decreases. This effect also influences the current–voltage characteristics.

The process conduction parameters, k'_n and k'_p , are directly related to the carrier mobility. We have assumed that the carrier mobilities and corresponding process conduction parameters are constant. However, the carrier mobility values are functions of the vertical electric field in the inversion layer. As the gate voltage and vertical electric field increase, the carrier mobility decreases. This result, again, directly influences the current–voltage characteristics of the device.

Another effect that occurs in short-channel devices is velocity saturation. As the horizontal electric field increases, the velocity of the carriers reaches a constant value and will no longer increase with an increase in drain voltage. Velocity saturation will lower the $V_{DS}(\text{sat})$ voltage value. The drain current will reach its saturation value at a smaller V_{DS} voltage. The drain current also becomes approximately a linear function of the gate voltage in the saturation region rather than the quadratic function of gate voltage in the long-channel characteristics.

Although the analysis of modern MOSFET circuits must take into account these short-channel effects, we will use the long-channel current–voltage relations in this introductory text. We will still be able to obtain a good basic understanding of the operation and characteristics of these devices, and we can still obtain a good basic understanding of the operation and characteristics of MOSFET circuits using the ideal long-channel current–voltage relations.

3.1.10 Additional Nonideal Current–Voltage Characteristics

The five nonideal effects in the current–voltage characteristics of MOS transistors are: the finite output resistance in the saturation region, the body effect, subthreshold conduction, breakdown effects, and temperature effects. This section will examine each of these effects.

Finite Output Resistance

In the ideal case, when a MOSFET is biased in the saturation region, the drain current i_D is independent of drain-to-source voltage v_{DS} . However, in actual MOSFET i_D versus v_{DS} characteristics, a nonzero slope does exist beyond the saturation point. For $v_{DS} > v_{DS}(\text{sat})$, the actual point in the channel at which the inversion charge goes to zero moves away from the drain terminal (see Figure 3.9(d)). The effective channel length decreases, producing the phenomenon called **channel length modulation**.

An exaggerated view of the current–voltage characteristics is shown in Figure 3.20. The curves can be extrapolated so that they intercept the voltage axis at a point $v_{DS} = -V_A$. The voltage V_A is usually defined as a positive quantity. The slope of the curve in the saturation region can be described by expressing the i_D versus v_{DS} characteristic in the form, for an n-channel device,

$$i_D = K_n[(v_{GS} - V_{TN})^2(1 + \lambda v_{DS})] \quad (3.7)$$

where λ is a positive quantity called the channel-length modulation parameter.

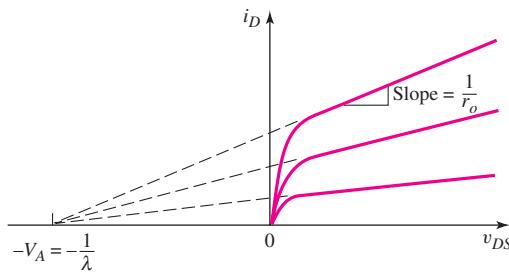


Figure 3.20 Family of i_D versus v_{DS} curves showing the effect of channel length modulation producing a finite output resistance

The parameters λ and V_A are related. From Equation (3.7), we have $(1 + \lambda v_{DS}) = 0$ at the extrapolated point where $i_D = 0$. At this point, $v_{DS} = -V_A$, which means that $V_A = 1/\lambda$.

The output resistance due to the channel length modulation is defined as

$$r_o = \left(\frac{\partial i_D}{\partial v_{DS}} \right)^{-1} \Big|_{v_{GS}=\text{const.}} \quad (3.8)$$

From Equation (3.7), the output resistance, evaluated at the Q -point, is

$$r_o = [\lambda K_n (V_{GSQ} - V_{TN})^2]^{-1} \quad (3.9(a))$$

or

$$r_o \approx [\lambda I_{DQ}]^{-1} = \frac{1}{\lambda I_{DQ}} = \frac{V_A}{I_{DQ}} \quad (3.9(b))$$

The output resistance r_o is also a factor in the small-signal equivalent circuit of the MOSFET, which is discussed in the next chapter.

Body Effect

Up to this point, we have assumed that the substrate, or body, is connected to the source. For this bias condition, the threshold voltage is a constant.

In integrated circuits, however, the substrates of all n-channel MOSFETs are usually common and are tied to the most negative potential in the circuit. An example of two n-channel MOSFETs in series is shown in Figure 3.21. The p-type

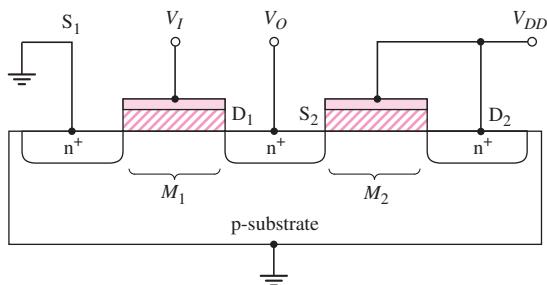


Figure 3.21 Two n-channel MOSFETs fabricated in series in the same substrate. The source terminal, S_2 , of the transistor M_2 is more than likely not at ground potential.

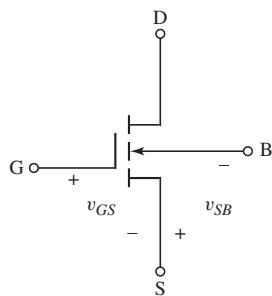


Figure 3.22 An n-channel enhancement-mode MOSFET with a substrate voltage

substrate is common to the two transistors, and the drain of M_1 is common to the source of M_2 . When the two transistors are conducting, there is a nonzero drain-to-source voltage on M_1 , which means that the source of M_2 is not at the same potential as the substrate. These bias conditions mean that a zero or reverse-bias voltage exists across the source–substrate pn junction, and a change in the source–substrate junction voltage changes the threshold voltage. This is called the **body effect**. The same situation exists in p-channel devices.

For example, consider the n-channel device shown in Figure 3.22. To maintain a zero- or reverse-biased source–substrate pn junction, we must have $v_{SB} \geq 0$. The threshold voltage for this condition is given by

$$V_{TN} = V_{TNO} + \gamma [\sqrt{2\phi_f + v_{SB}} - \sqrt{2\phi_f}] \quad (3.10)$$

where V_{TNO} is the threshold voltage for $v_{SB} = 0$; γ , called the bulk threshold or **body-effect parameter**, is related to device properties, and is typically on the order of $0.5 \text{ V}^{1/2}$; and ϕ_f is a semiconductor parameter, typically on the order of 0.35 V , and is a function of the semiconductor doping. We see from Equation (3.10) that the threshold voltage in n-channel devices increases due to this body effect.

The body effect can cause a degradation in circuit performance because of the changing threshold voltage. However, we will generally neglect the body effect in our circuit analyses, for simplicity.

Subthreshold Conduction

If we consider the ideal current-voltage relationship for the n-channel MOSFET biased in the saturation region, we have, from Equation (3.2(b)),

$$i_D = K_n(v_{GS} - V_{TN})^2$$

Taking the square root of both sides of the equation, we obtain

$$\sqrt{i_D} = \sqrt{K_n}(v_{GS} - V_{TN}) \quad (3.11)$$

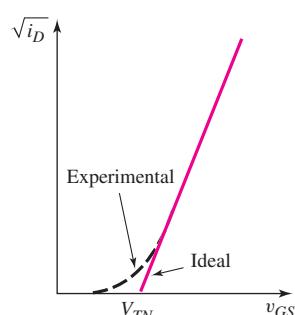


Figure 3.23 Plot of $\sqrt{i_D}$ versus v_{GS} for a MOSFET biased in the saturation region showing subthreshold conduction. Experimentally, a subthreshold current exists even for $v_{GS} < V_{TN}$.

From Equation (3.11), we see that $\sqrt{i_D}$ is a linear function of v_{GS} . Figure 3.23 shows a plot of this ideal relationship.

Also plotted in Figure 3.23 are experimental results, which show that when v_{GS} is slightly less than V_{TN} , the drain current is not zero, as previously assumed. This current is called the **subthreshold current**. The effect may not be significant for a single device, but if thousands or millions of devices on an integrated circuit are biased just slightly below the threshold voltage, the power supply current will not be zero but may contribute to significant power dissipation in the integrated circuit. One example of this is a dynamic random access memory (DRAM), as we will see in Chapter 16.

In this text, for simplicity we will not specifically consider the subthreshold current. However, when a MOSFET in a circuit is to be turned off, the “proper” design of the circuit must involve biasing the device at least a few tenths of a volt below the threshold voltage to achieve “true” cutoff.

Breakdown Effects

Several possible breakdown effects may occur in a MOSFET. The drain-to-substrate pn junction may break down if the applied drain voltage is too high and avalanche multiplication occurs. This breakdown is the same reverse-biased pn junction breakdown discussed in Chapter 1 in Section 1.2.5.

As the size of the device becomes smaller, another breakdown mechanism, called *punch-through*, may become significant. **Punch-through** occurs when the drain voltage is large enough for the depletion region around the drain to extend completely through the channel to the source terminal. This effect also causes the drain current to increase rapidly with only a small increase in drain voltage.

A third breakdown mechanism is called **near-avalanche** or **snapback breakdown**. This breakdown process is due to second-order effects within the MOSFET. The source-substrate-drain structure is equivalent to that of a bipolar transistor. As the device size shrinks, we may begin to see a parasitic bipolar transistor action with increases in the drain voltage. This parasitic action enhances the breakdown effect.

If the electric field in the oxide becomes large enough, breakdown can also occur in the oxide, which can lead to catastrophic failure. In silicon dioxide, the electric field at breakdown is on the order of 6×10^6 V/cm, which, to a first approximation, is given by $E_{ox} = V_G/t_{ox}$. A gate voltage of approximately 30 V would produce breakdown in an oxide with a thickness of $t_{ox} = 500 \text{ \AA}$. However, a safety margin of a factor of 3 is common, which means that the maximum safe gate voltage for $t_{ox} = 500 \text{ \AA}$ would be 10 V. A safety margin is necessary since there may be defects in the oxide that lower the breakdown field. We must also keep in mind that the input impedance at the gate is very high, and a small amount of static charge accumulating on the gate can cause the breakdown voltage to be exceeded. To prevent the accumulation of static charge on the gate capacitance of a MOSFET, a gate protection device, such as a reverse-biased diode, is usually included at the input of a MOS integrated circuit.

Temperature Effects

Both the threshold voltage V_{TN} and conduction parameter K_n are functions of temperature. The magnitude of the threshold voltage decreases with temperature, which means that the drain current increases with temperature at a given V_{GS} . However, the conduction parameter is a direct function of the inversion carrier mobility, which decreases as the temperature increases. Since the temperature dependence of mobility is larger than that of the threshold voltage, the net effect of increasing temperature is a decrease in drain current at a given V_{GS} . This particular result provides a negative feedback condition in power MOSFETs. A decreasing value of K_n inherently limits the channel current and provides stability for a power MOSFET.

Test Your Understanding

TYU 3.1 (a) An n-channel enhancement-mode MOSFET has a threshold voltage of $V_{TN} = 1.2$ V and an applied gate-to-source voltage of $v_{GS} = 2$ V. Determine the region of operation when: (i) $v_{DS} = 0.4$ V; (ii) $v_{DS} = 1$ V; and (iii) $v_{DS} = 5$ V. (b) Repeat part (a) for an n-channel depletion-mode MOSFET with a threshold voltage of $V_{TN} = -1.2$ V. (Ans. (a) (i) nonsaturation, (ii) saturation, (iii) saturation; (b) (i) nonsaturation, (ii) nonsaturation, (iii) saturation)

TYU 3.2 The NMOS devices described in Exercise TYU 3.1 have parameters $W = 20 \mu\text{m}$, $L = 0.8 \mu\text{m}$, $t_{ox} = 200 \text{ \AA}$, $\mu_n = 500 \text{ cm}^2/\text{V}\cdot\text{s}$, and $\lambda = 0$. (a) Calculate the conduction parameter K_n for each device. (b) Calculate the drain current for each bias condition listed in Exercise TYU 3.1. (Ans. (a) $K_n = 1.08 \text{ mA/V}^2$; (b) $i_D = 0.518 \text{ mA}, 0.691 \text{ mA}, \text{ and } 0.691 \text{ mA}; i_D = 2.59 \text{ mA}, 5.83 \text{ mA}, \text{ and } 11.1 \text{ mA}$)

TYU 3.3 (a) A p-channel enhancement-mode MOSFET has a threshold voltage of $V_{TP} = -1.2$ V and an applied source-to-gate voltage of $v_{SG} = 2$ V. Determine the region of operation when (i) $v_{SD} = 0.4$ V, (ii) $v_{SD} = 1$ V, and (iii) $v_{SD} = 5$ V. (b) Repeat part (a) for a p-channel depletion-mode MOSFET with a threshold voltage of $V_{TP} = +1.2$ V. (Ans. (a) (i) nonsaturation, (ii) saturation, (iii) saturation; (b) (i) nonsaturation, (ii) nonsaturation, (iii) saturation)

TYU 3.4 The PMOS devices described in Exercise TYU 3.3 have parameters $W = 10 \mu\text{m}$, $L = 0.8 \mu\text{m}$, $t_{ox} = 200 \text{ \AA}$, $\mu_p = 300 \text{ cm}^2/\text{V}\cdot\text{s}$, and $\lambda = 0$. (a) Calculate the conduction parameter K_p for each device. (b) Calculate the drain current for each bias condition listed in Exercise TYU 3.3. (Ans. (a) $K_p = 0.324 \text{ mA/V}^2$; (b) $i_D = 0.156 \text{ mA}$, 0.207 mA ; and 0.207 mA ; $i_D = 0.778 \text{ mA}$, 1.75 mA , and 3.32 mA)

TYU 3.5 The parameters of an NMOS enhancement-mode device are $V_{TN} = 0.25$ V and $K_n = 10 \mu\text{A/V}^2$. The device is biased at $v_{GS} = 0.5$ V. Calculate the drain current when (i) $v_{DS} = 0.5$ V and (ii) $v_{DS} = 1.2$ V for (a) $\lambda = 0$ and (b) $\lambda = 0.03 \text{ V}^{-1}$. (c) Calculate the output resistance r_o for parts (a) and (b). (Ans. (a) (i) and (ii) $i_D = 0.625 \mu\text{A}$; (b) (i) $i_D = 0.6344 \mu\text{A}$, (ii) $i_D = 0.6475 \mu\text{A}$; (c) (i) $r_o = \infty$, (ii) $r_o = 53.3 \text{ M}\Omega$).

TYU 3.6 An NMOS transistor has parameters $V_{TNO} = 0.4$ V, $\gamma = 0.15 \text{ V}^{1/2}$, and $\phi_f = 0.35$ V. Calculate the threshold voltage when (a) $v_{SB} = 0$, (b) $v_{SB} = 0.5$ V, and (c) $v_{SB} = 1.5$ V. (Ans. (a) 0.4 V, (b) 0.439 V, (c) 0.497 V)



3.2 MOSFET DC CIRCUIT ANALYSIS

Objective: • Understand and become familiar with the dc analysis and design techniques of MOSFET circuits.

In the last section, we considered the basic MOSFET characteristics and properties. We now start analyzing and designing the dc biasing of MOS transistor circuits. A primary purpose of the rest of the chapter is to continue to become familiar and comfortable with the MOS transistor and MOSFET circuits. The dc biasing of MOSFETs, the focus of this chapter, is an important part of the design of amplifiers. MOSFET amplifier design is the focus of the next chapter.

In most of the circuits presented in this chapter, resistors are used in conjunction with the MOS transistors. In a real MOSFET integrated circuit, however, the resistors are generally replaced by other MOSFETs, so the circuit is composed entirely of MOS devices. In general, a MOSFET device requires a smaller area than a resistor. As we go through the chapter, we will begin to see how this is accomplished and as we finish the text, we will indeed analyze and design circuits containing only MOSFETs.

In the dc analysis of MOSFET circuits, we can use the ideal current–voltage equations listed in Table 3.1 in Section 3.1.

3.2.1

Common-Source Circuit

One of the basic MOSFET circuit configurations is called the **common-source circuit**. Figure 3.24 shows one example of this type of circuit using an n-channel

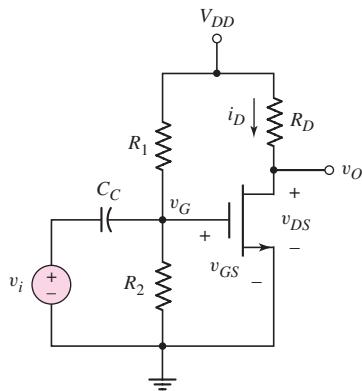


Figure 3.24 An NMOS common-source circuit

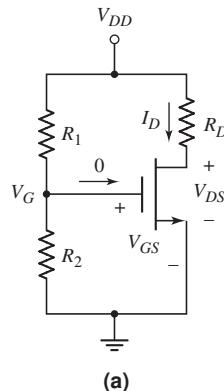
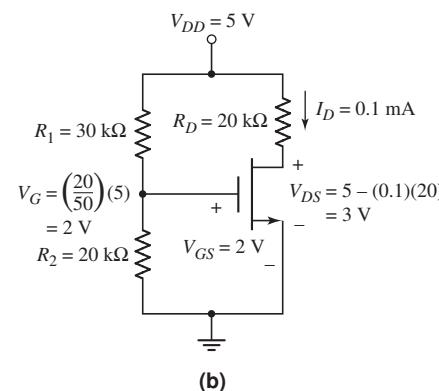


Figure 3.25 (a) The dc equivalent circuit of the NMOS common-source circuit and (b) the NMOS circuit for Example 3.3, showing current and voltage values



enhancement-mode MOSFET. The source terminal is at ground potential and is common to both the input and output portions of the circuit. The coupling capacitor C_C acts as an open circuit to dc but it allows the signal voltage to be coupled to the gate of the MOSFET.

The dc equivalent circuit is shown in Figure 3.25(a). In the following dc analyses, we again use the notation for dc currents and voltages. Since the gate current into the transistor is zero, the voltage at the gate is given by a voltage divider, which can be written as

$$V_G = V_{GS} = \left(\frac{R_2}{R_1 + R_2} \right) V_{DD} \quad (3.12)$$

Assuming that the gate-to-source voltage given by Equation (3.12) is greater than V_{TN} , and that the transistor is biased in the saturation region, the drain current is

$$I_D = K_n(V_{GS} - V_{TN})^2 \quad (3.13)$$

The drain-to-source voltage is

$$V_{DS} = V_{DD} - I_D R_D \quad (3.14)$$

If $V_{DS} > V_{DS(\text{sat})} = V_{GS} - V_{TN}$, then the transistor is biased in the saturation region, as we initially assumed, and our analysis is correct. If $V_{DS} < V_{DS(\text{sat})}$, then the transistor is biased in the nonsaturation region, and the drain current is given by the more complicated characteristic Equation (3.2(a)).

The power dissipated in the transistor, since there is no gate current, is simply given by

$$P_T = I_D V_{DS} \quad (3.15)$$

EXAMPLE 3.3

Objective: Calculate the drain current and drain-to-source voltage of a common-source circuit with an n-channel enhancement-mode MOSFET. Find the power dissipated in the transistor.

For the circuit shown in Figure 3.25(a), assume that $R_1 = 30 \text{ k}\Omega$, $R_2 = 20 \text{ k}\Omega$, $R_D = 20 \text{ k}\Omega$, $V_{DD} = 5 \text{ V}$, $V_{TN} = 1 \text{ V}$, and $K_n = 0.1 \text{ mA/V}^2$.

Solution: From the circuit shown in Figure 3.25(b) and Equation (3.12), we have

$$V_G = V_{GS} = \left(\frac{R_2}{R_1 + R_2} \right) V_{DD} = \left(\frac{20}{20 + 30} \right) (5) = 2 \text{ V}$$

Assuming the transistor is biased in the saturation region, the drain current is

$$I_D = K_n(V_{GS} - V_{TN})^2 = (0.1)(2 - 1)^2 = 0.1 \text{ mA}$$

and the drain-to-source voltage is

$$V_{DS} = V_{DD} - I_D R_D = 5 - (0.1)(20) = 3 \text{ V}$$

The power dissipated in the transistor is

$$P_T = I_D V_{DS} = (0.1)(3) = 0.3 \text{ mW}$$

Comment: Because $V_{DS} = 3 \text{ V} > V_{DS(\text{sat})} = V_{GS} - V_{TN} = 2 - 1 = 1 \text{ V}$, the transistor is indeed biased in the saturation region and our analysis is valid.

The dc analysis produces the quiescent values (Q -points) of drain current and drain-to-source voltage, usually indicated by I_{DQ} and V_{DSQ} .

EXERCISE PROBLEM

Ex 3.3: The transistor in Figure 3.25(a) has parameters $V_{TN} = 0.35 \text{ V}$ and $K_n = 25 \mu\text{A/V}^2$. The circuit parameters are $V_{DD} = 2.2 \text{ V}$, $R_1 = 355 \text{ k}\Omega$, $R_2 = 245 \text{ k}\Omega$, and $R_D = 100 \text{ k}\Omega$. Find I_D , V_{GS} , and V_{DS} . (Ans. $I_D = 7.52 \mu\text{A}$, $V_{GS} = 0.898 \text{ V}$, $V_{DS} = 1.45 \text{ V}$)

Figure 3.26 (a) shows a common-source circuit with a p-channel enhancement-mode MOSFET. The source terminal is tied to $+V_{DD}$, which becomes signal ground in the ac equivalent circuit. Thus the terminology common-source applies to this circuit.

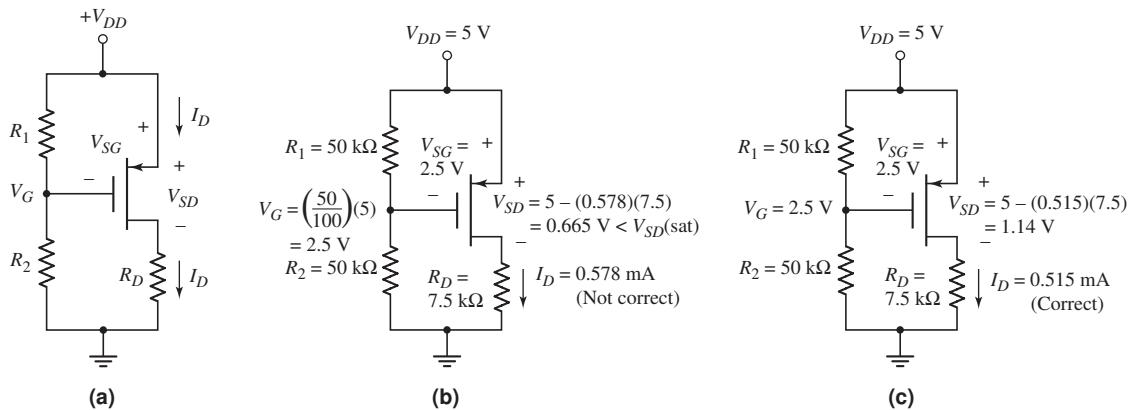


Figure 3.26 (a) A PMOS common-source circuit, (b) the PMOS common-source circuit for Example 3.4 showing current and voltage values when the saturation-region bias assumption is incorrect, and (c) the circuit for Example 3.4 showing current and voltage values when the nonsaturation-region bias assumption is correct

The dc analysis is essentially the same as for the n-channel MOSFET circuit. The gate voltage is

$$V_G = \left(\frac{R_2}{R_1 + R_2} \right) (V_{DD}) \quad (3.16)$$

and the source-to-gate voltage is

$$V_{SG} = V_{DD} - V_G \quad (3.17)$$

Assuming that $V_{GS} < V_{TP}$, or $V_{SG} > |V_{TP}|$, and that the device is biased in the saturation region, the drain current is given by

$$I_D = K_p (V_{SG} + V_{TP})^2 \quad (3.18)$$

and the source-to-drain voltage is

$$V_{SD} = V_{DD} - I_D R_D \quad (3.19)$$

If $V_{SD} > V_{SD}(\text{sat}) = V_{SG} + V_{TP}$, then the transistor is indeed biased in the saturation region, as we have assumed. However, if $V_{SD} < V_{SD}(\text{sat})$, the transistor is biased in the nonsaturation region.

EXAMPLE 3.4

Objective: Calculate the drain current and source-to-drain voltage of a common-source circuit with a p-channel enhancement-mode MOSFET.

Consider the circuit shown in Figure 3.26(a). Assume that $R_1 = R_2 = 50 \text{ k}\Omega$, $V_{DD} = 5 \text{ V}$, $R_D = 7.5 \text{ k}\Omega$, $V_{TP} = -0.8 \text{ V}$, and $K_p = 0.2 \text{ mA/V}^2$.

Solution: From the circuit shown in Figure 3.26(b) and Equation (3.16), we have

$$V_G = \left(\frac{R_2}{R_1 + R_2} \right) (V_{DD}) = \left(\frac{50}{50 + 50} \right) (5) = 2.5 \text{ V}$$

The source-to-gate voltage is therefore

$$V_{SG} = V_{DD} - V_G = 5 - 2.5 = 2.5 \text{ V}$$

Assuming the transistor is biased in the saturation region, the drain current is

$$I_D = K_p (V_{SG} + V_{TP})^2 = (0.2)(2.5 - 0.8)^2 = 0.578 \text{ mA}$$

and the source-to-drain voltage is

$$V_{SD} = V_{DD} - I_D R_D = 5 - (0.578)(7.5) = 0.665 \text{ V}$$

Since $V_{SD} = 0.665 \text{ V}$ is not greater than $V_{SD}(\text{sat}) = V_{SG} + V_{TP} = 2.5 - 0.8 = 1.7 \text{ V}$, the p-channel MOSFET is **not** biased in the saturation region, as we initially assumed.

In the nonsaturation region, the drain current is given by

$$I_D = K_p [2(V_{SG} + V_{TP})V_{SD} - V_{SD}^2]$$

and the source-to-drain voltage is

$$V_{SD} = V_{DD} - I_D R_D$$

Combining these two equations, we obtain

$$I_D = K_p [2(V_{SG} + V_{TP})(V_{DD} - I_D R_D) - (V_{DD} - I_D R_D)^2]$$

or

$$I_D = (0.2)[2(2.5 - 0.8)(5 - I_D(7.5)) - (5 - I_D(7.5))^2]$$

Solving this quadratic equation for I_D , we find

$$I_D = 0.515 \text{ mA}$$

We also find that

$$V_{SD} = 1.14 \text{ V}$$

Therefore, $V_{SD} < V_{SD}(\text{sat})$, which verifies that the transistor is biased in the nonsaturation region.

Comment: In solving the quadratic equation for I_D , we find a second solution that yields $V_{SD} = 2.93 \text{ V}$. However, this value of V_{SD} is greater than $V_{SD}(\text{sat})$, so it is not a valid solution since we assumed the transistor to be biased in the nonsaturation region.

EXERCISE PROBLEM

Ex 3.4: The transistor in Figure 3.26(a) has parameters $V_{TP} = -0.6 \text{ V}$ and $K_p = 0.2 \text{ mA/V}^2$. The circuit is biased at $V_{DD} = 3.3 \text{ V}$. Assume $R_1 \parallel R_2 = 300 \text{ k}\Omega$. Design the circuit such that $I_{DQ} = 0.5 \text{ mA}$ and $V_{SDQ} = 2.0 \text{ V}$. (Ans. $R_1 = 885 \text{ k}\Omega$, $R_2 = 454 \text{ k}\Omega$, $R_D = 2.6 \text{ k}\Omega$)

COMPUTER ANALYSIS EXERCISE

PS 3.1: Verify the results of Example 3.4 with a PSpice analysis.

As Example 3.4 illustrated, we may not know initially whether a transistor is biased in the saturation or nonsaturation region. The approach involves making an educated guess and then verifying that assumption. If the assumption proves incorrect, we must then change it and reanalyze the circuit.

In linear amplifiers containing MOSFETs, the transistors are biased in the saturation region.

DESIGN EXAMPLE 3.5

Objective: Design a MOSFET circuit biased with both positive and negative voltages to meet a set of specifications.

Specifications: The circuit configuration to be designed is shown in Figure 3.27. Design the circuit such that $I_{DQ} = 0.5 \text{ mA}$ and $V_{DSQ} = 4 \text{ V}$.

Choices: Standard resistors are to be used in the final design. A transistor with nominal parameters of $k'_n = 80 \mu\text{A/V}^2$, $(W/L) = 6.25$, and $V_{TN} = 1.2 \text{ V}$ is available.

Solution: Assuming the transistor is biased in the saturation region, we have $I_{DQ} = K_n(V_{GS} - V_{TN})^2$. The conduction parameter is

$$K_n = \frac{k'_n}{2} \cdot \frac{W}{L} = \left(\frac{0.080}{2} \right) (6.25) = 0.25 \text{ mA/V}^2$$

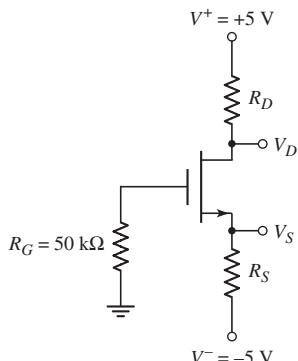


Figure 3.27 Circuit configuration for Example 3.5

Solving for the gate-to-source voltage, we find the required gate-to-source voltage to induce the specified drain current.

$$V_{GS} = \sqrt{\frac{I_{DQ}}{K_n}} + V_{TN} = \sqrt{\frac{0.5}{0.25}} + 1.2$$

or

$$V_{GS} = 2.614 \text{ V}$$

Since the gate current is zero, the gate is at ground potential. The voltage at the source terminal is then $V_S = -V_{GS} = -2.614 \text{ V}$. The value of the source resistor is found from

$$R_S = \frac{V_S - V^-}{I_{DQ}} = \frac{-2.614 - (-5)}{0.5}$$

or

$$R_S = 4.77 \text{ k}\Omega$$

The voltage at the drain terminal is determined to be

$$V_D = V_S + V_{DS} = -2.614 + 4 = 1.386 \text{ V}$$

The value of the drain resistor is

$$R_D = \frac{V^+ - V_D}{I_{DQ}} = \frac{5 - 1.386}{0.5}$$

or

$$R_D = 7.23 \text{ k}\Omega$$

We may note that

$$V_{DS} = 4 \text{ V} > V_{DS(\text{sat})} = V_{GS} - V_{TN} = 2.61 - 1.2 = 1.41 \text{ V}$$

which means that the transistor is indeed biased in the saturation region.

Trade-offs: The closest standard resistor values are $R_S = 4.7 \text{ k}\Omega$ and $R_D = 7.5 \text{ k}\Omega$. We may find the gate-to-source voltage from

$$V_{GS} + I_D R_S - 5 = 0$$

where

$$I_D = K_n(V_{GS} - V_{TN})^2$$

Using the standard resistor values, we find $V_{GS} = 2.622 \text{ V}$, $I_{DQ} = 0.506 \text{ mA}$, and $V_{DSQ} = 3.83 \text{ V}$. In this case, the drain current is within 1.2 percent of the design specification and the drain-to-source voltage is within 4.25 percent of the design specification.

Comment: It is important to keep in mind that the current into the gate terminal is zero. In this case, then, there is zero voltage drop across the R_G resistor.

Design Pointer: In an actual circuit design using discrete elements, we need to choose standard resistor values that are closest to the design values. In addition, the discrete resistors have tolerances that need to be taken into account. In the final design, then, the actual drain current and drain-to-source voltage are somewhat different from the specified values. In many applications, this slight deviation from the specified values will not cause a problem.

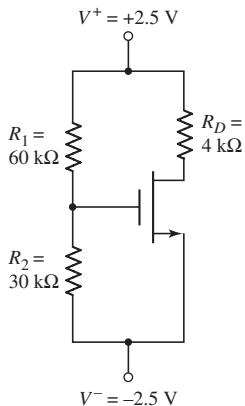


Figure 3.28 Circuit for Exercise Ex 3.5

EXERCISE PROBLEM

Ex 3.5: For the transistor in the circuit in Figure 3.28, the nominal parameter values are $V_{TN} = 0.6$ V and $K_n = 0.5$ mA/V². (a) Determine the quiescent values V_{GSQ} , I_{DQ} , and V_{DSQ} . (b) Determine the range in I_D and V_{DS} values for a ± 5 percent variation in V_{TN} and K_n . (Ans. (a) $V_{GSQ} = 1.667$ V, $I_{DQ} = 0.5689$ mA, $V_{DSQ} = 2.724$ V; (b) $0.5105 \leq I_D \leq 0.6314$ mA, $2.474 \leq V_{DS} \leq 2.958$ V)

Now consider an example of a p-channel device biased with both positive and negative voltages.

DESIGN EXAMPLE 3.6

Objective: Design a circuit with a p-channel MOSFET that is biased with both negative and positive supply voltages and that contains a source resistor R_S to meet a set of specifications.

Specifications: The circuit to be designed is shown in Figure 3.29. Design the circuit such that $I_{DQ} = 100 \mu\text{A}$, $V_{SDQ} = 3$ V, and $V_{RS} = 0.8$ V. Note that V_{RS} is the voltage across the source resistor R_S . The value of the larger bias resistor, either R_1 or R_2 , is to be $200 \text{ k}\Omega$.

Choices: A transistor with nominal parameter values of $K_p = 100 \mu\text{A/V}^2$ and $V_{TP} = -0.4$ V is available. The conduction parameter may vary by ± 5 percent.

Solution: Assuming that the transistor is biased in the saturation region, we have $I_{DQ} = K_p(V_{SG} + V_{TP})^2$. Solving for the source-to-gate voltage, we find the required value of source-to-gate voltage to be

$$V_{SG} = \sqrt{\frac{I_{DQ}}{K_p}} - V_{TP} = \sqrt{\frac{100}{100}} - (-0.4)$$

or

$$V_{SG} = 1.4 \text{ V}$$

We may note that the design value of

$$V_{SDQ} = 3 \text{ V} > V_{SDQ}(\text{sat}) = V_{SGQ} + V_{TP} = 1.4 - 0.4 = 1 \text{ V}$$

so that the transistor will be biased in the saturation region.

The voltage at the gate with respect to ground potential is found to be

$$V_G = V^+ - V_{RS} - V_{SG} = 2.5 - 0.8 - 1.4 = 0.3 \text{ V}$$

With $V_G > 0$, the resistor R_2 will be the larger of the two bias resistors, so set $R_2 = 200 \text{ k}\Omega$. The current through R_2 is then

$$I_{\text{Bias}} = \frac{V_G - V^-}{R_2} = \frac{0.3 - (-2.5)}{200} = 0.014 \text{ mA}$$

Since the current through R_1 is the same, we can find the value of R_1 to be

$$R_1 = \frac{V^+ - V_G}{I_{\text{Bias}}} = \frac{2.5 - 0.3}{0.014}$$

which yields

$$R_1 = 157 \text{ k}\Omega$$

The source resistor value is found from

$$R_S = \frac{V_{RS}}{I_{DQ}} = \frac{0.8}{0.1}$$

or

$$R_S = 8 \text{ k}\Omega$$

The voltage at the drain terminal is

$$V_D = V^+ - V_{RS} - V_{SD} = 2.5 - 0.8 - 3 = -1.3 \text{ V}$$

Then the drain resistor value is found as

$$R_D = \frac{V_D - V^-}{I_{DQ}} = \frac{-1.3 - (-2.5)}{0.1}$$

or

$$R_D = 12 \text{ k}\Omega$$

Trade-offs: If the conduction parameter K_p varies by $\pm 5\%$, the quiescent drain current I_{DQ} and the source-to-drain voltage V_{SDQ} will change. Using the resistor values found in the previous design, we find the following:

K_p	V_{SGQ}	I_{DQ}	V_{SDQ}
95 $\mu\text{A}/\text{V}^2$	1.416 V	98.0 μA	3.04 V
105 $\mu\text{A}/\text{V}^2$	1.385 V	101.9 μA	2.962 V
$\pm 5\%$	$\pm 1.14\%$	$\pm 2\%$	$\pm 1.33\%$

Comment: We may note that the variation in the Q -point values is smaller than the variation in K_p . Including the source resistor R_S tends to stabilize the Q -point.

EXERCISE PROBLEM

Ex 3.6: Consider the circuit shown in Figure 3.30. The nominal transistor parameters are $V_{TP} = -0.30 \text{ V}$ and $K_p = 120 \mu\text{A}/\text{V}^2$. (a) Calculate V_{SG} , I_D , and V_{SD} . (b) Determine the variation in I_D if the threshold voltage varies by ± 5 percent. (Ans. (a) $V_{SG} = 1.631 \text{ V}$, $I_D = 0.2126 \text{ mA}$, $V_{SD} = 3.295 \text{ V}$; (b) $0.2091 \leq I_D \leq 0.2160 \text{ mA}$)

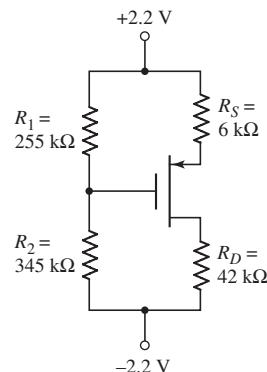


Figure 3.30 Figure for Exercise Ex 3.6

COMPUTER ANALYSIS EXERCISE

PS 3.2 Verify the circuit design in Example 3.6 with a PSpice simulation. Also investigate the change in Q -point values with ± 10 percent variations in resistor values.

3.2.2 Load Line and Modes of Operation

The load line is helpful in visualizing the region in which the MOSFET is biased. Consider again the common-source circuit shown in Figure 3.25(b). Writing a Kirchhoff's voltage law equation around the drain-source loop results in Equation (3.14), which is the load line equation, showing a linear relationship between the drain current and drain-to-source voltage.

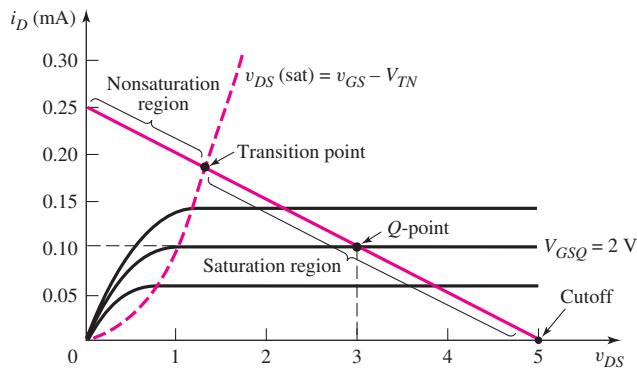


Figure 3.31 Transistor characteristics, $v_{DS}(\text{sat})$ curve, load line, and Q -point for the NMOS common-source circuit in Figure 3.25(b)

Figure 3.31 shows the $v_{DS}(\text{sat})$ characteristic for the transistor described in Example 3.3. The load line is given by

$$V_{DS} = V_{DD} - I_D R_D = 5 - I_D(20) \quad (3.20\text{(a)})$$

or

$$I_D = \frac{5}{20} - \frac{V_{DS}}{20} (\text{mA}) \quad (3.20\text{(b)})$$

and is also plotted in the figure. The two end points of the load line are determined in the usual manner. If $I_D = 0$, then $V_{DS} = 5 \text{ V}$; if $V_{DS} = 0$, then $I_D = 5/20 = 0.25 \text{ mA}$. The Q -point of the transistor is given by the dc drain current and drain-to-source voltage, and it is always on the load line, as shown in the figure. A few transistor characteristics are also shown on the figure.

If the gate-to-source voltage is less than V_{TN} , the drain current is zero and the transistor is in cutoff. As the gate-to-source voltage becomes just greater than V_{TN} , the transistor turns on and is biased in the saturation region. As V_{GS} increases, the Q -point moves up the load line. The **transition point** is the boundary between the saturation and nonsaturation regions and is defined as the point where $V_{DS} = V_{DS}(\text{sat}) = V_{GS} - V_{TN}$. As V_{GS} increases above the transition point value, the transistor becomes biased in the nonsaturation region.

EXAMPLE 3.7

Objective: Determine the transition point parameters for a common-source circuit.

Consider the circuit shown in Figure 3.25(b). Assume transistor parameters of $V_{TN} = 1 \text{ V}$ and $K_n = 0.1 \text{ mA/V}^2$.

Solution: At the transition point,

$$V_{DS} = V_{DS}(\text{sat}) = V_{GS} - V_{TN} = V_{DD} - I_D R_D$$

The drain current is still

$$I_D = K_n(V_{GS} - V_{TN})^2$$

Combining the last two equations, we obtain

$$V_{GS} - V_{TN} = V_{DD} - K_n R_D (V_{GS} - V_{TN})^2$$

Rearranging this equation produces

$$K_n R_D (V_{GS} - V_{TN})^2 + (V_{GS} - V_{TN}) - V_{DD} = 0$$

or

$$(0.1)(20)(V_{GS} - V_{TN})^2 + (V_{GS} - V_{TN}) - 5 = 0$$

Solving the quadratic equation, we find that

$$V_{GS} - V_{TN} = 1.35 \text{ V} = V_{DS}$$

Therefore,

$$V_{GS} = 2.35 \text{ V}$$

and

$$I_D = (0.1)(2.35 - 1)^2 = 0.182 \text{ mA}$$

Comment: For $V_{GS} < 2.35 \text{ V}$, the transistor is biased in the saturation region; for $V_{GS} > 2.35 \text{ V}$, the transistor is biased in the nonsaturation region.

EXERCISE PROBLEM

Ex 3.7: Consider the circuit in Figure 3.30. Using the nominal transistor parameters described in Exercise Ex 3.6, draw the load line and determine the transition point parameters. (Ans. $V_{SG} = 2.272 \text{ V}$, $I_D = 0.4668 \text{ mA}$, $V_{SD} = 1.972 \text{ V}$)

Problem-Solving Technique: MOSFET DC Analysis

Analyzing the dc response of a MOSFET circuit requires knowing the bias condition (saturation or nonsaturation) of the transistor. In some cases, the bias condition may not be obvious, which means that we have to guess the bias condition, then analyze the circuit to determine if we have a solution consistent with our initial guess. To do this, we can:

1. Assume that the transistor is biased in the saturation region, in which case $V_{GS} > V_{TN}$, $I_D > 0$, and $V_{DS} \geq V_{DS}(\text{sat})$.
2. Analyze the circuit using the saturation current-voltage relations.
3. Evaluate the resulting bias condition of the transistor. If the assumed parameter values in step 1 are valid, then the initial assumption is correct. If $V_{GS} < V_{TN}$, then the transistor is probably cutoff, and if $V_{DS} < V_{DS}(\text{sat})$, the transistor is likely biased in the nonsaturation region.
4. If the initial assumption is proved incorrect, then a new assumption must be made and the circuit reanalyzed. Step 3 must then be repeated.

3.2.3 Additional MOSFET Configurations: DC Analysis

There are other MOSFET circuits, in addition to the basic common-source circuits just considered, that are biased with the basic four-resistor configuration.

However, MOSFET integrated circuit amplifiers are generally biased with constant current sources. Example 3.8 demonstrates this technique using an ideal current source.

DESIGN EXAMPLE 3.8

Objective: Design a MOSFET circuit biased with a constant-current source to meet a set of specifications.

Specifications: The circuit configuration to be designed is shown in Figure 3.32(a). Design the circuit such that the quiescent values are $I_{DQ} = 250 \mu\text{A}$ and $V_D = 2.5 \text{ V}$.

Choices: A transistor with nominal values of $V_{TN} = 0.8 \text{ V}$, $k'_n = 80 \mu\text{A/V}^2$, and $W/L = 3$ is available. Assume k'_n varies by ± 5 percent.

Solution: The dc equivalent circuit is shown in Figure 3.32(b). Since $v_i = 0$, the gate is at ground potential and there is no current through R_G . We have that $I_Q = I_{DQ} = 250 \mu\text{A}$.

Assuming the transistor is biased in the saturation region, we have

$$I_D = \frac{k'_n}{2} \cdot \frac{W}{L} (V_{GS} - V_{TN})^2$$

or

$$250 = \left(\frac{80}{2}\right) \cdot (3)(V_{GS} - 0.8)^2$$

which yields

$$V_{GS} = 2.24 \text{ V}$$

The voltage at the source terminal is $V_S = -V_{GS} = -2.24 \text{ V}$.

The drain current can also be written as

$$I_D = \frac{5 - V_D}{R_D}$$

For $V_D = 2.5 \text{ V}$, we have

$$R_D = \frac{5 - 2.5}{0.25} = 10 \text{ k}\Omega$$

The drain-to-source voltage is

$$V_{DS} = V_D - V_S = 2.5 - (-2.24) = 4.74 \text{ V}$$

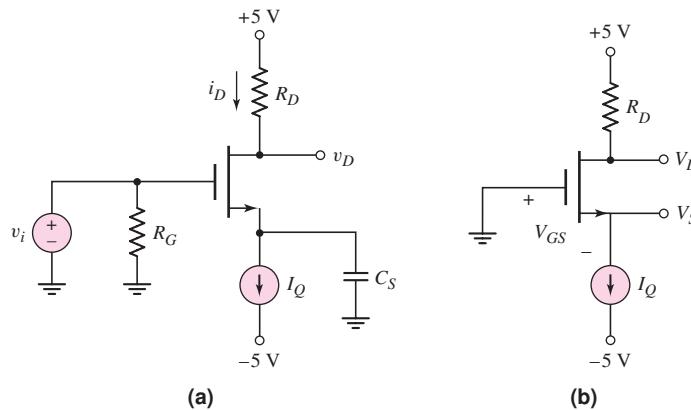


Figure 3.32 (a) NMOS common-source circuit biased with a constant-current source and (b) equivalent dc circuit

Since $V_{DS} = 4.74 \text{ V} > V_{DS(\text{sat})} = V_{GS} - V_{TN} = 2.24 - 0.8 = 1.44 \text{ V}$, the transistor is biased in the saturation region, as initially assumed.

Trade-offs: Note that even if k'_n changes, the drain current remains constant. For $76 \leq k'_n \leq 84 \mu\text{A/V}^2$, the variation in V_{GSQ} is $2.209 \leq V_{GSQ} \leq 2.281 \text{ V}$ and the variation in V_{DSQ} is $4.709 \leq V_{DSQ} \leq 4.781 \text{ V}$. The variation in V_{DSQ} is ± 0.87 percent even with a ± 5 percent variation in k'_n . This stability effect is one advantage of using constant current biasing.

Comment: MOSFET circuits can be biased by using constant-current sources, which in turn are designed by using other MOS transistors, as we will see. Biasing with current sources tends to stabilize circuits against variations in device or circuit parameters.

EXERCISE PROBLEM

Ex 3.8: (a) Consider the circuit shown in Figure 3.33. The transistor parameters are $V_{TP} = -0.40 \text{ V}$ and $K_p = 30 \mu\text{A/V}^2$. Design the circuit such that $I_{DQ} = 60 \mu\text{A}$ and $V_{SDQ} = 2.5 \text{ V}$. (b) Determine the variation in Q -point values if the V_{TP} and K_p parameters vary by ± 5 percent. (Ans. (a) $R_S = 19.77 \text{ k}\Omega$, $R_D = 38.57 \text{ k}\Omega$; (b) $58.2 \leq I_{DQ} \leq 61.08 \mu\text{A}$, $2.437 \leq V_{SDQ} \leq 2.605 \text{ V}$)

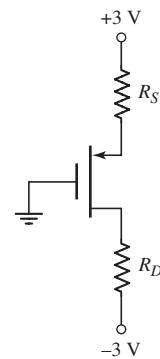


Figure 3.33 Circuit for Exercise Ex 3.8

n-Channel Enhancement-Load Device

An enhancement-mode MOSFET connected in a configuration such as that shown in Figure 3.34 can be used as a nonlinear resistor. A transistor with this connection is called an **enhancement-load device**. Since the transistor is an enhancement mode device, $V_{TN} > 0$. Also, for this circuit, $v_{DS} = v_{GS} > v_{DS(\text{sat})} = v_{GS} - V_{TN}$, which means that the transistor is always biased in the saturation region. The general i_D versus v_{DS} characteristics can then be written as

$$i_D = K_n(v_{GS} - V_{TN})^2 = K_n(v_{DS} - V_{TN})^2 \quad (3.21)$$

Figure 3.35 shows a plot of Equation (3.21) for the case when $K_n = 1 \text{ mA/V}^2$ and $V_{TN} = 1 \text{ V}$.

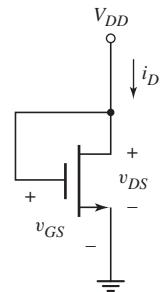


Figure 3.34 Enhancement-mode NMOS device with the gate connected to the drain

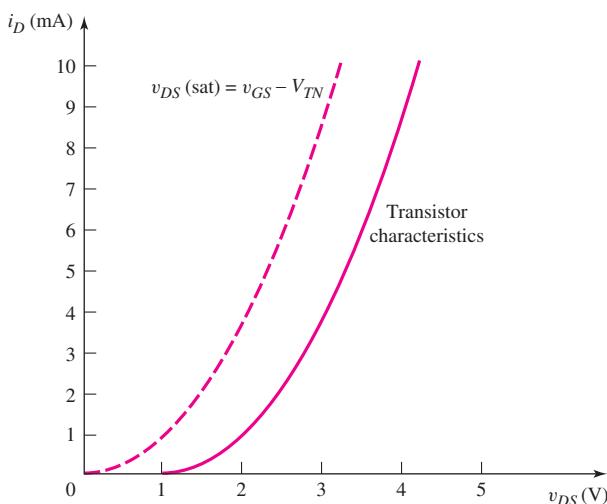


Figure 3.35 Current–voltage characteristic of an enhancement load device

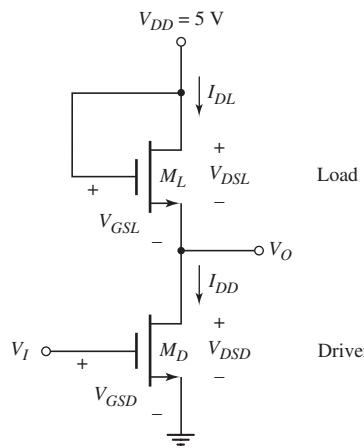


Figure 3.36 Circuit with enhancement-load device and NMOS driver

If an enhancement-load device is connected in a circuit with another MOSFET in the configuration shown in Figure 3.36, the circuit can be used as an amplifier or as an inverter in a digital logic circuit. The load device, M_L , is always biased in the saturation region, and the transistor M_D , called the **driver transistor**, can be biased in either the saturation or nonsaturation region, depending on the value of the input voltage. The next example addresses the dc analysis of this circuit for dc input voltages to the gate of M_D .

EXAMPLE 3.9

Objective: Determine the dc transistor currents and voltages in a circuit containing an enhancement load device.

The transistors in the circuit shown in Figure 3.36 have parameters $V_{TND} = V_{TNL} = 1 \text{ V}$, $K_{nD} = 50 \mu\text{A/V}^2$, and $K_{nL} = 10 \mu\text{A/V}^2$. Also assume $\lambda_{nD} = \lambda_{nL} = 0$. (The subscript D applies to the driver transistor and the subscript L applies to the load transistor.) Determine V_O for $V_I = 5 \text{ V}$ and $V_I = 1.5 \text{ V}$.

Solution: ($V_I = 5 \text{ V}$) For an inverter circuit with a resistive load, when the input voltage is large, the output voltage drops to a low value. Therefore, we assume that the driver transistor is biased in the nonsaturation region since the drain-to-source voltage will be small. The drain current in the load device is equal to the drain current in the driver transistor. Writing these currents in generic form, we have

$$I_{DD} = I_{DL}$$

or

$$K_{nD}[2(V_{GSD} - V_{TND})V_{DSD} - V_{DSD}^2] = K_{nL}[V_{GSL} - V_{TNL}]^2$$

Since $V_{GSD} = V_I$, $V_{DSD} = V_O$, and $V_{GSL} = V_{DSL} = V_{DD} - V_O$, then

$$K_{nD}[2(V_I - V_{TND})V_O - V_O^2] = K_{nL}[V_{DD} - V_O - V_{TNL}]^2$$

Substituting numbers, we find

$$(50)[2(5 - 1)V_O - V_O^2] = (10)[5 - V_O - 1]^2$$

Rearranging the terms provides

$$3V_O^2 - 24V_O + 8 = 0$$

Using the quadratic formula, we obtain two possible solutions:

$$V_O = 7.65 \text{ V} \quad \text{or} \quad V_O = 0.349 \text{ V}$$

Since the output voltage cannot be greater than the supply voltage $V_{DD} = 5 \text{ V}$, the valid solution is $V_O = 0.349 \text{ V}$.

Also, since $V_{DSD} = V_O = 0.349 \text{ V} < V_{GSD} - V_{TND} = 5 - 1 = 4 \text{ V}$, the driver M_D is biased in the nonsaturation region, as initially assumed.

The current can be determined from

$$I_D = K_{nL}(V_{GSL} - V_{TNL})^2 = K_{nL}(V_{DD} - V_O - V_{TNL})^2$$

or

$$I_D = (10)(5 - 0.349 - 1)^2 = 133 \mu\text{A}$$

Solution: ($V_I = 1.5 \text{ V}$) Since the threshold voltage of the driver transistor is $V_{TN} = 1 \text{ V}$, an input voltage of 1.5 V means the transistor current is going to be relatively small so the output voltage should be relatively large. For this reason, we will assume that the driver transistor M_D is biased in the saturation region. Equating the currents in the two transistors and writing the current equations in generic form, we have

$$I_{DD} = I_{DL}$$

or

$$K_{nD}[V_{GSD} - V_{TND}]^2 = K_{nL}[V_{GSL} - V_{TNL}]^2$$

Again, since $V_{GSD} = V_I$ and $V_{GSL} = V_{DSL} = V_{DD} - V_O$, then

$$K_{nD}[V_I - V_{TND}]^2 = K_{nL}[V_{DD} - V_O - V_{TNL}]^2$$

Substituting numbers and taking the square root, we find

$$\sqrt{50}[1.5 - 1] = \sqrt{10}[5 - V_O - 1]$$

which yields $V_O = 2.88 \text{ V}$.

Since $V_{DSD} = V_O = 2.88 \text{ V} > V_{GSD} - V_{TND} = 1.5 - 1 = 0.5 \text{ V}$, the driver transistor M_D is biased in the saturation region, as initially assumed.

The current is

$$I_D = K_{nD}(V_{GSD} - V_{TND})^2 = (50)(1.5 - 1)^2 = 12.5 \mu\text{A}$$

Comment: For this example, we made an initial guess as to whether the driver transistor was biased in the saturation or nonsaturation region. A more analytical approach is shown following this example.

Computer Simulation: The voltage transfer characteristics of the NMOS inverter with enhancement load shown in Figure 3.36 were obtained by a PSpice analysis. These results are shown in Figure 3.37. As the input voltage decreases from its high state, the output voltage increases, charging and discharging capacitances in the transistors. The current in the circuit goes to zero when the driver transistor is cutoff. This occurs when $V_I = V_{GSD} = V_{TN} = 1 \text{ V}$. At this point, the output voltage is $V_O = 4 \text{ V}$. Since there is no current, the capacitances cease charging and discharging so the output voltage cannot get to the full $V_{DD} = 5 \text{ V}$ value. The maximum output voltage is $V_O(\max) = V_{DD} - V_{TNL} = 5 - 1 = 4 \text{ V}$.

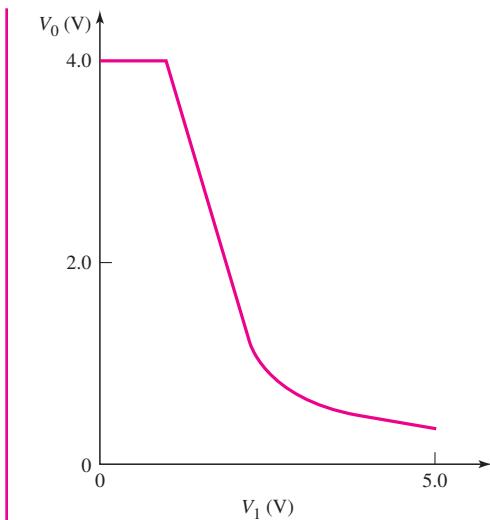


Figure 3.37 Voltage transfer characteristics of NMOS inverter with enhancement load device

When the input voltage is just greater than 1 V, both transistors are biased in the saturation region as the previous analysis for $V_I = 1.5$ V showed. The output voltage is a linear function of input voltage as we will see in Equation (3.24).

For an input voltage greater than approximately 2.25 V, the driver transistor is biased in the nonsaturation region and the output voltage is a nonlinear function of input voltage.

EXERCISE PROBLEM

Ex 3.9: Consider the NMOS inverter shown in Figure 3.36 with transistor parameters described in Example 3.9. Determine the output voltage V_O for input voltages (a) $V_I = 4$ V and (b) $V_I = 2$ V. (Ans. (a) 0.454 V, (b) 1.76 V)

COMPUTER ANALYSIS EXERCISE

PS 3.3: Consider the NMOS circuit shown in Figure 3.36. Plot the voltage transfer characteristics, using a PSpice simulation. Use transistor parameters similar to those in Example 3.9. What are the values of V_O for $V_I = 1.5$ V and $V_I = 5$ V?

In the circuit shown in Figure 3.36, we can determine the transition point for the driver transistor that separates the saturation and nonsaturation regions. The transition point is determined by the equation

$$V_{DSD}(\text{sat}) = V_{GSD} - V_{TND} \quad (3.22)$$

Again, the drain currents in the two transistors are equal. Using the saturation drain current relationship for the driver transistor, we have

$$I_{DD} = I_{DL} \quad (3.23(a))$$

or

$$K_{nD}[V_{GSD} - V_{TND}]^2 = K_{nL}[V_{GSL} - V_{TNL}]^2 \quad (3.23(b))$$

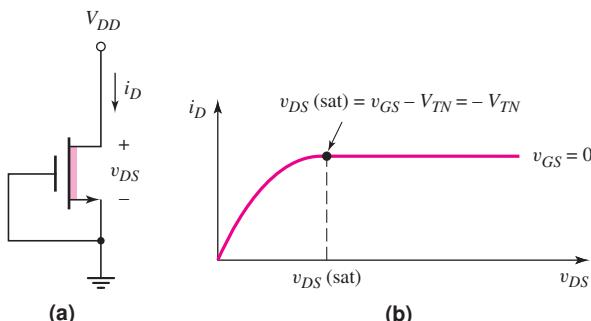


Figure 3.38 (a) Depletion-mode NMOS device with the gate connected to the source and (b) current–voltage characteristics

Again, noting that $V_{GSD} = V_I$ and $V_{GSL} = V_{DSL} = V_{DD} - V_O$, and taking the square root, we have

$$\sqrt{\frac{K_{nD}}{K_{nL}}}(V_I - V_{TND}) = (V_{DD} - V_O - V_{TNL}) \quad (3.24)$$

At the transition point, we can define the input voltage as $V_I = V_{It}$ and the output voltage as $V_{Ot} = V_{DS(\text{sat})} = V_{It} - V_{TND}$. Then, from Equation (3.24), the input voltage at the transition point is

$$V_{It} = \frac{V_{DD} - V_{TNL} + V_{TND}(1 + \sqrt{K_{nD}/K_{nL}})}{1 + \sqrt{K_{nD}/K_{nL}}} \quad (3.25)$$

If we apply Equation (3.25) to the previous example, we can show that our initial assumptions were correct.

n-Channel Depletion-Load Device

An n-channel depletion-mode MOSFET can also be used as a load device. Consider the depletion-mode MOSFET with the gate and source connected together shown in Figure 3.38(a). The current–voltage characteristics are shown in Figure 3.38(b). The transistor may be biased in either the saturation or nonsaturation regions. The transition point is also shown on the plot. The threshold voltage of the n-channel depletion-mode MOSFET is negative so that $v_{DS(\text{sat})}$ is positive.

A depletion-load device can be used in conjunction with another MOSFET, as shown in Figure 3.39, to create a circuit that can be used as an amplifier or as an inverter in a digital logic circuit. Both the load device M_L and driver transistor M_D may be biased in either the saturation or nonsaturation region, depending on the value of the input voltage. We will perform the dc analysis of this circuit for a particular dc input voltage to the gate of the driver transistor.

EXAMPLE 3.10

Objective: Determine the dc transistor currents and voltages in a circuit containing a depletion load device.

Consider the circuit shown in Figure 3.39. Let $V_{DD} = 5$ V and assume transistor parameters of $V_{TND} = 1$ V, $V_{TNL} = -2$ V, $K_{nD} = 50 \mu\text{A/V}^2$, and $K_{nL} = 10 \mu\text{A/V}^2$. Determine V_O for $V_I = 5$ V.

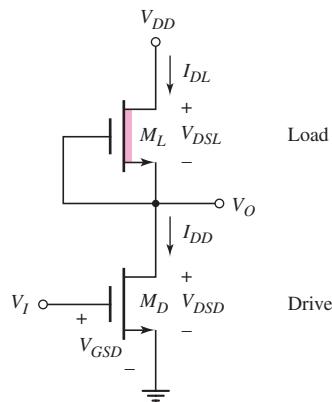


Figure 3.39 Circuit with depletion-load device and NMOS driver

Solution: Assume the driver transistor M_D is biased in the nonsaturation region and the load transistor M_L is biased in the saturation region. The drain currents in the two transistors are equal. In generic form, these currents are

$$I_{DD} = I_{DL}$$

or

$$K_{nD}[2(V_{GSD} - V_{TND})V_{DSD} - V_{DSD}^2] = K_{nL}[V_{GSL} - V_{TNL}]^2$$

Since $V_{GSD} = V_I$, $V_{DSD} = V_O$, and $V_{GSL} = 0$, then

$$K_{nD}[2(V_I - V_{TND})V_O - V_O^2] = K_{nL}[-V_{TNL}]^2$$

Substituting numbers, we find

$$(50)[2(5 - 1)V_O - V_O^2] = (10)[-(-2)]^2$$

Rearranging the terms produces

$$5V_O^2 - 40V_O + 4 = 0$$

Using the quadratic formula, we obtain two possible solutions:

$$V_O = 7.90 \text{ V} \quad \text{or} \quad V_O = 0.10 \text{ V}$$

Since the output voltage cannot be greater than the supply voltage $V_{DD} = 5 \text{ V}$, the valid solution is $V_O = 0.10 \text{ V}$.

The current is

$$I_D = K_{nL}(-V_{TNL})^2 = (10)[-(-2)]^2 = 40 \mu\text{A}$$

Comment: Since $V_{DSD} = V_O = 0.10 \text{ V} < V_{GSD} - V_{TND} = 5 - 1 = 4 \text{ V}$, M_D is biased in the nonsaturation region, as assumed. Similarly, since $V_{DSL} = V_{DD} - V_O = 4.9 \text{ V} > V_{GSL} - V_{TNL} = 0 - (-2) = 2 \text{ V}$, M_L is biased in the saturation region, as originally assumed.

Computer Simulation: The voltage transfer characteristics of the NMOS inverter circuit with depletion load in Figure 3.39 were obtained using a PSpice analysis. These results are shown in Figure 3.40. For an input voltage less than 1 V, the driver is cut off and the output voltage is $V_O = V_{DD} = 5 \text{ V}$.

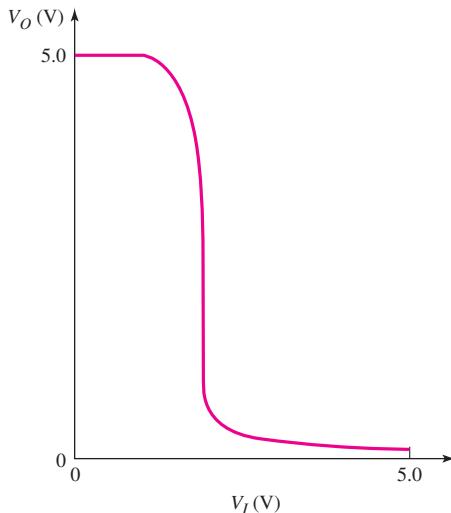


Figure 3.40 Voltage transfer characteristics of NMOS inverter with depletion load device

When the input voltage is just greater than 1 V, the driver transistor is biased in the saturation region and the load device in the nonsaturation region. When the input voltage is approximately 1.9 V, both transistors are biased in the saturation region. If the channel length modulation parameter λ is assumed to be zero as in this example, there is no change in the input voltage during this transition region. As the input voltage becomes larger than 1.9 V, the driver is biased in the nonsaturation region and the load in the saturation region.

EXERCISE PROBLEM

Ex 3.10: Consider the circuit shown in Figure 3.39 with transistor parameters $V_{TND} = 1\text{ V}$ and $V_{TNL} = -2\text{ V}$. (a) Design the ratio K_{nD}/K_{nL} that will produce an output voltage of $V_O = 0.25\text{ V}$ at $V_I = 5\text{ V}$. (b) Find K_{nD} and K_{nL} if the transistor currents are 0.2 mA when $V_I = 5\text{ V}$. (Ans. (a) $K_{nD}/K_{nL} = 2.06$ (b) $K_{nL} = 50 \mu\text{A/V}^2$, $K_{nD} = 103 \mu\text{A/V}^2$)

COMPUTER ANALYSIS EXERCISE

PS 3.4: Consider the NMOS circuit shown in Figure 3.39. Plot the voltage transfer characteristics using a PSpice simulation. Use transistor parameters similar to those in Example 3.10. What are the values of V_O for $V_I = 1.5\text{ V}$ and $V_I = 5\text{ V}$?

p-Channel Enhancement-Load Device

A p-channel enhancement-mode transistor can also be used as a load device to form a **complementary MOS (CMOS)** inverter. The term complementary implies that both n-channel and p-channel transistors are used in the same circuit. The CMOS technology is used extensively in both analog and digital electronic circuits.

Figure 3.41 shows one example of a CMOS inverter. The NMOS transistor is used as the amplifying device, or the driver, and the PMOS device is the load, which is referred to as an active load. This configuration is typically used in analog applications.

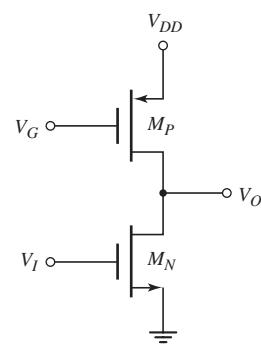


Figure 3.41 Example of CMOS inverter

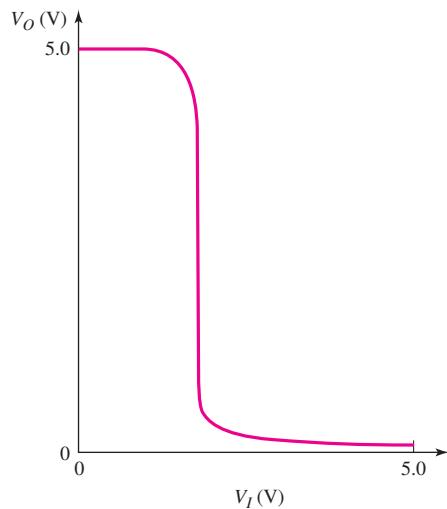


Figure 3.42 Voltage transfer characteristics of CMOS inverter in Figure 3.41

In another configuration, the two gates are tied together and form the input. This configuration will be discussed in detail in Chapter 16.

As with the previous two NMOS inverters, the two transistors shown in Figure 3.41 may be biased in either the saturation or nonsaturation region, depending on the value of the input voltage. The voltage transfer characteristic is most easily determined from a PSpice analysis.

EXAMPLE 3.11

Objective: Determine the voltage transfer characteristic of the CMOS inverter using a PSpice analysis.

For the circuit shown in Figure 3.41, assume transistor parameters of $V_{TN} = 1$ V, $V_{TP} = -1$ V, and $K_n = K_p$. Also assume $V_{DD} = 5$ V and $V_G = 3.25$ V.

Solution: The voltage transfer characteristics are shown in Figure 3.42. In this case, there is a region, as was the case for an NMOS inverter with depletion load, in which both transistors are biased in the saturation region, and the input voltage is a constant over this transition region for the assumption that the channel length modulation parameter λ is zero.

Comment: In this example, the source-to-gate voltage of the PMOS device is only $V_{SG} = 1.75$ V. The effective resistance looking into the drain of the PMOS device is then relatively large. This is a desirable characteristic for an amplifier, as we will see in the next chapter.

EXERCISE PROBLEM

Ex 3.11: Consider the circuit in Figure 3.41. Assume the same transistor parameters and circuit parameters as given in Example 3.11. Determine the transition point parameters for the transistors M_N and M_P . (Ans. M_P : $V_{Ot} = 4.25$ V, $V_{It} = 1.75$ V; M_N : $V_{Ot} = 0.75$ V, $V_{It} = 1.75$ V)

Test Your Understanding

TYU 3.7 The transistor in the circuit shown in Figure 3.25(a) has parameters $V_{TN} = 0.25$ V and $K_n = 30 \mu\text{A/V}^2$. The circuit is biased at $V_{DD} = 2.2$ V. Let $R_1 + R_2 = 500 \text{ k}\Omega$. Redesign the circuit such that $I_{DQ} = 70 \mu\text{A}$ and $V_{DSQ} = 1.2$ V. (Ans. $R_1 = 96 \text{ k}\Omega$, $R_2 = 404 \text{ k}\Omega$, $R_D = 14.3 \text{ k}\Omega$)

TYU 3.8 Consider the circuit in Figure 3.43. The transistor parameters are $V_{TN} = 0.4$ V and $k'_n = 100 \mu\text{A/V}^2$. Design the transistor width-to-length ratio such that $V_{DS} = 1.6$ V. (Ans. 2.36)

TYU 3.9 For the circuit shown in Figure 3.36, use the transistor parameters given in Example 3.9. (a) Determine V_I and V_O at the transition point for the driver transistor. (b) Calculate the transistor currents at the transition point. (Ans. (a) $V_{It} = 2.236$ V, $V_{Ot} = 1.236$ V; (b) $I_D = 76.4 \mu\text{A}$)

TYU 3.10 Consider the circuit shown in Figure 3.44. The transistor parameters are $V_{TN} = -1.2$ V and $k'_n = 80 \mu\text{A/V}^2$. (a) Design the transistor width-to-length ratio such that $V_{DS} = 1.8$ V. Is the transistor biased in the saturation or nonsaturation region? (b) Repeat part (a) for $V_{DS} = 0.8$ V. (Ans. (a) 3.26, (b) 6.10)

TYU 3.11 For the circuit shown in Figure 3.39, use the transistor parameters given in Example 3.10. (a) Determine V_I and V_O at the transition point for the load transistor. (b) Determine V_I and V_O at the transition point for the driver transistor. (Ans. (a) $V_{It} = 1.89$ V, $V_{Ot} = 3$ V; (b) $V_{It} = 1.89$ V, $V_{Ot} = 0.89$ V)

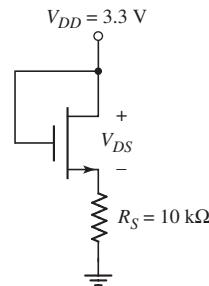


Figure 3.43 Circuit for Exercise TYU 3.8

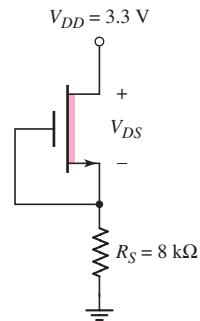


Figure 3.44 Circuit for Exercise TYU 3.10



3.3 BASIC MOSFET APPLICATIONS: SWITCH, DIGITAL LOGIC GATE, AND AMPLIFIER

Objective: • Examine three applications of MOSFET circuits: a switch circuit, digital logic circuit, and an amplifier circuit.

MOSFETs may be used to: switch currents, voltages, and power; perform digital logic functions; and amplify small time-varying signals. In this section, we will examine the switching properties of an NMOS transistor, analyze a simple NMOS transistor digital logic circuit, and discuss how the MOSFET can be used to amplify small signals.

3.3.1 NMOS Inverter

The MOSFET can be used as a switch in a wide variety of electronic applications. The transistor switch provides an advantage over mechanical switches in both speed and reliability. The transistor switch considered in this section is also called an inverter. Two other switch configurations, the NMOS transmission gate and the CMOS transmission gate, are discussed in Chapter 16.

Figure 3.45 shows the n-channel enhancement-mode MOSFET inverter circuit. If $v_I < V_{TN}$, the transistor is in cutoff and $i_D = 0$. There is no voltage drop across R_D , and the output voltage is $v_O = V_{DD}$. Also, since $i_D = 0$, no power is dissipated in the transistor.

If $v_I > V_{TN}$, the transistor is on and initially is biased in the saturation region, since $v_{DS} > v_{GS} - V_{TN}$. As the input voltage increases, the drain-to-source voltage

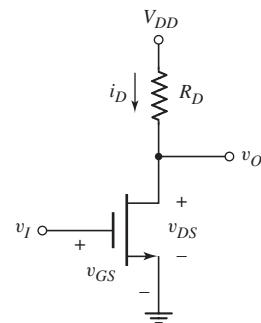


Figure 3.45 NMOS inverter circuit

decreases, and the transistor eventually becomes biased in the nonsaturation region. When $v_I = V_{DD}$, the transistor is biased in the nonsaturation region, v_O reaches a minimum value, and the drain current reaches a maximum value. The current and voltage are given by

$$i_D = K_n [2(v_I - V_{TN})v_O - v_O^2] \quad (3.26)$$

and

$$v_O = v_{DD} - i_D R_D \quad (3.27)$$

where $v_O = v_{DS}$ and $v_I = v_{GS}$.

DESIGN EXAMPLE 3.12

Objective: Design the size of a power MOSFET to meet the specification of a particular switch application.

The load in the inverter circuit in Figure 3.45 is a coil of an electromagnet that requires a current of 0.5 A when turned on. The effective load resistance varies between 8 and 10 Ω, depending on temperature and other variables. A 10 V power supply is available. The transistor parameters are $k'_n = 80 \mu\text{A/V}^2$ and $V_{TN} = 1 \text{ V}$.

Solution: One solution is to bias the transistor in the saturation region so that the current is constant, independent of the load resistance.

The minimum V_{DS} value is 5 V. We need $V_{DS} > V_{DS(\text{sat})} = V_{GS} - V_{TN}$. If we bias the transistor at $V_{GS} = 5 \text{ V}$, then the transistor will always be biased in the saturation region. We can then write

$$I_D = \frac{k'_n}{2} \cdot \frac{W}{L} (V_{GS} - V_{TN})^2$$

or

$$0.5 = \frac{80 \times 10^{-6}}{2} \left(\frac{W}{L} \right) \cdot (5 - 1)^2$$

which yields $W/L = 781$.

The maximum power dissipation in the transistor occurs when the load resistance is 8 Ω and $V_{DS} = 6 \text{ V}$. Then

$$P(\max) = V_{DS}(\max) \cdot I_D = (6) \cdot (0.5) = 3 \text{ W}$$

Comment: We see that we can switch a relatively large drain current with essentially no input current to the transistor. The size of the transistor required is fairly large, which implies a power transistor is necessary. If a transistor with a slightly different width-to-length ratio is available, the applied V_{GS} can be changed to meet the specification.

EXERCISE PROBLEM

Ex 3.12: For the MOS inverter circuit shown in Figure 3.45, assume the circuit values are $V_{DD} = 5 \text{ V}$ and $R_D = 500 \Omega$. The threshold voltage of the transistor is $V_{TN} = 1 \text{ V}$. (a) Determine the value of the conduction parameter K_n such that $v_O = 0.2 \text{ V}$ when $v_I = 5 \text{ V}$. (b) What is the power dissipated in the transistor? (Ans. (a) $K_n = 6.15 \text{ mA/V}^2$, (b) $P = 1.92 \text{ mW}$)

3.3.2 Digital Logic Gate

For the transistor inverter circuit in Figure 3.45, when the input is low and approximately zero volts, the transistor is cut off, and the output is high and equal to V_{DD} . When the input is high and equal to V_{DD} , the transistor is biased in the nonsaturation region and the output reaches a low value. Since the input voltages will be either high or low, we can analyze the circuit in terms of dc parameters.

Now consider the case when a second transistor is connected in parallel, as shown in Figure 3.46. If the two inputs are zero, both M_1 and M_2 are cut off, and $V_O = 5$ V. When $V_1 = 5$ V and $V_2 = 0$, the transistor M_1 turns on and M_2 is still cut off. Transistor M_1 is biased in the nonsaturation region, and V_O reaches a low value. If we reverse the input voltages such that $V_1 = 0$ and $V_2 = 5$ V, then M_1 is cut off and M_2 is biased in the nonsaturation region. Again, V_O is at a low value. If both inputs are high, at $V_1 = V_2 = 5$ V, then both transistors are biased in the nonsaturation region and V_O is low.

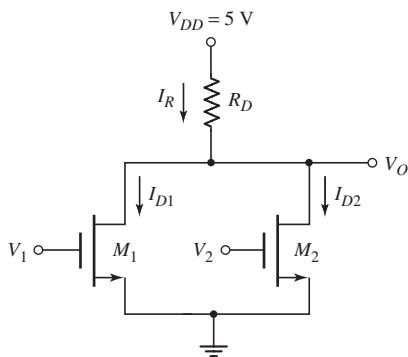


Figure 3.46 A two-input NMOS NOR logic gate

Table 3.2 shows these various conditions for the circuit in Figure 3.46. In a positive logic system, these results indicate that this circuit performs the NOR logic function, and, it is therefore called a two-input NOR logic circuit. In actual NMOS logic circuits, the resistor R_D is replaced by another NMOS transistor.

Table 3.2 NMOS NOR logic circuit response

$V_1(V)$	$V_2(V)$	$V_O(V)$
0	0	High
5	0	Low
0	5	Low
5	5	Low

EXAMPLE 3.13

Objective: Determine the currents and voltages in a digital logic gate, for various input conditions.

Consider the circuit shown in Figure 3.46 with circuit and transistor parameters $R_D = 20 \text{ k}\Omega$, $K_n = 0.1 \text{ mA/V}^2$, $V_{TN} = 0.8 \text{ V}$, and $\lambda = 0$.

Solution: For $V_1 = V_2 = 0$, both M_1 and M_2 are cut off and $V_O = V_{DD} = 5\text{ V}$. For $V_1 = 5\text{ V}$ and $V_2 = 0$, the transistor M_1 is biased in the nonsaturation region, and we can write

$$I_R = I_{D1} = \frac{5 - V_O}{R_D} = K_n [2(V_1 - V_{TN})V_O - V_O^2]$$

Solving for the output voltage V_O , we obtain $V_O = 0.29\text{ V}$.

The currents are

$$I_R = I_{D1} = \frac{5 - 0.29}{20} = 0.236\text{ mA}$$

For $V_1 = 0$ and $V_2 = 5\text{ V}$, we have $V_O = 0.29\text{ V}$ and $I_R = I_{D2} = 0.236\text{ mA}$. When both inputs go high to $V_1 = V_2 = 5\text{ V}$, we have $I_R = I_{D1} + I_{D2}$, or

$$\frac{5 - V_O}{R_D} = K_n [2(V_1 - V_{TN})V_O - V_O^2] + K_n [2(V_2 - V_{TN})V_O - V_O^2]$$

which can be solved for V_O to yield $V_O = 0.147\text{ V}$.

The currents are

$$I_R = \frac{5 - 0.147}{20} = 0.243\text{ mA}$$

and

$$I_{D1} = I_{D2} = \frac{I_R}{2} = 0.121\text{ mA}$$

Comment: When either transistor is biased on, it is biased in the nonsaturation region, since $V_{DS} < V_{DS(\text{sat})}$, and the output voltage reaches a low state.

EXERCISE PROBLEM

Ex 3.13: For the circuit in Figure 3.46, assume the circuit and transistor parameters are: $R_D = 30\text{ k}\Omega$, $V_{TN} = 1\text{ V}$, and $K_n = 50\text{ }\mu\text{A/V}^2$. Determine V_O , I_R , I_{D1} , and I_{D2} for: (a) $V_1 = 5\text{ V}$, $V_2 = 0$; and (b) $V_1 = V_2 = 5\text{ V}$. (Ans. (a) $V_O = 0.40\text{ V}$, $I_R = I_{D1} = 0.153\text{ mA}$, $I_{D2} = 0$ (b) $V_O = 0.205\text{ V}$, $I_R = 0.16\text{ mA}$, $I_{D1} = I_{D2} = 0.080\text{ mA}$)

This example and discussion illustrates that MOS transistors can be configured in a circuit to perform logic functions. A more detailed analysis and design of MOSFET logic gates and circuits is presented in Chapter 16. As we will see in that chapter, most MOS logic gate circuits are fabricated by using CMOS, which means designing circuits with both n-channel and p-channel transistors and no resistors.

3.3.3 MOSFET Small-Signal Amplifier

The MOSFET, in conjunction with other circuit elements, can amplify small time-varying signals. Figure 3.47(a) shows the MOSFET small-signal amplifier, which is a common-source circuit in which a time-varying signal is coupled to the gate through a coupling capacitor. Figure 3.47(b) shows the transistor characteristics and the load line. The load line is determined for $v_i = 0$.

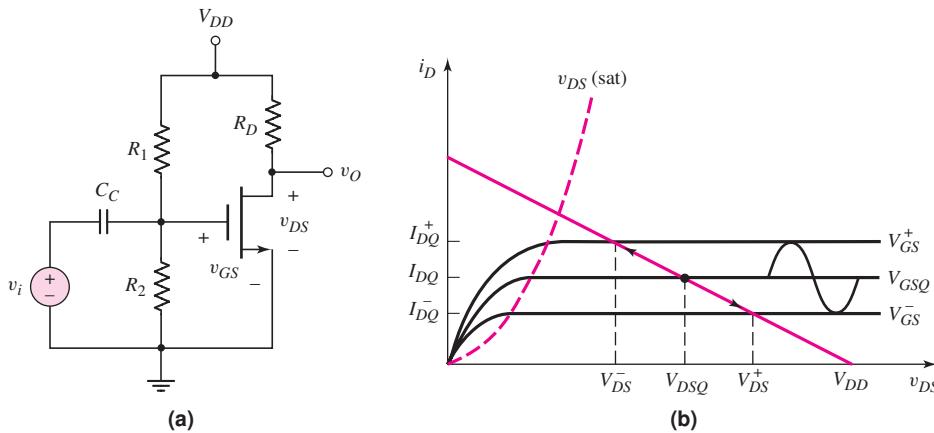


Figure 3.47 (a) An NMOS common-source circuit with a time-varying signal coupled to the gate and (b) transistor characteristics, load line, and superimposed sinusoidal signals

We can establish a particular Q -point on the load line by designing the ratio of the bias resistors R_1 and R_2 . If we assume that $v_i = V_i \sin \omega t$, the gate-to-source voltage will have a sinusoidal signal superimposed on the dc quiescent value. As the gate-to-source voltage changes over time, the Q -point will move up and down the line, as indicated in the figure.

Moving up and down the load line translates into a sinusoidal variation in the drain current and in the drain-to-source voltage. The variation in output voltage can be larger than the input signal voltage, which means the input signal is amplified. The actual signal gain depends on both the transistor parameters and the circuit element values.

In the next chapter, we will develop an equivalent circuit for the transistor used to determine the time-varying small-signal gain and other characteristics of the circuit.

Test Your Understanding

TYU 3.12 The circuit shown in Figure 3.45 is biased at $V_{DD} = 10$ V, and the transistor parameters are $V_{TN} = 0.7$ V and $K_n = 4$ mA/V². Design the value of R_D such that the output voltage will be $v_O = 0.20$ V when $v_I = 10$ V. (Ans. 0.666 k Ω)

TYU 3.13 The transistor in the circuit shown in Figure 3.48 has parameters $K_n = 4$ mA/V² and $V_{TN} = 0.8$ V, and is used to switch the LED on and off. The LED cutin voltage is $V_y = 1.5$ V. The LED is turned on by applying an input voltage of $v_I = 5$ V. (a) Determine the value of R such that the diode current is 12 mA. (b) From the results of part (a), what is the value of v_{DS} ? (Ans. (a) $R = 261 \Omega$, (b) $v_{DS} = 0.374$ V)

TYU 3.14 In the circuit in Figure 3.46, let $R_D = 25$ k Ω and $V_{TN} = 1$ V. (a) Determine the value of the conduction parameter K_n required such that $V_O = 0.10$ V when $V_1 = 0$ and $V_2 = 5$ V. (b) Using the results of part (a), find the value of V_O when $V_1 = V_2 = 5$ V. (Ans. (a) $K_n = 0.248$ mA/V², (b) $V_O = 0.0502$ V)

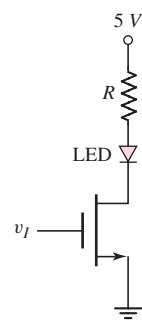


Figure 3.48


3.4

CONSTANT-CURRENT BIASING

Objective: • Investigate current biasing of MOSFET devices.

As was shown in Figure 3.32, a MOSFET can be biased with a constant-current source I_Q . The gate-to-source voltage of the transistor in this circuit then adjusts itself to correspond to the current I_Q .

We can implement the current source by using MOSFET devices. The circuits shown in Figures 3.49(a) and 3.49(b) are a first step toward this design. The transistors M_2 and M_3 in Figure 3.49(a) form a **current mirror** and are used to bias the NMOS transistor M_1 . Similarly, the transistors M_B and M_C in Figure 3.49(b) form a current mirror and are used to bias the PMOS transistor M_A .

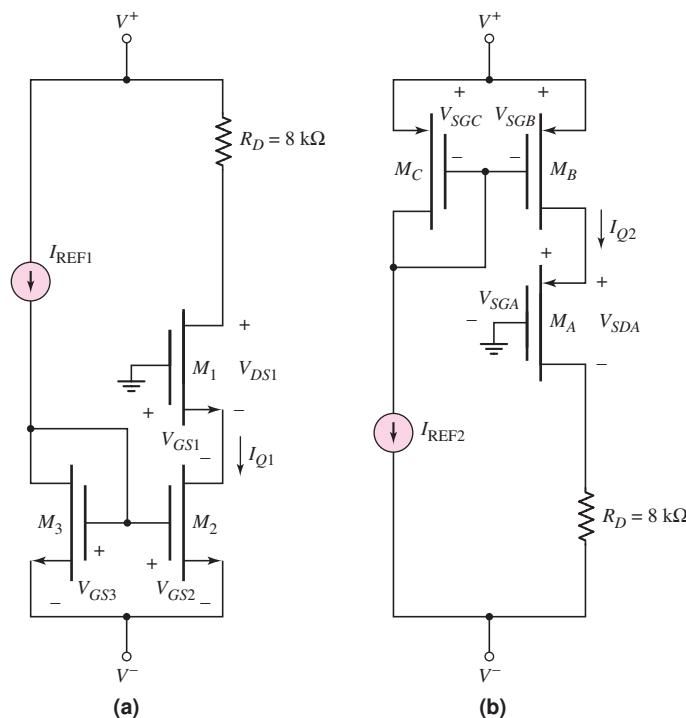


Figure 3.49 (a) NMOS current mirror and (b) PMOS current mirror

The operation and characteristics of these circuits are demonstrated in the following two examples.

EXAMPLE 3.14

Objective: Analyze the circuit shown in Figure 3.49(a). Determine the bias current I_{Q1} , the gate-to-source voltages of the transistors, and the drain-to-source voltage of M_1 .

Assume circuit parameters of $I_{REF1} = 200 \mu\text{A}$, $V^+ = 2.5 \text{ V}$, and $V^- = -2.5 \text{ V}$. Assume transistor parameters of $V_{TN} = 0.4 \text{ V}$ (all transistors), $\lambda = 0$ (all transistors), $K_{n1} = 0.25 \text{ mA/V}^2$, and $K_{n2} = K_{n3} = 0.15 \text{ mA/V}^2$.

Solution: The drain current in M_3 is $I_{D3} = I_{\text{REF1}} = 200 \mu\text{A}$ and is given by the relation $I_{D3} = K_{n3}(V_{GS3} - V_{TN})^2$ (the transistor is biased in the saturation region). Solving for the gate-to-source voltage, we find

$$V_{GS3} = \sqrt{\frac{I_{D3}}{K_{n3}}} + V_{TN} = \sqrt{\frac{0.2}{0.15}} + 0.4$$

or

$$V_{GS3} = 1.555 \text{ V}$$

We note that $V_{GS3} = V_{GS2} = 1.555 \text{ V}$. We can write

$$I_{D2} = I_{Q1} = K_{n2}(V_{GS2} - V_{TN})^2 = 0.15(1.555 - 0.4)^2$$

or

$$I_{Q1} = 200 \mu\text{A}$$

The gate-to-source voltage V_{GS1} (assuming M_1 is biased in the saturation region) can be written as

$$V_{GS1} = \sqrt{\frac{I_{Q1}}{K_{n1}}} + V_{TN} = \sqrt{\frac{0.2}{0.25}} + 0.4$$

or

$$V_{GS1} = 1.29 \text{ V}$$

The drain-to-source voltage is found from

$$\begin{aligned} V_{DS1} &= V^+ - I_{Q1}R_D - (-V_{GS1}) \\ &= 2.5 - (0.2)(8) - (-1.29) \end{aligned}$$

or

$$V_{DS1} = 2.19 \text{ V}$$

We may note that M_1 is indeed biased in the saturation region.

Comment: Since the current mirror transistors M_2 and M_3 are matched (identical parameters) and since the gate-to-source voltages are the same in the two transistors, the bias current, I_{Q1} , is equal to (i.e., mirrors) the reference current, I_{REF1} .

EXERCISE PROBLEM

Ex 3.14: For the circuit shown in Figure 3.49(a), assume circuit parameters of $I_{\text{REF1}} = 120 \mu\text{A}$, $V^+ = 3 \text{ V}$, and $V^- = -3 \text{ V}$; and assume transistor parameters of $V_{TN} = 0.4 \text{ V}$, $\lambda = 0$, $K_{n1} = 50 \mu\text{A/V}^2$, $K_{n2} = 30 \mu\text{A/V}^2$, and $K_{n3} = 60 \mu\text{A/V}^2$. Determine I_{Q1} and all gate-to source voltages. (Ans. $I_{Q1} = 60 \mu\text{A}$, $V_{GS1} = 1.495 \text{ V}$, $V_{GS2} = V_{GS3} = 1.814 \text{ V}$)

We will now consider a current mirror in which the bias current and reference current are not equal.

EXAMPLE 3.15

Objective: Design the circuit shown in Figure 3.49(b) to provide a bias current of $I_{Q2} = 150 \mu\text{A}$.

Assume circuit parameters of $I_{REF2} = 250 \mu\text{A}$, $V^+ = 3 \text{ V}$, and $V^- = -3 \text{ V}$. Assume transistor parameters of $V_{TP} = -0.6 \text{ V}$ (all transistors), $\lambda = 0$ (all transistors), $k'_p = 40 \mu\text{A/V}^2$ (all transistors), $W/L_C = 15$, and $W/L_A = 25$.

Solution: Since the bias current I_{Q2} and reference current I_{REF2} are not equal, the W/L ratios of the current mirror transistors, M_C and M_B , will not be the same.

For M_C , since the transistor is biased in the saturation region, we have

$$I_{DC} = I_{REF2} = \frac{k'_p}{2} \cdot \left(\frac{W}{L}\right)_C (V_{SGC} + V_{TP})^2$$

or

$$250 = \frac{40}{2}(15)[V_{SGC} + (-0.6)]^2 = 300(V_{SGC} - 0.6)^2$$

Then

$$V_{SGC} = \sqrt{\frac{250}{300}} + 0.6$$

or

$$V_{SGC} = 1.513 \text{ V}$$

Since $V_{SGC} = V_{SGB} = 1.513 \text{ V}$, we obtain

$$I_B = I_{Q2} = \frac{k'_p}{2} \cdot \left(\frac{W}{L}\right)_B (V_{SGB} + V_{TP})^2$$

or

$$150 = \frac{40}{2} \cdot \left(\frac{W}{L}\right)_B [1.513 + (-0.6)]^2$$

We find

$$\left(\frac{W}{L}\right)_B = 9$$

For M_A , we have

$$I_{DA} = I_{Q2} = \frac{k'_p}{2} \cdot \left(\frac{W}{L}\right)_A (V_{SGA} + V_{TP})^2$$

or

$$150 = \frac{40}{2}(25)(V_{SGA} + (-0.6))^2 = 500(V_{SGA} - 0.6)^2$$

Now

$$V_{SGA} = \sqrt{\frac{150}{500}} + 0.6$$

or

$$V_{SGA} = 1.148 \text{ V}$$

The source-to-drain voltage of M_A is found from

$$V_{SDA} = V_{SGA} - I_{Q2}R_D - V^- = 1.148 - (0.15)(8) - (-3)$$

or

$$V_{SDA} = 2.95 \text{ V}$$

We may note that the transistor M_A is biased in the saturation region.

Comment: By designing the W/L ratios of the current mirror transistors, we can obtain different reference current and bias current values.

EXERCISE PROBLEM

Ex 3.15: Consider the circuit shown in Figure 3.49(b). Assume circuit parameters of $I_{REF2} = 0.1$ mA, $V^+ = 5$ V, and $V^- = -5$ V. The transistor parameters are the same as given in Example 3.15. Design the circuit such that $I_{Q2} = 0.2$ mA. Also determine all source-to-gate voltages. (Ans. $V_{SGC} = V_{SGB} = 1.18$ V, $(W/L)_B = 30$, $V_{SGA} = 1.23$ V)

The constant-current source can be implemented by using MOSFETs as shown in Figure 3.50. The transistors M_2 , M_3 , and M_4 form the current source. Transistors M_3 and M_4 are each connected in a diode-type configuration, and they establish a reference current. We noted in the last section that this diode-type connection implies the transistor is always biased in the saturation region. Transistors M_3 and M_4 are therefore biased in the saturation region, and M_2 is assumed to be biased in the saturation region. The resulting gate-to-source voltage on M_3 is applied to M_2 , and this establishes the bias current I_O .

Since the reference current is the same in transistors M_3 and M_4 , we can write

$$K_{n3}(V_{GS3} - V_{TN3})^2 = K_{n4}(V_{GS4} - V_{TN4})^2 \quad (3.28)$$

We also know that

$$V_{GS4} + V_{GS3} = (-V^-) \quad (3.29)$$

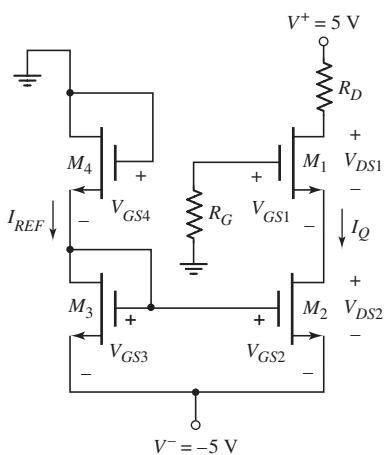


Figure 3.50 Implementation of a MOSFET constant-current source

Solving Equation (3.29) for V_{GS4} and substituting the result into Equation (3.28) yields

$$V_{GS3} = \frac{\sqrt{\frac{K_{n4}}{K_{n3}}}[-V^-] - V_{TN4} + V_{TN3}}{1 + \sqrt{\frac{K_{n4}}{K_{n3}}}} \quad (3.30)$$

Since $V_{GS3} = V_{GS2}$, the bias current is

$$I_Q = K_{n2}(V_{GS3} - V_{TN2})^2 \quad (3.31)$$

EXAMPLE 3.16

Objective: Determine the currents and voltages in a MOSFET constant-current source.

For the circuit shown in Figure 3.50, the transistor parameters are: $K_{n1} = 0.2 \text{ mA/V}^2$, $K_{n2} = K_{n3} = K_{n4} = 0.1 \text{ mA/V}^2$, and $V_{TN1} = V_{TN2} = V_{TN3} = V_{TN4} = 1\text{V}$.

Solution: From Equation (3.30), we can determine V_{GS3} , as follows:

$$V_{GS3} = \frac{\sqrt{\frac{0.1}{0.1}}[5 - 1] + 1}{1 + \sqrt{\frac{0.1}{0.1}}} = 2.5 \text{ V}$$

Since M_3 and M_4 are identical transistors, V_{GS3} should be one-half of the bias voltage. The bias current I_Q is then

$$I_Q = (0.1) \cdot (2.5 - 1)^2 = 0.225 \text{ mA}$$

The gate-to-source voltage on M_1 is found from

$$I_Q = K_{n1}(V_{GS1} - V_{TN1})^2$$

or

$$0.225 = (0.2) \cdot (V_{GS1} - 1)^2$$

which yields

$$V_{GS1} = 2.06 \text{ V}$$

The drain-to-source voltage on M_2 is

$$V_{DS2} = (-V^-) - V_{GS1} = 5 - 2.06 = 2.94 \text{ V}$$

Since $V_{DS2} = 2.94 \text{ V} > V_{DS(\text{sat})} = V_{GS2} - V_{TN2} = 2.5 - 1 = 1.5 \text{ V}$, M_2 is biased in the saturation region.

Design Consideration: Since in this example M_2 and M_3 are identical transistors, the reference current I_{REF} and bias current I_Q are equal. By redesigning the width-to-length ratios of M_2 , M_3 , and M_4 , we can obtain a specific bias current I_Q . If M_2 and

M_3 are not identical, then I_Q and I_{REF} will not be equal. A variety of design options are possible with such a circuit configuration.

EXERCISE PROBLEM

Ex 3.16: Consider the constant-current source shown in Figure 3.50. Assume that the threshold voltage of each transistor is $V_{TN} = 1$ V. (a) Design the ratio of K_{n4}/K_{n3} such that $V_{GS3} = 2$ V. (b) Determine K_{n2} such that $I_Q = 100 \mu\text{A}$. (c) Find K_{n3} and K_{n4} such that $I_{\text{REF}} = 200 \mu\text{A}$. (Ans. (a) $K_{n4}/K_{n3} = \frac{1}{4}$ (b) $K_{n2} = 0.1 \text{ mA/V}^2$ (c) $K_{n3} = 0.2 \text{ mA/V}^2$, $K_{n4} = 0.05 \text{ mA/V}^2$)

Test Your Understanding

TYU 3.15 Consider the circuit in Figure 3.49(b). Assume circuit parameters of $I_{\text{REF2}} = 40 \mu\text{A}$, $V^+ = 2.5$ V, $V^- = -2.5$ V, and $R_D = 20 \text{ k}\Omega$. The transistor parameters are $V_{TP} = -0.30$ V, $K_{pC} = 40 \mu\text{A/V}^2$, $K_{pB} = 60 \mu\text{A/V}^2$, and $K_{pA} = 75 \mu\text{A/V}^2$. Determine I_{Q2} and all source-to-gate voltages. (Ans. $I_{Q2} = 60 \mu\text{A}$, $V_{SGC} = V_{SGB} = 1.30$ V, $V_{SGA} = 1.19$ V)

TYU 3.16 Consider the circuit shown in Figure 3.50. Assume all transistor threshold voltages are 0.7 V. Determine the values of K_{n1} , K_{n2} , K_{n3} , and K_{n4} such that $I_{\text{REF}} = 80 \mu\text{A}$, $I_Q = 120 \mu\text{A}$, $V_{GS3} = 2$ V, and $V_{GS1} = 1.5$ V. (Ans. $K_{n1} = 187.5 \mu\text{A/V}^2$, $K_{n2} = 71.0 \mu\text{A/V}^2$, $K_{n3} = 47.3 \mu\text{A/V}^2$, $K_{n4} = 15.12 \mu\text{A/V}^2$)



3.5

MULTISTAGE MOSFET CIRCUITS

Objective: • Consider the dc biasing of multistage or multitransistor circuits.

In most applications, a single-transistor amplifier will not be able to meet the combined specifications of a given amplification factor, input resistance, and output resistance. For example, the required voltage gain may exceed that which can be obtained in a single-transistor circuit.

Transistor amplifier circuits can be connected in series, or **cascaded**, as shown in Figure 3.51. This may be done either to increase the overall small-signal voltage gain, or provide an overall voltage gain greater than 1, with a very low output resistance.

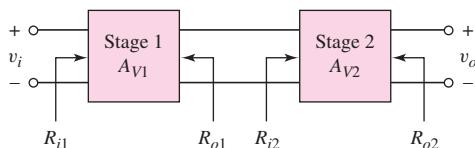


Figure 3.51 Generalized two-stage amplifier

The overall voltage gain may not simply be the product of the individual amplification factors. Loading effects, in general, need to be taken into account.

There are many possible multistage configurations; we will examine a few here, in order to understand the type of analysis required.

3.5.1 Multitransistor Circuit: Cascade Configuration

The circuit shown in Figure 3.52 is a cascade of a common-source amplifier followed by a source-follower amplifier. We will show in the next chapter that the common-source amplifier provides a small-signal voltage gain and the source-follower has a low output impedance.

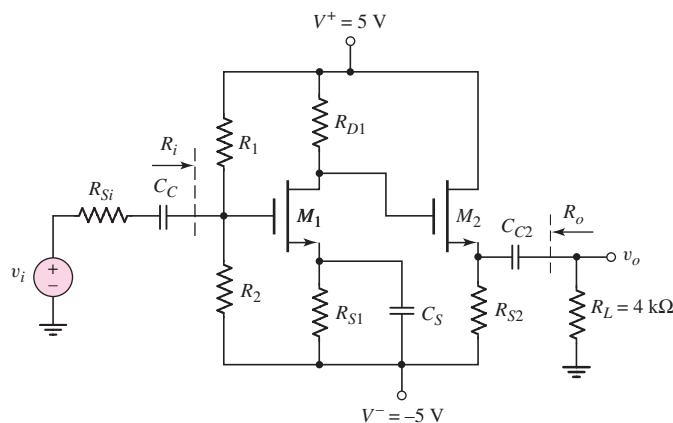


Figure 3.52 Common-source amplifier in cascade with source follower

DESIGN EXAMPLE 3.17

Objective: Design the biasing of a multistage MOSFET circuit to meet specific requirements.

Consider the circuit shown in Figure 3.52 with transistor parameters $K_{n1} = 500 \mu\text{A/V}^2$, $K_{n2} = 200 \mu\text{A/V}^2$, $V_{TN1} = V_{TN2} = 1.2 \text{ V}$, and $\lambda_1 = \lambda_2 = 0$. Design the circuit such that $I_{DQ1} = 0.2 \text{ mA}$, $I_{DQ2} = 0.5 \text{ mA}$, $V_{DSQ1} = V_{DSQ2} = 6 \text{ V}$, and $R_i = 100 \text{ k}\Omega$. Let $R_{Si} = 4 \text{ k}\Omega$.

Solution: For output transistor M_2 , we have

$$V_{DSQ2} = 5 - (-5) - I_{DQ2}R_{S2}$$

or

$$6 = 10 - (0.5)R_{S2}$$

which yields $R_{S2} = 8 \text{ k}\Omega$. Also, assuming transistors are biased in the saturation region,

$$I_{DQ2} = K_{n2}(K_{GS2} - V_{TN2})^2$$

or

$$0.5 = 0.2(V_{GS2} - 1.2)^2$$

which yields

$$V_{GS2} = 2.78 \text{ V}$$

Since $V_{DSQ2} = 6 \text{ V}$, the source voltage of M_2 is $V_{S2} = -1 \text{ V}$. With $V_{GS2} = 2.78 \text{ V}$, the gate voltage on M_2 must be

$$V_{G2} = -1 + 2.78 = 1.78 \text{ V}$$

The resistor R_{D1} is then

$$R_{D1} = \frac{5 - 1.78}{0.2} = 16.1 \text{ k}\Omega$$

For $V_{DSQ1} = 6 \text{ V}$, the source voltage of M_1 is

$$V_{S1} = 1.78 - 6 = -4.22 \text{ V}$$

The resistor R_{S1} is then

$$R_{S1} = \frac{-4.22 - (-5)}{0.2} = 3.9 \text{ k}\Omega$$

For transistor M_1 , we have

$$I_{DQ1} = K_{n1}(V_{GS1} - V_{TN1})^2$$

or

$$0.2 = 0.50(V_{GS1} - 1.2)^2$$

which yields

$$V_{GS1} = 1.83 \text{ V}$$

To find R_1 and R_2 , we can write

$$V_{GS1} = \left(\frac{R_2}{R_1 + R_2} \right) (10) - I_{DQ1} R_{S1}$$

Since

$$\frac{R_2}{R_1 + R_2} = \frac{1}{R_1} \cdot \left(\frac{R_1 R_2}{R_1 + R_2} \right) = \frac{1}{R_1} \cdot R_i$$

then, since the input resistance is specified to be $100 \text{ k}\Omega$, we have

$$1.83 = \frac{1}{R_1} (100)(10) - (0.2)(3.9)$$

which yields $R_1 = 383 \text{ k}\Omega$. From $R_i = 100 \text{ k}\Omega$, we find that $R_2 = 135 \text{ k}\Omega$.

Comment: Both transistors are biased in the saturation region, as assumed, which is desired for linear amplifiers as we will see in the next chapter.

EXERCISE PROBLEM

Ex 3.17: The transistor parameters for the circuit shown in Figure 3.52 are the same as described in Example 3.17. Design the circuit such that $I_{DQ1} = 0.1 \text{ mA}$, $I_{DQ2} = 0.3 \text{ mA}$, $V_{DSQ1} = V_{DSQ2} = 5 \text{ V}$, and $R_i = 200 \text{ k}\Omega$. (Ans. $R_{S2} = 16.7 \text{ k}\Omega$, $R_{D1} = 25.8 \text{ k}\Omega$, $R_{S1} = 24.3 \text{ k}\Omega$, $R_1 = 491 \text{ k}\Omega$, and $R_2 = 337 \text{ k}\Omega$)

3.5.2 Multitransistor Circuit: Cascode Configuration

Figure 3.53 shows a **cascode** circuit with n-channel MOSFETs. Transistor M_1 is connected in a common-source configuration and M_2 is connected in a common-gate configuration. The advantage of this type of circuit is a higher frequency response, which is discussed in a later chapter.

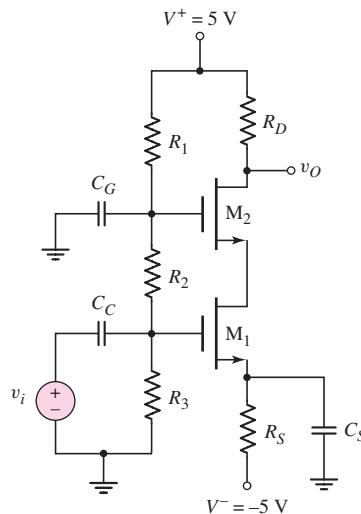


Figure 3.53 NMOS cascode circuit

DESIGN EXAMPLE 3.18

Objective: Design the biasing of the cascode circuit to meet specific requirements.

For the circuit shown in Figure 3.53, the transistor parameters are: $V_{TN1} = V_{TN2} = 1.2$ V, $K_{n1} = K_{n2} = 0.8$ mA/V², and $\lambda_1 = \lambda_2 = 0$. Let $R_1 + R_2 + R_3 = 300$ k Ω and $R_S = 10$ k Ω . Design the circuit such that $I_{DQ} = 0.4$ mA and $V_{DSQ1} = V_{DSQ2} = 2.5$ V.

Solution: The dc voltage at the source of M_1 is

$$V_{S1} = I_{DQ} R_S - 5 = (0.4)(10) - 5 = -1 \text{ V}$$

Since M_1 and M_2 are identical transistors, and since the same current exists in the two transistors, the gate-to-source voltage is the same for both devices. We have

$$I_D = K_n(V_{GS} - V_{TN})^2$$

or

$$0.4 = 0.8(V_{GS} - 1.2)^2$$

which yields

$$V_{GS} = 1.907 \text{ V}$$

Then,

$$V_{G1} = \left(\frac{R_3}{R_1 + R_2 + R_3} \right) (5) = V_{GS} + V_{S1}$$

or

$$\left(\frac{R_3}{300}\right)(5) = 1.907 - 1 = 0.907$$

which yields

$$R_3 = 54.4 \text{ k}\Omega$$

The voltage at the source of M_2 is

$$V_{S2} = V_{DSQ1} + V_{S1} = 2.5 - 1 = 1.5 \text{ V}$$

Then,

$$V_{G2} = \left(\frac{R_2 + R_3}{R_1 + R_2 + R_3}\right)(5) = V_{GS} + V_{S2}$$

or

$$\left(\frac{R_2 + R_3}{300}\right)(5) = 1.907 + 1.5 = 3.407 \text{ V}$$

which yields

$$R_2 + R_3 = 204.4 \text{ k}\Omega$$

and

$$R_2 = 150 \text{ k}\Omega$$

Therefore

$$R_1 = 95.6 \text{ k}\Omega$$

The voltage at the drain of M_2 is

$$V_{D2} = V_{DSQ2} + V_{S2} = 2.5 + 1.5 = 4 \text{ V}$$

The drain resistor is therefore

$$R_D = \frac{5 - V_{D2}}{I_{DQ}} = \frac{5 - 4}{0.4} = 2.5 \text{ k}\Omega$$

Comment: Since $V_{DS} = 2.5 \text{ V} > V_{GS} - V_{TN} = 1.91 - 1.2 = 0.71 \text{ V}$, each transistor is biased in the saturation region.

EXERCISE PROBLEM

Ex 3.18: The transistor parameters for the circuit shown in Figure 3.53 are $V_{TN1} = V_{TN2} = 0.8 \text{ V}$, $K_{n1} = K_{n2} = 0.5 \text{ mA/V}^2$, and $\lambda_1 = \lambda_2 = 0$. Let $R_1 + R_2 + R_3 = 500 \text{ k}\Omega$ and $R_S = 16 \text{ k}\Omega$. Design the circuit such that $I_{DQ} = 0.25 \text{ mA}$ and $V_{DSQ1} = V_{DSQ2} = 2.5 \text{ V}$. (Ans. $R_3 = 50.7 \text{ k}\Omega$, $R_2 = 250 \text{ k}\Omega$, $R_1 = 199.3 \text{ k}\Omega$, $R_D = 4 \text{ k}\Omega$)

We will encounter many more examples of multitransistor and multistage amplifiers in later chapters of this text. Specifically in Chapter 11, we will consider the differential amplifier and in Chapter 13, we will analyze circuits that form the operational amplifier.


3.6

JUNCTION FIELD-EFFECT TRANSISTOR

Objective: • Understand the operation and characteristics of the pn junction FET (JFET) and the Schottky barrier junction FET (MESFET), and understand the dc analysis techniques of JFET and MESFET circuits.

The two general categories of **junction field-effect transistor (JFET)** are the pn junction FET, or **pn JFET**, and the **metal-semiconductor field-effect transistor (MESFET)**, which is fabricated with a Schottky barrier junction.

The current in a JFET is through a semiconductor region known as the channel, with ohmic contacts at each end. The basic transistor action is the modulation of the channel conductance by an electric field perpendicular to the channel. Since the modulating electric field is induced in the space-charge region of a reverse-biased pn junction or Schottky barrier junction, the field is a function of the gate voltage. Modulation of the channel conductance by the gate voltage modulates the channel current.

JFETs were developed before MOSFETs, but the applications and uses of the MOSFET have far surpassed those of the JFET. One reason is that the voltages applied to the gate and drain of a MOSFET are the same polarity (both positive or both negative), whereas the voltages applied to the gate and drain of most JFETs must have opposite polarities. Since the JFET is used only in specialized applications, our discussion will be brief.

3.6.1 pn JFET and MESFET Operation

pn JFET

A simplified cross section of a symmetrical pn JFET is shown in Figure 3.54. In the n-region channel between the two p-regions, majority carrier electrons flow from the source to the drain terminal; thus, the JFET is called a majority-carrier device. The two gate terminals shown in Figure 3.54 are connected to form a single gate.

In a p-channel JFET, the p- and n-regions are reversed from those of the n-channel device, and holes flow in the channel from the source to the drain. The current direction and voltage polarities in the p-channel JFET are reversed from

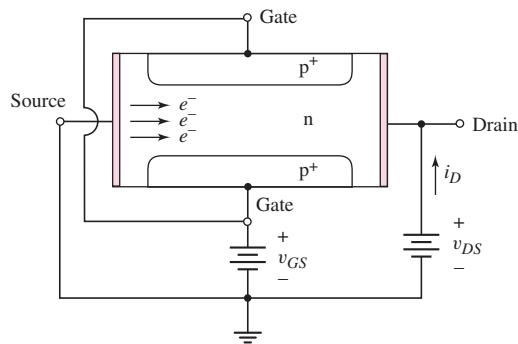


Figure 3.54 Cross section of a symmetrical n-channel pn junction field-effect transistor

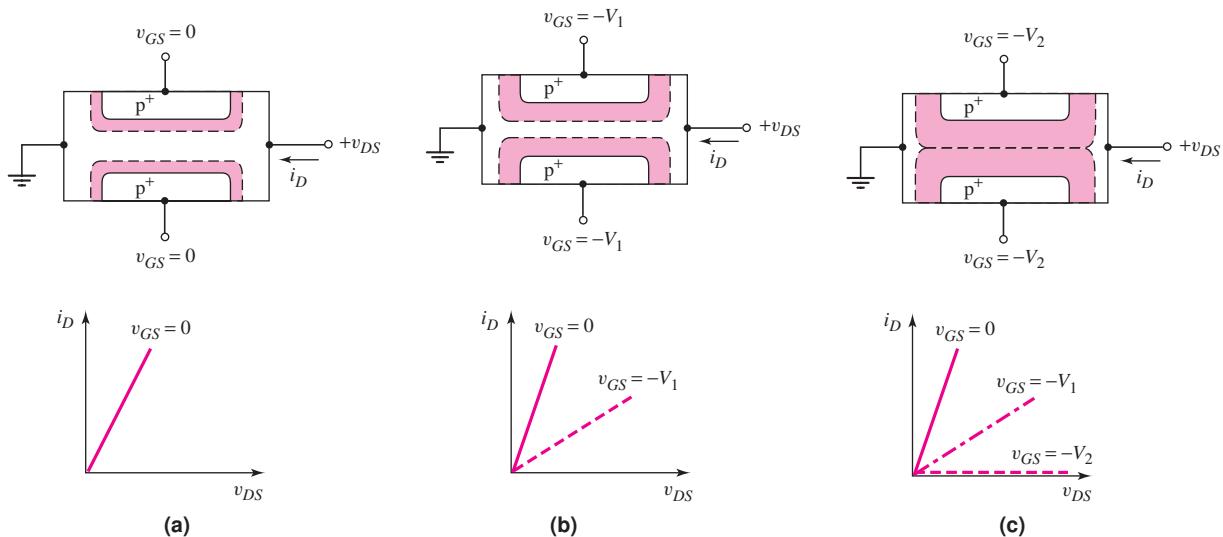


Figure 3.55 Gate-to-channel space-charge regions and current–voltage characteristics for small drain-to-source voltages and for: (a) zero gate voltage, (b) small reverse-biased gate voltage, and (c) a gate voltage that achieves pinchoff

those in the n-channel device. Also, the p-channel JFET is generally a lower-frequency device than the n-channel JFET, because hole mobility is lower than electron mobility.

Figure 3.55(a) shows an n-channel JFET with zero volts applied to the gate. If the source is at ground potential, and if a small positive drain voltage is applied, a drain current i_D is produced between the source and drain terminals. Since the n-channel acts essentially as a resistance, the i_D versus v_{DS} characteristic for small v_{DS} values is approximately linear, as shown in the figure.

If a voltage is applied to the gate of a pn JFET, the channel conductance changes. If a negative gate voltage is applied to the n-channel pn JFET in Figure 3.55, the gate-to-channel pn junction becomes reverse biased. The space-charge region widens, the channel region narrows, the resistance of the n-channel increases, and the slope of the i_D versus v_{DS} curve, for small v_{DS} , decreases. These effects are shown in Figure 3.55(b). If a larger negative gate voltage is applied, the condition shown in Figure 3.55(c) can be achieved. The reverse-biased gate-to-channel space-charge region completely fills the channel region. This condition is known as **pinchoff**. Since the depletion region isolates the source and drain terminals, the drain current at pinchoff is essentially zero. The i_D versus v_{DS} curves are shown in Figure 3.55(c). The current in the channel is controlled by the gate voltage. The control of the current in one part of the device by a voltage in another part of the device is the basic transistor action. The pn JFET is a “normally on,” or depletion mode, device; that is, a voltage must be applied to the gate terminal to turn the device off.

Consider the situation in which the gate voltage is zero, $v_{GS} = 0$, and the drain voltage changes, as shown in Figure 3.56(a). As the drain voltage increases (positive), the gate-to-channel pn junction becomes reverse biased near the drain terminal, and the space-charge region widens, extending farther into the channel. The channel acts essentially as a resistor, and the effective channel resistance increases as the space-charge region widens; therefore, the slope of the i_D versus v_{DS} characteristic

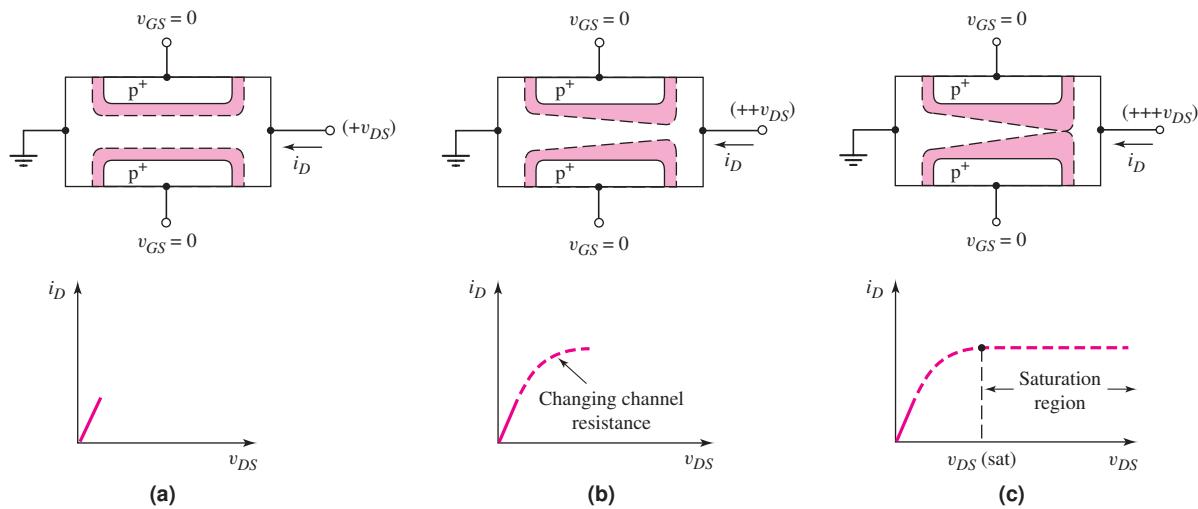


Figure 3.56 Gate-to-channel space-charge regions and current–voltage characteristics for zero gate voltage and for: (a) a small drain voltage, (b) a larger drain voltage, and (c) a drain voltage that achieves pinchoff at the drain terminal

decreases, as shown in Figure 3.56(b). The effective channel resistance now varies along the channel length, and, since the channel current must be constant, the voltage drop through the channel becomes dependent on position.

If the drain voltage increases further, the condition shown in Figure 3.56(c) can result. The channel is pinched off at the drain terminal. Any further increase in drain voltage will not increase the drain current. The $i_D - v_{DS}$ characteristic for this condition is also shown in the figure. The drain voltage at pinchoff is $v_{DS}(\text{sat})$. Therefore, for $v_{DS} > v_{DS}(\text{sat})$, the transistor is biased in the saturation region, and the drain current for this ideal case is independent of v_{DS} .

MESFET

In the MESFET, the gate junction is a Schottky barrier junction, instead of a pn junction. Although MESFETs can be fabricated in silicon, they are usually associated with gallium arsenide or other compound-semiconductor materials.

A simplified cross section of a GaAs MESFET is shown in Figure 3.57. A thin, epitaxial layer of GaAs is used for the active region; the substrate is a very high resistivity GaAs material, referred to as a semi-insulating substrate. The advantages of

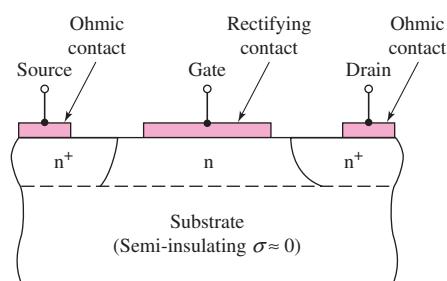


Figure 3.57 Cross section of an n-channel MESFET with a semi-insulating substrate

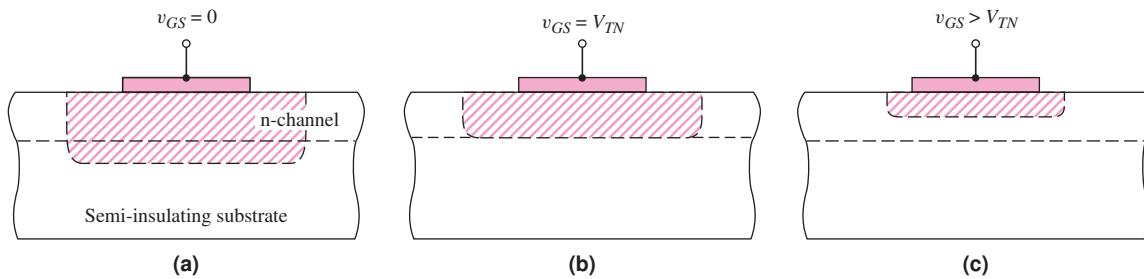


Figure 3.58 Channel space-charge region of an enhancement-mode MESFET for:
(a) $v_{GS} = 0$, (b) $v_{GS} = V_{TN}$, and (c) $v_{GS} > V_{TN}$

these devices include: higher electron mobility in GaAs, hence smaller transit time and faster response; and decreased parasitic capacitance and a simplified fabrication process, resulting from the semi-insulating GaAs substrate.

In the MESFET in Figure 3.57, a reverse-bias gate-to-source voltage induces a space-charge region under the metal gate, which modulates the channel conductance, as in the case of the pn JFET. If a negative applied gate voltage is sufficiently large, the space-charge region will eventually reach the substrate. Again, pinchoff will occur. Also, the device shown in the figure is a depletion mode device, since a gate voltage must be applied to pinch off the channel, that is, to turn the device off.

In another type of MESFET, the channel is pinched off even at $v_{GS} = 0$, as shown in Figure 3.58(a). For this MESFET, the channel thickness is smaller than the zero-biased space-charge width. To open a channel, the depletion region must be reduced; that is, a forward-biased voltage must be applied to the gate–semiconductor junction. When a slightly forward-bias voltage is applied, the depletion region extends just to the width of the channel as shown in Figure 3.58(b). The threshold voltage is the gate-to-source voltage required to create the pinchoff condition. The threshold voltage for this n-channel MESFET is positive, in contrast to the negative threshold voltage of the n-channel depletion-mode device. If a larger forward-bias voltage is applied, the channel region opens, as shown in Figure 3.58(c). The applied forward-bias gate voltage is limited to a few tenths of a volt before a significant gate current occurs.

This device is an **n-channel enhancement-mode MESFET**. Enhancement-mode p-channel MESFETs and enhancement-mode pn JFETs have also been fabricated. The advantage of enhancement-mode MESFETs is that circuits can be designed in which the voltage polarities on the gate and drain are the same. However, the output voltage swing of these devices is quite small.

3.6.2 Current-Voltage Characteristics

The circuit symbols for the n-channel and p-channel JFETs are shown in Figure 3.59, along with the gate-to-source voltages and current directions. The ideal current-voltage characteristics, when the transistor is biased in the saturation region, are described by

$$i_D = I_{DSS} \left(1 - \frac{v_{GS}}{V_P} \right)^2 \quad (3.32)$$

where I_{DSS} is the saturation current when $v_{GS} = 0$, and V_P is the **pinchoff voltage**.

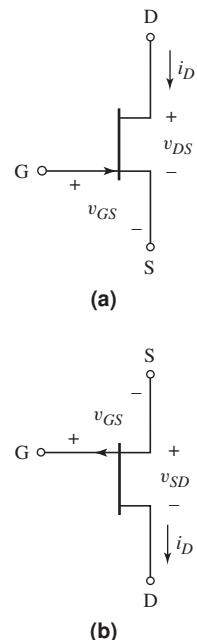


Figure 3.59 Circuit symbols for: (a) n-channel JFET and (b) p-channel JFET

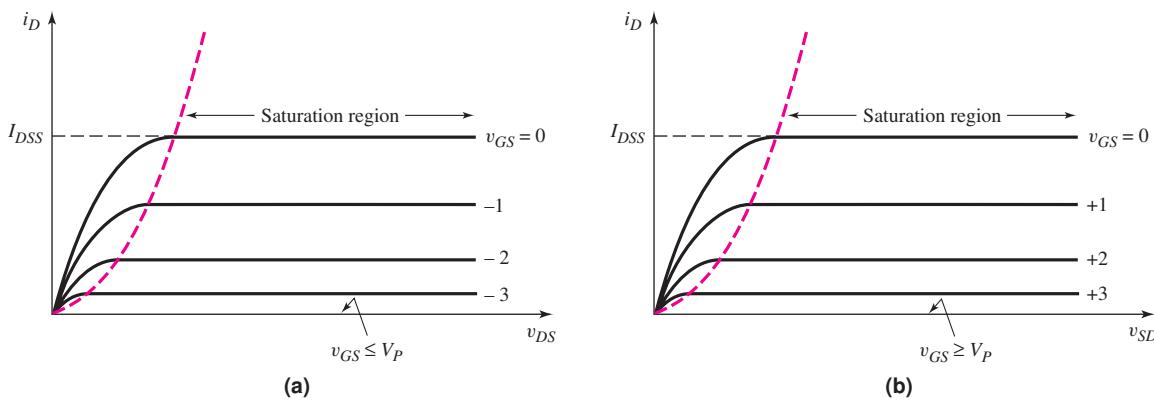


Figure 3.60 Current–voltage characteristics for: (a) n-channel JFET and (b) p-channel JFET

The current–voltage characteristics for n-channel and p-channel JFETs are shown in Figures 3.60(a) and 3.60(b), respectively. Note that the pinchoff voltage V_P for the n-channel JFET is negative and the gate-to-source voltage v_{GS} is usually negative; therefore, the ratio v_{GS}/V_P is positive. Similarly, the pinchoff voltage V_P for the p-channel JFET is positive and the gate-to-source voltage v_{GS} must be positive, and therefore the ratio v_{GS}/V_P is positive.

For the n-channel device, the saturation region occurs when $v_{DS} \geq v_{DS}(\text{sat})$ where

$$v_{DS}(\text{sat}) = v_{GS} - V_P \quad (3.33)$$

For the p-channel device, the saturation region occurs when $v_{SD} \geq v_{SD}(\text{sat})$ where

$$v_{SD}(\text{sat}) = V_P - v_{GS} \quad (3.34)$$

EXAMPLE 3.19

Objective: Calculate i_D and $v_{DS}(\text{sat})$ in an n-channel pn JFET.

Assume the saturation current is $I_{DSS} = 2 \text{ mA}$ and the pinchoff voltage is $V_P = -3.5 \text{ V}$. Calculate i_D and $v_{DS}(\text{sat})$ for $v_{GS} = 0, V_P/4$, and $V_P/2$.

Solution: From Equation (3.32), we have

$$i_D = I_{DSS} \left(1 - \frac{v_{GS}}{V_P}\right)^2 = (2) \left(1 - \frac{v_{GS}}{(-3.5)}\right)^2$$

Therefore, for $v_{GS} = 0, V_P/4$, and $V_P/2$, we obtain

$$i_D = 2, 1.13, \text{ and } 0.5 \text{ mA}$$

From Equation (3.33), we have

$$v_{DS}(\text{sat}) = v_{GS} - V_P = v_{GS} - (-3.5)$$

Therefore, for $v_{GS} = 0, V_P/4$, and $V_P/2$, we obtain

$$v_{DS}(\text{sat}) = 3.5, 2.63, \text{ and } 1.75 \text{ V}$$

Comment: The current capability of a JFET can be increased by increasing the value of I_{DSS} , which is a function of the transistor width.

EXERCISE PROBLEM

Ex 3.19: The parameters of an n-channel JFET are $I_{DSS} = 12 \text{ mA}$, $V_P = -4.5 \text{ V}$, and $\lambda = 0$. Determine $V_{DS}(\text{sat})$ for $V_{GS} = -1.2 \text{ V}$, and calculate I_D for $V_{DS} > V_{DS}(\text{sat})$. (Ans. $V_{DS}(\text{sat}) = 3.3 \text{ V}$, $I_D = 6.45 \text{ mA}$)

As in the case of the MOSFET, the i_D versus v_{DS} characteristic for the JFET may have a nonzero slope beyond the saturation point. This nonzero slope can be described through the following equation:

$$i_D = I_{DSS} \left(1 - \frac{v_{GS}}{V_P} \right)^2 (1 + \lambda v_{DS}) \quad (3.35)$$

The output resistance r_o is defined as

$$r_o = \left. \left(\frac{\partial i_D}{\partial v_{DS}} \right) \right|_{v_{GS}=\text{const.}}^{-1} \quad (3.36)$$

Using Equation (3.35), we find that

$$r_o = \left[\lambda I_{DSS} \left(1 - \frac{V_{GSQ}}{V_P} \right)^2 \right]^{-1} \quad (3.37(a))$$

or

$$r_o \cong [\lambda I_{DQ}]^{-1} = \frac{1}{\lambda I_{DQ}} \quad (3.37(b))$$

The output resistance will be considered again when we discuss the small-signal equivalent circuit of the JFET in the next chapter.

Enhancement-mode GaAs MESFETs can be fabricated with current–voltage characteristics much like those of the enhancement-mode MOSFET. Therefore, for the ideal enhancement-mode MESFET biased in the saturation region, we can write

$$i_D = K_n (v_{GS} - V_{TN})^2 \quad (3.38(a))$$

For the ideal enhancement-mode MESFET biased in the nonsaturation region,

$$i_D = K_n [2(v_{GS} - V_{TN})v_{DS} - v_{DS}^2] \quad (3.38(b))$$

where K_n is the conduction parameter and V_{TN} is the threshold voltage, which in this case is equivalent to the pinchoff voltage. For an n-channel enhancement-mode MESFET, the threshold voltage is positive.

3.6.3 Common JFET Configurations: dc Analysis

There are several common JFET circuit configurations. We will look at a few of these, using examples, and will illustrate the dc analysis and design of such circuits.

DESIGN EXAMPLE 3.20

Objective: Design a JFET circuit with a voltage divider biasing circuit.

Consider the circuit shown in Figure 3.61(a) with transistor parameters $I_{DSS} = 12 \text{ mA}$, $V_P = -3.5 \text{ V}$, and $\lambda = 0$. Let $R_1 + R_2 = 100 \text{ k}\Omega$. Design the circuit such that the dc drain current is $I_D = 5 \text{ mA}$ and the dc drain-to-source voltage is $V_{DS} = 5 \text{ V}$.

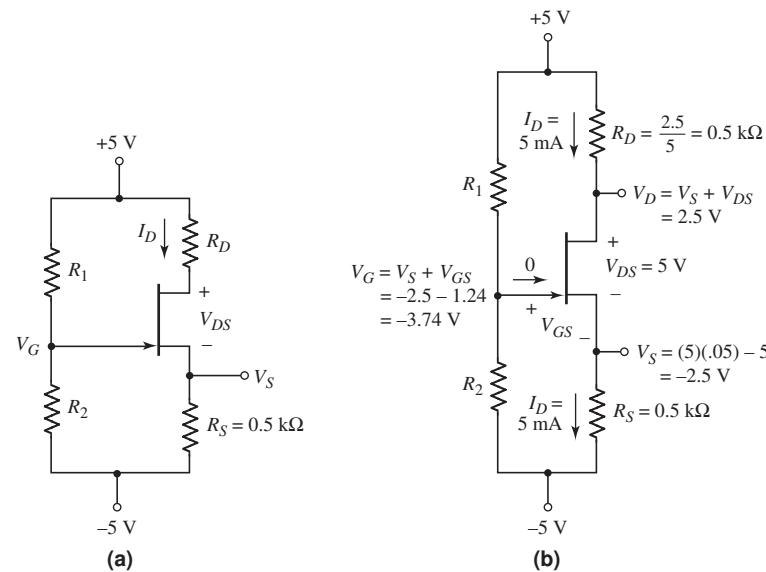


Figure 3.61 (a) An n-channel JFET circuit with voltage divider biasing and
(b) the n-channel JFET circuit for Example 3.20

Solution: Assume the transistor is biased in the saturation region. The dc drain current is then given by

$$I_D = I_{DSS} \left(1 - \frac{V_{GS}}{V_P} \right)^2$$

Therefore,

$$5 = 12 \left(1 - \frac{V_{GS}}{(-3.5)} \right)^2$$

which yields

$$V_{GS} = -1.24 \text{ V}$$

From Figure 3.61(b), the voltage at the source terminal is

$$V_S = I_D R_S - 5 = (5)(0.5) - 5 = -2.5 \text{ V}$$

which means that the gate voltage is

$$V_G = V_{GS} + V_S = -1.24 - 2.5 = -3.74 \text{ V}$$

We can also write the gate voltage as

$$V_G = \left(\frac{R_2}{R_1 + R_2} \right) (10) - 5$$

or

$$-3.74 = \frac{R_2}{100} (10) - 5.$$

Therefore,

$$R_2 = 12.6 \text{ k}\Omega$$

and

$$R_1 = 87.4 \text{ k}\Omega$$

The drain-to-source voltage is

$$V_{DS} = 5 - I_D R_D - I_D R_S - (-5)$$

Therefore,

$$R_D = \frac{10 - V_{DS} - I_D R_S}{I_D} = \frac{10 - 5 - (5)(0.5)}{5} = 0.5 \text{ k}\Omega$$

We also see that

$$V_{DS} = 5 \text{ V} > V_{GS} - V_P = -1.24 - (-3.5) = 2.26 \text{ V}$$

which shows that the JFET is indeed biased in the saturation region, as initially assumed.

Comment: The dc analysis of the JFET circuit is essentially the same as that of the MOSFET circuit, since the gate current is assumed to be zero.

EXERCISE PROBLEM

Ex 3.20: The transistor in the circuit in Figure 3.62 has parameters $I_{DSS} = 6 \text{ mA}$, $V_P = -4 \text{ V}$, and $\lambda = 0$. Design the circuit such that $I_{DQ} = 2.5 \text{ mA}$ and $V_{DS} = 6 \text{ V}$, and the total power dissipated in R_1 and R_2 is 2 mW. (Ans. $R_D = 1.35 \text{ k}\Omega$, $R_1 = 158 \text{ k}\Omega$, $R_2 = 42 \text{ k}\Omega$)

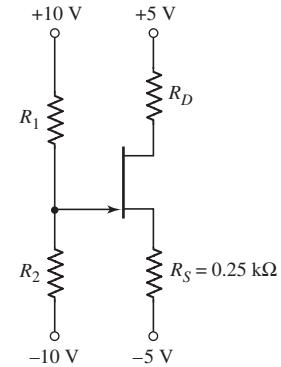


Figure 3.62 Circuit for Exercise Ex 3.20

EXAMPLE 3.21

Objective: Calculate the quiescent current and voltage values in a p-channel JFET circuit.

The parameters of the transistor in the circuit shown in Figure 3.63 are: $I_{DSS} = 2.5 \text{ mA}$, $V_P = +2.5 \text{ V}$, and $\lambda = 0$. The transistor is biased with a constant-current source.

Solution: From Figure 3.63, we can write the dc drain current as

$$I_D = I_Q = 0.8 \text{ mA} = \frac{V_D - (-9)}{R_D}$$

which yields

$$V_D = (0.8)(4) - 9 = -5.8 \text{ V}$$

Now, assume the transistor is biased in the saturation region. We then have

$$I_D = I_{DSS} \left(1 - \frac{V_{GS}}{V_P}\right)^2$$

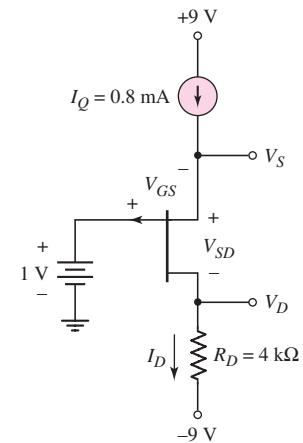


Figure 3.63 A p-channel JFET circuit biased with a constant-current source

or

$$0.8 = 2.5 \left(1 - \frac{V_{GS}}{2.5}\right)^2$$

which yields

$$V_{GS} = 1.086 \text{ V}$$

Then

$$V_S = 1 - V_{GS} = 1 - 1.086 = -0.086 \text{ V}$$

and

$$V_{SD} = V_S - V_D = -0.086 - (-5.8) = 5.71 \text{ V}$$

Again, we see that

$$V_{SD} = 5.71 \text{ V} > V_P - V_{GS} = 2.5 - 1.086 = 1.41 \text{ V}$$

which verifies that the transistor is biased in the saturation region, as assumed.

Comment: In the same way as bipolar or MOS transistors, junction field-effect transistors can be biased using constant-current sources.

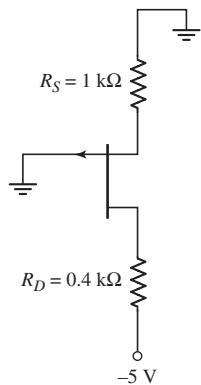


Figure 3.64 Circuit for Exercise Ex 3.21

EXERCISE PROBLEM

Ex 3.21: For the p-channel transistor in the circuit in Figure 3.64, the parameters are: $I_{DSS} = 6 \text{ mA}$, $V_P = 4 \text{ V}$, and $\lambda = 0$. Calculate the quiescent values of I_D , V_{GS} , and V_{SD} . Is the transistor biased in the saturation or nonsaturation region? (Ans. $V_{GS} = 1.81 \text{ V}$, $I_D = 1.81 \text{ mA}$, $V_{SD} = 2.47 \text{ V}$, saturation region)

DESIGN EXAMPLE 3.22

Objective: Design a circuit with an enhancement-mode MESFET.

Consider the circuit shown in Figure 3.65(a). The transistor parameters are: $V_{TN} = 0.24 \text{ V}$, $K_n = 1.1 \text{ mA/V}^2$, and $\lambda = 0$. Let $R_1 + R_2 = 50 \text{ k}\Omega$. Design the circuit such that $V_{GS} = 0.50 \text{ V}$ and $V_{DS} = 2.5 \text{ V}$.

Solution: From Equation (3.38(a)) the drain current is

$$I_D = K_n(V_{GS} - V_{TN})^2 = (1.1)(0.5 - 0.24)^2 = 74.4 \mu\text{A}$$

From Figure 3.65(b), the voltage at the drain is

$$V_D = V_{DD} - I_D R_D = 4 - (0.0744)(6.7) = 3.5 \text{ V}$$

Therefore, the voltage at the source is

$$V_S = V_D - V_{DS} = 3.5 - 2.5 = 1 \text{ V}$$

The source resistance is then

$$R_S = \frac{V_S}{I_D} = \frac{1}{0.0744} = 13.4 \text{ k}\Omega$$

The voltage at the gate is

$$V_G = V_{GS} + V_S = 0.5 + 1 = 1.5 \text{ V}$$

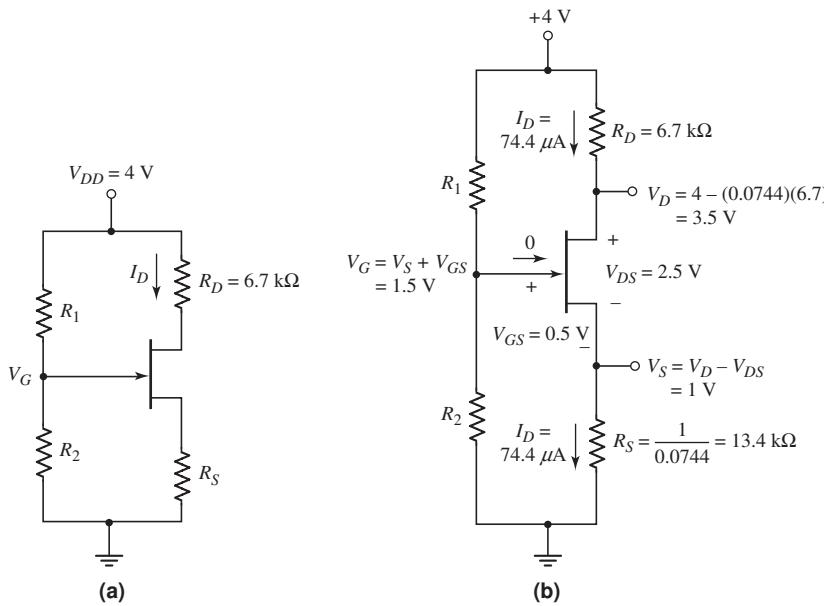


Figure 3.65 (a) An n-channel enhancement-mode MESFET circuit and (b) the n-channel MESFET circuit for Example 3.22

Since the gate current is zero, the gate voltage is also given by a voltage divider equation, as follows:

$$V_G = \left(\frac{R_2}{R_1 + R_2} \right) (V_{DD})$$

or

$$1.5 = \left(\frac{R_2}{50} \right) (4)$$

which yields

$$R_2 = 18.75 \text{ k}\Omega$$

and

$$R_1 = 31.25 \text{ k}\Omega$$

Again, we see that

$$V_{DS} = 2.5 \text{ V} > V_{GS} - V_{TN} = 0.5 - 0.24 = 0.26 \text{ V}$$

which confirms that the transistor is biased in the saturation region, as initially assumed.

Comment: The dc analysis and design of an enhancement-mode MESFET circuit is similar to that of MOSFET circuits, except that the gate-to-source voltage of the MESFET must be held to no more than a few tenths of a volt.

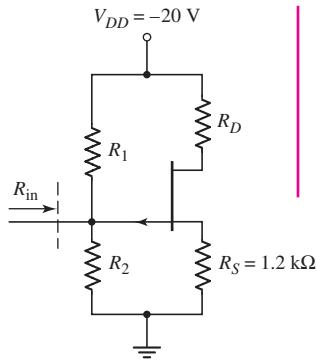


Figure 3.66 Circuit for Exercise Ex3.22

EXERCISE PROBLEM

Ex 3.22: Consider the circuit shown in Figure 3.66 with transistor parameters $I_{DSS} = 8 \text{ mA}$, $V_P = 4 \text{ V}$, and $\lambda = 0$. Design the circuit such that $R_{in} = 100 \text{ k}\Omega$, $I_{DQ} = 5 \text{ mA}$, and $V_{SDQ} = 12 \text{ V}$. (Ans. $R_D = 0.4 \text{ k}\Omega$, $R_1 = 387 \text{ k}\Omega$, $R_2 = 135 \text{ k}\Omega$)

Test Your Understanding

TYU 3.17 The n-channel enhancement-mode MESFET in the circuit shown in Figure 3.67 has parameters $K_n = 50 \mu\text{A}/\text{V}^2$ and $V_{TN} = 0.15 \text{ V}$. Find the value of V_{GG} so that $I_{DQ} = 5 \mu\text{A}$. What are the values of V_{GS} and V_{DS} ? (Ans. $V_{GG} = 0.516 \text{ V}$, $V_{GS} = 0.466 \text{ V}$, $V_{DS} = 4.45 \text{ V}$)

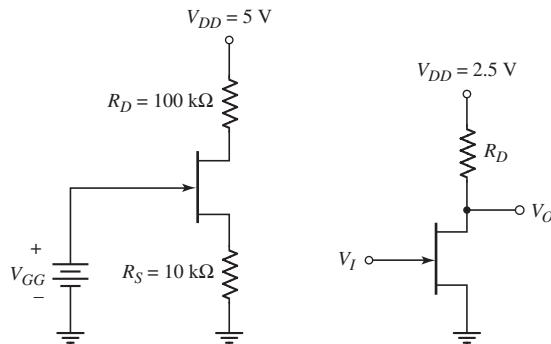


Figure 3.67 Circuit for Exercise TYU 3.17

Figure 3.68 Circuit for Exercise TYU 3.18

TYU 3.18 For the inverter circuit shown in Figure 3.68, the n-channel enhancement-mode MESFET parameters are $K_n = 100 \mu\text{A}/\text{V}^2$ and $V_{TN} = 0.2 \text{ V}$. Determine the value of R_D required to produce $V_O = 0.10 \text{ V}$ when $V_I = 0.7 \text{ V}$. (Ans. $R_D = 267 \text{ k}\Omega$)



DESIGN APPLICATION: DIODE THERMOMETER WITH AN MOS TRANSISTOR

Objective: • Incorporate an MOS transistor in a design application that enhances the simple diode thermometer design discussed in Chapter 1.

Specifications: The electronic thermometer is to operate over a temperature range of 0 to 100 °F.

Design Approach: The output diode voltage developed in the diode thermometer in Figure 1.47 is to be applied between the gate–source terminals of an NMOS transistor to enhance the voltage over the temperature range. The NMOS transistor is to be held at a constant temperature.

Choices: Assume an n-channel, depletion-mode MOSFET is available with the parameters $k'_n = 80 \mu\text{A/V}^2$, $W/L = 10$, and $V_{TN} = -1 \text{ V}$.

Solution: From the design in Chapter 1, the diode voltage is given by

$$V_D = 1.12 - 0.522 \left(\frac{T}{300} \right)$$

where T is in kelvins.

Consider the circuit shown in Figure 3.69. We assume that the diode is in a variable temperature environment while the rest of the circuit is held at room temperature.

From the circuit, we see that $V_{GS} = V_D$, where V_D is the diode voltage and not the drain voltage. We want the MOSFET biased in the saturation region, so

$$I_D = K_n (V_{GS} - V_{TN})^2 = \frac{k'_n}{2} \cdot \frac{W}{L} (V_D - V_{TN})^2$$

We find the output voltage as

$$\begin{aligned} V_O &= 15 - I_D R_D \\ &= 15 - \frac{k'_n}{2} \cdot \frac{W}{L} \cdot R_D (V_D - V_{TN})^2 \end{aligned}$$

The diode current and output voltage can be written as

$$I_D = \frac{0.080}{2} \cdot \frac{10}{1} (V_D + 1)^2 = 0.4(V_D + 1)^2 (\text{mA})$$

and

$$V_O = 15 - [0.4(V_D + 1)^2](10) = 15 - 4(V_D + 1)^2 (\text{V})$$

From Chapter 1, we have the following:

$T (\text{°F})$	$V_D (\text{V})$
0	0.6760
40	0.6372
80	0.5976
100	0.5790

We find the circuit response as:

$T (\text{°F})$	$I_D (\text{mA})$	$V_O (\text{V})$
0	1.124	3.764
40	1.072	4.278
80	1.021	4.791
100	0.9973	5.027

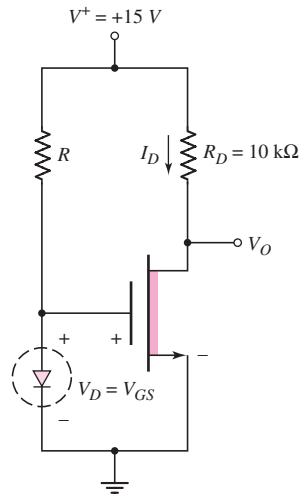


Figure 3.69 Design application circuit to measure output voltage of diode versus temperature

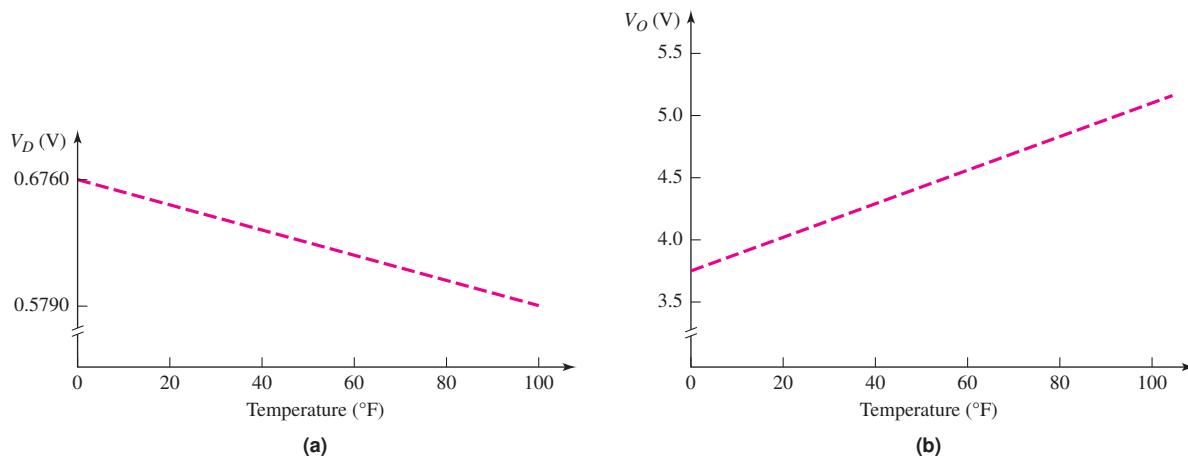


Figure 3.70 (a) Diode voltage versus temperature and (b) circuit output voltage versus temperature

Comment: Figure 3.70(a) shows the diode voltage versus temperature and Figure 3.70(b) now shows the output voltage versus temperature from the MOSFET circuit. We can see that the transistor circuit provides a voltage gain. This voltage gain is the desired characteristic of the transistor circuit.

Discussion: We can see from the equations that the diode current and output voltage are not linear functions of the diode voltage. This effect implies that the transistor output voltage is also not a linear function of temperature. We will see a better circuit design using operational amplifiers in Chapter 9.

We can note from the results that $V_O = V_{DS} > V_{DS(\text{sat})}$ in all cases, so the transistor is biased in the saturation region as desired.

3.8 SUMMARY

- In this chapter, we have considered the physical structure and dc electrical characteristics of the MOSFET.
- The current in the MOSFET is controlled by the gate voltage. In the nonsaturation bias region of operation, the drain current is also a function of drain voltage, whereas in the saturation bias region of operation the drain current is essentially independent of the drain voltage. The drain current is directly proportional to the width-to-length ratio of the transistor, so this parameter becomes the primary design variable in MOSFET circuit design.
- The dc analysis and design techniques of MOSFET circuits were emphasized in this chapter. The use of MOSFETs in place of resistors was investigated. This leads to the design of all-MOSFET circuits.
- Basic applications of the MOSFET include switching currents and voltages, performing digital logic functions, and amplifying time-varying signals. The amplifying characteristics will be considered in the next chapter and the digital applications will be considered in Chapter 16.
- MOSFET circuits that provide constant-current biasing to other MOSFET circuits were analyzed and designed. This current biasing technique is used in integrated circuits.

- The dc analysis and design of multistage MOSFET circuits were considered.
- The physical structure and dc electrical characteristics of JFET and MESFET devices as well as the analysis and design of JFET and MESFET circuits were considered.
- As an application, a MOSFET transistor was incorporated in a circuit design that enhances the simple diode electronic thermometer discussed in Chapter 1.

CHECKPOINT

After studying this chapter, the reader should have the ability to:

- ✓ Understand and describe the structure and general operation of n-channel and p-channel enhancement-mode and depletion-mode MOSFETs.
- ✓ Apply the ideal current–voltage relations in the dc analysis and design of various MOSFET circuits using any of the four basic MOSFETs.
- ✓ Understand how MOSFETs can be used in place of resistor load devices to create all-MOSFET circuits.
- ✓ Qualitatively understand how MOSFETs can be used to switch currents and voltages, to perform digital logic functions, and to amplify time-varying signals.
- ✓ Understand the basic operation of a MOSFET constant-current circuit.
- ✓ Understand the dc analysis and design of a multistage MOSFET circuit.
- ✓ Understand the general operation and characteristics of junction FETs.

REVIEW QUESTIONS

1. Describe the basic structure and operation of a MOSFET. Define enhancement mode and depletion mode.
2. Sketch the general current–voltage characteristics for both enhancement-mode and depletion-mode MOSFETs. Define the saturation and nonsaturation bias regions.
3. Describe what is meant by threshold voltage, width-to-length ratio, and drain-to-source saturation voltage.
4. Describe the channel length modulation effect and define the parameter λ . Describe the body effect and define the parameter γ .
5. Describe a simple common-source MOSFET circuit with an n-channel enhancement-mode device and discuss the relation between the drain-to-source voltage and gate-to-source voltage.
6. How do you prove that a MOSFET is biased in the saturation region?
7. In the dc analysis of some MOSFET circuits, quadratic equations in gate-to-source voltage are developed. How do you determine which of the two possible solutions is the correct one?
8. How can the Q -point be stabilized against variations in transistor parameters?
9. Describe the current–voltage relation of an n-channel enhancement-mode MOSFET with the gate connected to the drain.
10. Describe the current–voltage relation of an n-channel depletion-mode MOSFET with the gate connected to the source.
11. Describe a MOSFET NOR logic circuit.
12. Describe how a MOSFET can be used to amplify a time-varying voltage.
13. Describe the basic operation of a junction FET.
14. What is the difference between a MESFET and a pn junction FET?

 PROBLEMS

[Note: In all problems, assume the transistor parameter $\lambda = 0$, unless otherwise stated.]

Section 3.1 MOS Field-Effect Transistor

- 3.1 (a) Calculate the drain current in an NMOS transistor with parameters $V_{TN} = 0.4$ V, $k'_n = 120 \mu\text{A/V}^2$, $W = 10 \mu\text{m}$, $L = 0.8 \mu\text{m}$, and with applied voltages of $V_{DS} = 0.1$ V and (i) $V_{GS} = 0$, (ii) $V_{GS} = 1$ V, (iii) $V_{GS} = 2$ V, and (iv) $V_{GS} = 3$ V. (b) Repeat part (a) for $V_{DS} = 4$ V.
- 3.2 The current in an NMOS transistor is 0.5 mA when $V_{GS} - V_{TN} = 0.6$ V and 1.0 mA when $V_{GS} - V_{TN} = 1.0$ V. The device is operating in the nonsaturation region. Determine V_{DS} and K_n .
- 3.3 The transistor characteristics i_D versus v_{DS} for an NMOS device are shown in Figure P3.3. (a) Is this an enhancement-mode or depletion-mode device? (b) Determine the values for K_n and V_{TN} . (c) Determine $i_D(\text{sat})$ for $v_{GS} = 3.5$ V and $v_{GS} = 4.5$ V.

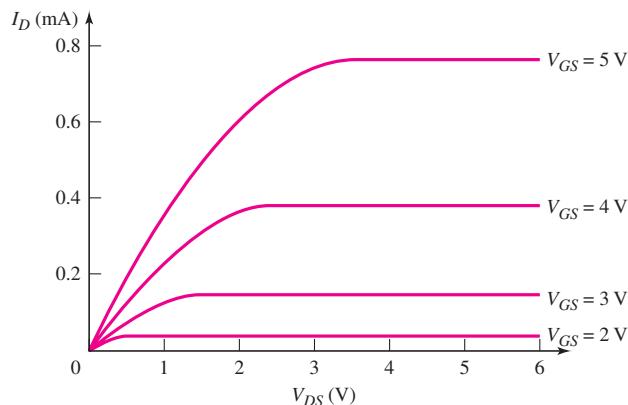


Figure P3.3

- 3.4 For an n-channel depletion-mode MOSFET, the parameters are $V_{TN} = -2.5$ V and $K_n = 1.1 \text{ mA/V}^2$. (a) Determine I_D for $V_{GS} = 0$; and: (i) $V_{DS} = 0.5$ V, (ii) $V_{DS} = 2.5$ V, and (iii) $V_{DS} = 5$ V. (b) Repeat part (a) for $V_{GS} = 2$ V.
- 3.5 The threshold voltage of each transistor in Figure P3.5 is $V_{TN} = 0.4$ V. Determine the region of operation of the transistor in each circuit.

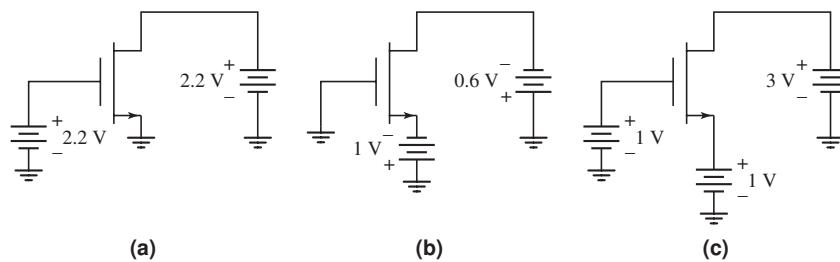


Figure P3.5

- 3.6 The threshold voltage of each transistor in Figure P3.6 is $V_{TP} = -0.4 \text{ V}$. Determine the region of operation of the transistor in each circuit.

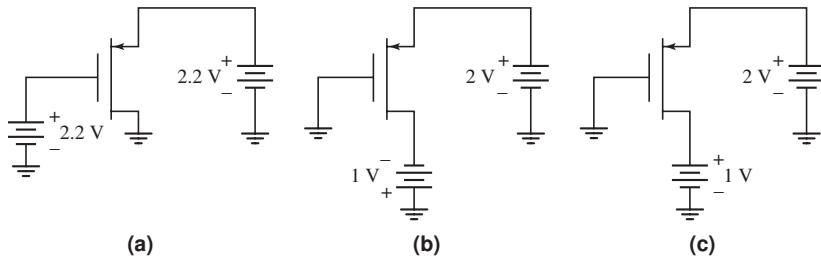


Figure P3.6

- 3.7 Consider an n-channel depletion-mode MOSFET with parameters $V_{TN} = -1.2 \text{ V}$ and $k'_n = 120 \mu\text{A/V}^2$. The drain current is $I_D = 0.5 \text{ mA}$ at $V_{GS} = 0$ and $V_{DS} = 2 \text{ V}$. Determine the W/L ratio.
- 3.8 Determine the value of the process conduction parameter k'_n for an NMOS transistor with $\mu_n = 600 \text{ cm}^2/\text{V}\cdot\text{s}$ and for an oxide thickness t_{ox} of (a) 500 \AA , (b) 250 \AA , (c) 100 \AA , (d) 50 \AA , and (e) 25 \AA .
- 3.9 An n-channel enhancement-mode MOSFET has parameters $V_{TN} = 0.4 \text{ V}$, $W = 20 \mu\text{m}$, $L = 0.8 \mu\text{m}$, $t_{ox} = 200 \text{ \AA}$, and $\mu_n = 650 \text{ cm}^2/\text{V}\cdot\text{s}$. (a) Calculate the conduction parameter K_n . (b) Determine the drain current when $V_{GS} = V_{DS} = 2 \text{ V}$. (c) With $V_{GS} = 2 \text{ V}$, what value of V_{DS} puts the device at the edge of saturation?
- 3.10 An NMOS device has parameters $V_{TN} = 0.8 \text{ V}$, $L = 0.8 \mu\text{m}$, and $k'_n = 120 \mu\text{A/V}^2$. When the transistor is biased in the saturation region with $V_{GS} = 1.4 \text{ V}$, the drain current is $I_D = 0.6 \text{ mA}$. (a) What is the channel width W ? (b) Determine the drain current when $V_{DS} = 0.4 \text{ V}$. (c) What value of V_{DS} puts the device at the edge of saturation?
- 3.11 A particular NMOS device has parameters $V_{TN} = 0.6 \text{ V}$, $L = 0.8 \mu\text{m}$, $t_{ox} = 200 \text{ \AA}$, and $\mu_n = 600 \text{ cm}^2/\text{V}\cdot\text{s}$. A drain current of $I_D = 1.2 \text{ mA}$ is required when the device is biased in the saturation region at $V_{GS} = 3 \text{ V}$. Determine the required channel width of the device.
- 3.12 MOS transistors with very short channels do not exhibit the square law voltage relation in saturation. The drain current is instead given by
- $$I_D = WC_{\text{ox}}(V_{GS} - V_{TN})v_{\text{sat}}$$
- where v_{sat} is a saturation velocity. Assuming $v_{\text{sat}} = 2 \times 10^7 \text{ cm/s}$ and using the parameters in Problem 3.11, determine the current.
- 3.13 For a p-channel enhancement-mode MOSFET, $k'_p = 50 \mu\text{A/V}^2$. The device has drain currents of $I_D = 0.225 \text{ mA}$ at $V_{SG} = V_{SD} = 2 \text{ V}$ and $I_D = 0.65 \text{ mA}$ at $V_{SG} = V_{SD} = 3 \text{ V}$. Determine the W/L ratio and the value of V_{TP} .
- 3.14 For a p-channel enhancement-mode MOSFET, the parameters are $K_P = 2 \text{ mA/V}^2$ and $V_{TP} = -0.5 \text{ V}$. The gate is at ground potential, and the source and substrate terminals are at $+5 \text{ V}$. Determine I_D when the drain terminal voltage is: (a) $V_D = 0 \text{ V}$, (b) $V_D = 2 \text{ V}$, (c) $V_D = 4 \text{ V}$, and (d) $V_D = 5 \text{ V}$.
- 3.15 The transistor characteristics i_D versus v_{SD} for a PMOS device are shown in Figure P3.15. (a) Is this an enhancement-mode or depletion-mode device?

- (b) Determine the values for K_p and V_{TP} . (c) Determine $i_D(\text{sat})$ for $v_{SG} = 3.5 \text{ V}$ and $v_{SG} = 4.5 \text{ V}$.

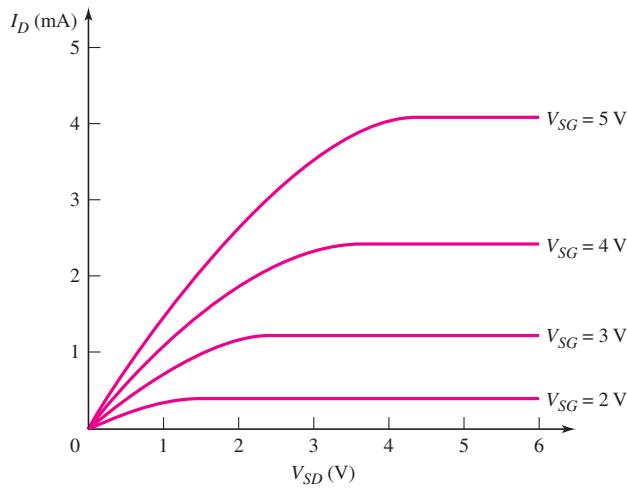


Figure P3.15

- 3.16 A p-channel depletion-mode MOSFET has parameters $V_{TP} = +2 \text{ V}$, $k'_p = 40 \mu\text{A}/\text{V}^2$, and $W/L = 6$. Determine $V_{SD}(\text{sat})$ for: (a) $V_{SG} = -1 \text{ V}$, (b) $V_{SG} = 0$, and (c) $V_{SG} = +1 \text{ V}$. If the transistor is biased in the saturation region, calculate the drain current for each value of V_{SG} .
- 3.17 Calculate the drain current in a PMOS transistor with parameters $V_{TP} = -0.5 \text{ V}$, $k'_p = 50 \mu\text{A}/\text{V}^2$, $W = 12 \mu\text{m}$, $L = 0.8 \mu\text{m}$, and with applied voltages of $V_{SG} = 2 \text{ V}$ and (a) $V_{SD} = 0.2 \text{ V}$, (b) $V_{SD} = 0.8 \text{ V}$, (c) $V_{SD} = 1.2 \text{ V}$, (d) $V_{SD} = 2.2 \text{ V}$, and (e) $V_{SD} = 3.2 \text{ V}$.
- 3.18 Determine the value of the process conduction parameter k'_p for a PMOS transistor with $\mu_p = 250 \text{ cm}^2/\text{V}\cdot\text{s}$ and for an oxide thickness t_{ox} of (a) 500 \AA , (b) 250 \AA , (c) 100 \AA , (d) 50 \AA , and (e) 25 \AA .
- 3.19 Enhancement-mode NMOS and PMOS devices both have parameters $L = 4 \mu\text{m}$ and $t_{\text{ox}} = 500 \text{ \AA}$. For the NMOS transistor, $V_{TN} = +0.6 \text{ V}$, $\mu_n = 675 \text{ cm}^2/\text{V}\cdot\text{s}$, and the channel width is W_n ; for the PMOS transistor, $V_{TP} = -0.6 \text{ V}$, $\mu_p = 375 \text{ cm}^2/\text{V}\cdot\text{s}$, and the channel width is W_p . Design the widths of the two transistors such that they are electrically equivalent and the drain current in the PMOS transistor is $I_D = 0.8 \text{ mA}$ when it is biased in the saturation region at $V_{SG} = 5 \text{ V}$. What are the values of K_n , K_p , W_n , and W_p ?
- 3.20 For an NMOS enhancement-mode transistor, the parameters are: $V_{TN} = 1.2 \text{ V}$, $K_n = 0.20 \text{ mA}/\text{V}^2$, and $\lambda = 0.01 \text{ V}^{-1}$. Calculate the output resistance r_o for $V_{GS} = 2.0 \text{ V}$ and for $V_{GS} = 4.0 \text{ V}$. What is the value of V_A ?
- 3.21 The parameters of an n-channel enhancement-mode MOSFET are $V_{TN} = 0.5 \text{ V}$, $k'_n = 120 \mu\text{A}/\text{V}^2$, and $W/L = 4$. What is the maximum value of λ and the minimum value of V_A such that for $V_{GS} = 2 \text{ V}$, $r_o \geq 200 \text{ k}\Omega$?
- 3.22 An enhancement-mode NMOS transistor has parameters $V_{TNO} = 0.8 \text{ V}$, $\gamma = 0.8 \text{ V}^{1/2}$, and $\phi_f = 0.35 \text{ V}$. At what value of V_{SB} will the threshold voltage change by 2V due to the body effect?
- 3.23 An NMOS transistor has parameters $V_{TO} = 0.75 \text{ V}$, $k'_n = 80 \mu\text{A}/\text{V}^2$, $W/L = 15$, $\phi_f = 0.37 \text{ V}$, and $\gamma = 0.6 \text{ V}^{1/2}$. (a) The transistor is biased at

- $V_{GS} = 2.5 \text{ V}$, $V_{SB} = 3 \text{ V}$, and $V_{DS} = 3 \text{ V}$. Determine the drain current I_D .
- (b) Repeat part (a) for $V_{DS} = 0.25 \text{ V}$.
- 3.24 (a) A silicon dioxide gate insulator of an MOS transistor has a thickness of $t_{\text{ox}} = 120 \text{ \AA}$. (i) Calculate the ideal oxide breakdown voltage. (ii) If a safety factor of three is required, determine the maximum safe gate voltage that may be applied. (b) Repeat part (a) for an oxide thickness of $t_{\text{ox}} = 200 \text{ \AA}$.
- 3.25 In a power MOS transistor, the maximum applied gate voltage is 24 V. If a safety factor of three is specified, determine the minimum thickness necessary for the silicon dioxide gate insulator.

Section 3.2 Transistor dc Analysis

- 3.26 In the circuit in Figure P3.26, the transistor parameters are $V_{TN} = 0.8 \text{ V}$ and $K_n = 0.5 \text{ mA/V}^2$. Calculate V_{GS} , I_D , and V_{DS} .
- 3.27 The transistor in the circuit in Figure P3.27 has parameters $V_{TN} = 0.8 \text{ V}$ and $K_n = 0.25 \text{ mA/V}^2$. Sketch the load line and plot the Q -point for (a) $V_{DD} = 4 \text{ V}$, $R_D = 1 \text{ k}\Omega$ and (b) $V_{DD} = 5 \text{ V}$, $R_D = 3 \text{ k}\Omega$. What is the operating bias region for each condition?

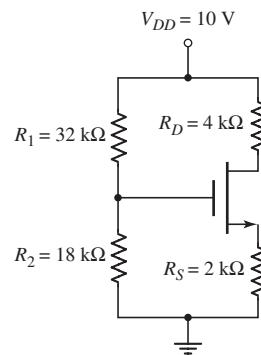


Figure P3.26

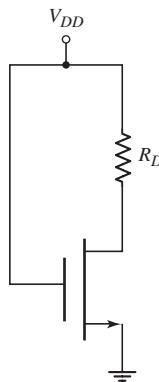


Figure P3.27

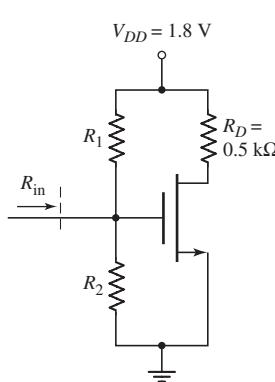


Figure P3.28

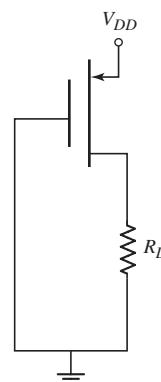


Figure P3.29

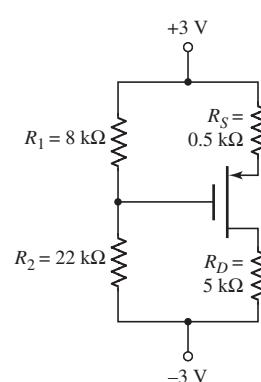


Figure P3.30

- D3.28 The transistor in Figure P3.28 has parameters $V_{TN} = 0.4 \text{ V}$, $k'_n = 120 \mu\text{A/V}^2$, and $W/L = 80$. Design the circuit such that $I_Q = 0.8 \text{ mA}$ and $R_{in} = 200 \text{ k}\Omega$.
- 3.29 The transistor in the circuit in Figure P3.29 has parameters $V_{TP} = -0.8 \text{ V}$ and $K_p = 0.20 \text{ mA/V}^2$. Sketch the load line and plot the Q -point for (a) $V_{DD} = 3.5 \text{ V}$, $R_D = 1.2 \text{ k}\Omega$ and (b) $V_{DD} = 5 \text{ V}$, $R_D = 4 \text{ k}\Omega$. What is the operating bias region for each condition?
- 3.30 Consider the circuit in Figure P3.30. The transistor parameters are $V_{TP} = -0.8 \text{ V}$ and $K_p = 0.5 \text{ mA/V}^2$. Determine I_D , V_{SG} , and V_{SD} .
- 3.31 For the circuit in Figure P3.31, the transistor parameters are $V_{TP} = -0.8 \text{ V}$ and $K_p = 200 \mu\text{A/V}^2$. Determine V_S and V_{SD} .
- D3.32 Design a MOSFET circuit in the configuration shown in Figure P3.26. The transistor parameters are $V_{TN} = 0.4 \text{ V}$ and $k'_n = 120 \mu\text{A/V}^2$, and $\lambda = 0$. The circuit parameters are $V_{DD} = 3.3 \text{ V}$ and $R_D = 5 \text{ k}\Omega$. Design the circuit so that $V_{DSQ} \cong 1.6 \text{ V}$ and the voltage across R_S is approximately 0.8 V. Set $V_{GS} = 0.8 \text{ V}$. The current through the bias resistors is to be approximately 5 percent of the drain current.

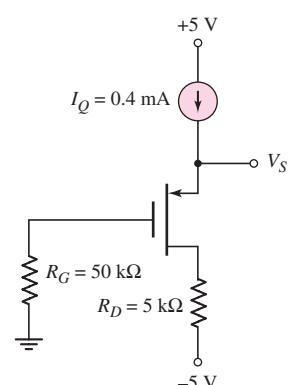


Figure P3.31

- 3.33 Consider the circuit shown in Figure P3.33. The transistor parameters are $V_{TN} = 0.4$ V and $k'_n = 120 \mu\text{A}/\text{V}^2$. The voltage drop across R_S is to be 0.20 V. Design the transistor W/L ratio such that $V_{DS} = V_{DS(\text{sat})} + 0.4$ V, and determine R_1 and R_2 such that $R_{in} = 200 \text{ k}\Omega$.

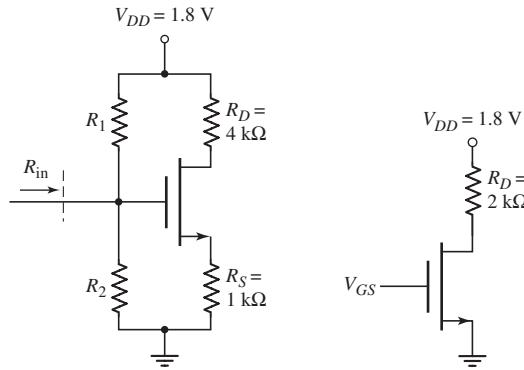


Figure P3.33

Figure P3.34

- 3.34 The transistor parameters for the transistor in Figure P3.34 are $V_{TN} = 0.4$ V, $k'_n = 120 \mu\text{A}/\text{V}^2$, and $W/L = 50$. (a) Determine V_{GS} such that $I_D = 0.35$ mA. (b) Determine V_{DS} and $V_{DS(\text{sat})}$.
- 3.35 For the transistor in the circuit in Figure P3.35, the parameters are $V_{TN} = 0.4$ V, $k'_n = 120 \mu\text{A}/\text{V}^2$, and $W/L = 25$. Determine V_{GS} , I_D , and V_{DS} . Sketch the load line and plot the Q -point.
- D3.36 Design a MOSFET circuit with the configuration shown in Figure P3.30. The transistor parameters are $V_{TP} = -0.6$ V, $k'_p = 50 \mu\text{A}/\text{V}^2$, and $\lambda = 0$. The circuit bias is ± 3 V, the drain current is to be 0.2 mA, the drain-to-source voltage is to be approximately 3 V, and the voltage across R_S is to be approximately equal to V_{SG} . In addition, the current through the bias resistors is to be no more than 10 percent of the drain current. (*Hint:* choose a reasonable value of width-to-length ratio for the transistor.)
- 3.37 The parameters of the transistors in Figures P3.37 (a) and (b) are $K_n = 0.5 \text{ mA}/\text{V}^2$, $V_{TN} = 1.2$ V, and $\lambda = 0$. Determine v_{GS} and v_{DS} for each transistor when (i) $I_Q = 50 \mu\text{A}$ and (ii) $I_Q = 1$ mA.

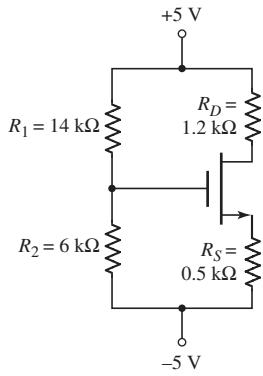


Figure P3.35

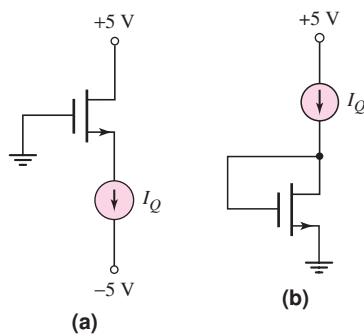


Figure P3.37

- 3.38 For the circuit in Figure P3.38, the transistor parameters are $V_{TN} = 0.6$ V and $K_n = 200 \mu\text{A/V}^2$. Determine V_S and V_D .

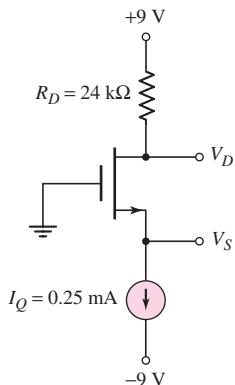


Figure P3.38

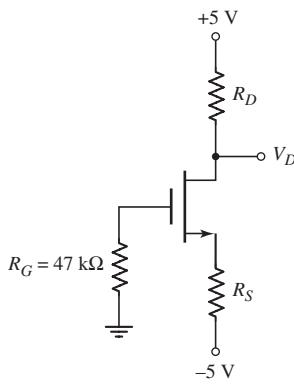


Figure P3.39

- *3.39 (a) Design the circuit in Figure P3.39 such that $I_{DQ} = 0.50 \text{ mA}$ and $V_D = 1 \text{ V}$. The transistor parameters are $K_n = 0.25 \text{ mA/V}^2$ and $V_{TN} = 1.4 \text{ V}$. Sketch the load line and plot the Q -point. (b) Choose standard resistor values that are closest to the ideal designed values. What are the resulting Q -point values? (c) If the resistors in part (b) have tolerances of ± 10 percent, determine the maximum and minimum values of I_{DQ} .
- 3.40 The PMOS transistor in Figure P3.40 has parameters $V_{TP} = -0.7 \text{ V}$, $k'_p = 50 \mu\text{A/V}^2$, $L = 0.8 \mu\text{m}$, and $\lambda = 0$. Determine the values of W and R such that $I_D = 0.1 \text{ mA}$ and $V_{SD} = 2.5 \text{ V}$.

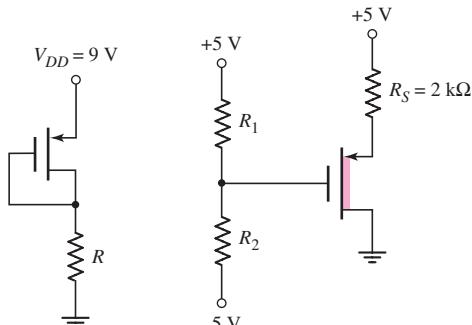


Figure P3.40

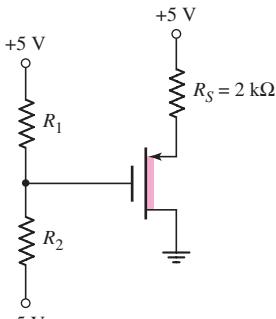


Figure P3.41

- 3.41 Design the circuit in Figure P3.41 so that $V_{SD} = 2.5 \text{ V}$. The current in the bias resistors should be no more than 10 percent of the drain current. The transistor parameters are $V_{TP} = +1.5 \text{ V}$ and $K_p = 0.5 \text{ mA/V}^2$.
- *3.42 (a) Design the circuit in Figure P3.42 such that $I_{DQ} = 0.25 \text{ mA}$ and $V_D = -2 \text{ V}$. The nominal transistor parameters are $V_{TP} = -1.2 \text{ V}$, $k'_p = 35 \mu\text{A/V}^2$, and $W/L = 15$. Sketch the load line and plot the Q -point. (b) Determine the maximum and minimum Q -point values if the tolerance of the k'_p parameter is ± 5 percent.

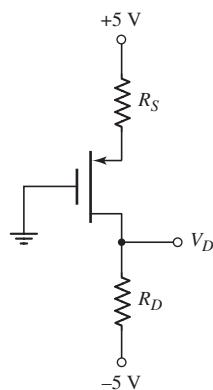


Figure P3.42

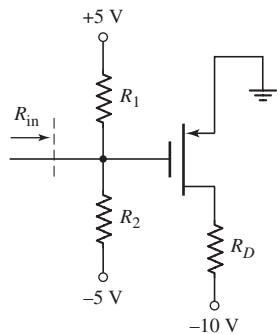


Figure P3.43

- 3.43 The parameters of the transistor in the circuit in Figure P3.43 are $V_{TP} = -1.75$ V and $k_p = 3$ mA/V². Design the circuit such that $I_D = 5$ mA, $V_{SD} = 6$ V, and $R_{in} = 80$ k Ω .

- 3.44 For each transistor in the circuit in Figure P3.44, $k'_n = 120 \mu\text{A}/\text{V}^2$. Also for M_1 , $W/L = 4$ and $V_{TN} = 0.4$ V, and for M_2 , $W/L = 1$ and $V_{TN} = -0.6$ V. (a) Determine the input voltage such that both M_1 and M_2 are biased in the saturation region. (b) Determine the region of operation of each transistor and the output voltage v_O for: (i) $v_I = 0.6$ V and (ii) $v_I = 1.5$ V.

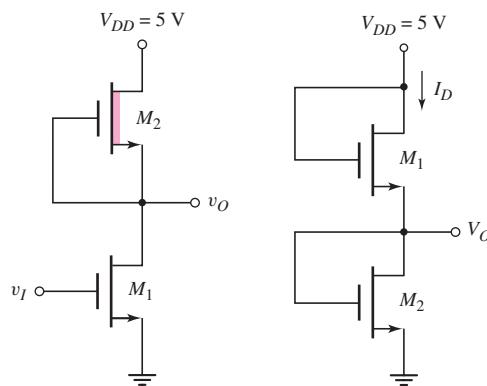


Figure P3.44

Figure P3.46

- 3.45 Consider the circuit in Figure P3.44. The transistor parameters for M_1 are $V_{TN} = 0.4$ V and $k'_n = 120 \mu\text{A}/\text{V}^2$, and for M_2 are $V_{TN} = -0.6$ V, $k'_n = 120 \mu\text{A}/\text{V}^2$, and $W/L = 1$. Determine the W/L ratio of M_1 such that $v_O = 0.025$ V when $v_I = 3$ V.

- 3.46 The transistors in the circuit in Figure P3.46 both have parameters $V_{TN} = 0.4$ V and $k'_n = 120 \mu\text{A}/\text{V}^2$. (a) If the width-to-length ratios of M_1 and M_2 are $(W/L)_1 = (W/L)_2 = 30$, determine V_{GS1} , V_{GS2} , V_O , and I_D . (b) Repeat part (a) if the width-to-length ratios are changed to $(W/L)_1 = 30$ and $(W/L)_2 = 15$. (c) Repeat part (a) if the width-to-length ratios are changed to $(W/L)_1 = 15$ and $(W/L)_2 = 30$.

- 3.47 Consider the circuit in Figure P3.47. (a) The nominal transistor parameters are $V_{TN} = 0.6$ V and $k'_n = 120 \mu\text{A}/\text{V}^2$. Design the width-to-length ratio required in each transistor such that $I_{DQ} = 0.8$ mA, $V_1 = 2.5$ V, and $V_2 = 6$ V. (b) Determine the change in the values of V_1 and V_2 if the k'_n parameter in each transistor changes by (i) +5 percent and (ii) -5 percent. (c) Determine the values of V_1 and V_2 if the k'_n parameter of M_1 decreases by 5 percent and the k'_n parameter of M_2 and M_3 increases by 5 percent.

- 3.48 The transistors in the circuit in Figure 3.36 in the text have parameters $V_{TN} = 0.6$ V, $k'_n = 120 \mu\text{A}/\text{V}^2$, and $\lambda = 0$. The width-to-length ratio of M_L is $(W/L)_L = 2$. Design the width-to-length ratio of the driver transistor such that $V_O = 0.15$ V when $V_I = 5$ V.

- 3.49 For the circuit in Figure 3.39 in the text, the transistor parameters are: $V_{TND} = 0.6$ V, $V_{TNL} = -1.2$ V, $\lambda = 0$, and $k'_n = 120 \mu\text{A}/\text{V}^2$. Let $V_{DD} = 5$ V. The width-to-length ratio of M_L is $(W/L)_L = 2$. Design the width-to-length ratio of the driver transistor such that $V_O = 0.10$ V when $V_I = 5$ V.

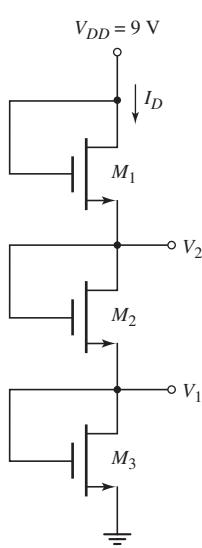


Figure P3.47

Section 3.3 MOSFET Switch and Amplifier

- 3.50 Consider the circuit in Figure P3.50. The circuit parameters are $V_{DD} = 3\text{ V}$ and $R_D = 30\text{ k}\Omega$. The transistor parameters are $V_{TN} = 0.4\text{ V}$ and $k'_n = 120\text{ }\mu\text{A/V}^2$. (a) Determine the transistor width-to-length ratio such that $V_O = 0.08\text{ V}$ when $V_I = 2.6\text{ V}$. (b) Repeat part (a) for $V_I = 3\text{ V}$.

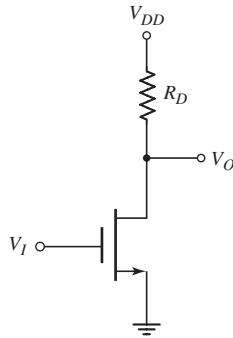


Figure P3.50

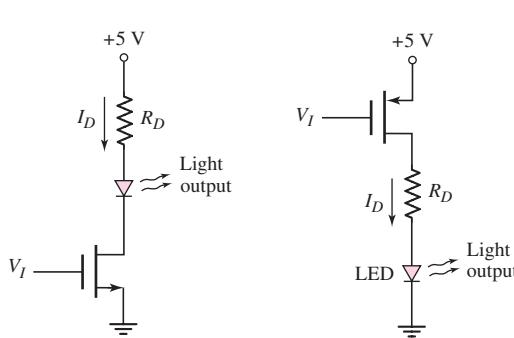


Figure P3.51

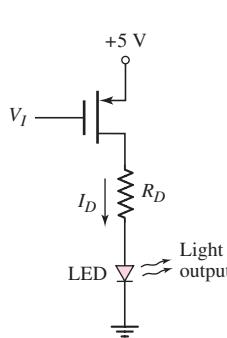


Figure P3.52

- 3.51 The transistor in the circuit in Figure P3.51 is used to turn the LED on and off. The transistor parameters are $V_{TN} = 0.6\text{ V}$, $k'_n = 80\text{ }\mu\text{A/V}^2$, and $\lambda = 0$. The diode cut-in voltage is $V_\gamma = 1.6\text{ V}$. Design R_D and the transistor width-to-length ratio such that $I_D = 12\text{ mA}$ for $V_I = 5\text{ V}$ and $V_{DS} = 0.15\text{ V}$.
- 3.52 The circuit in Figure P3.52 is another configuration used to switch an LED on and off. The transistor parameters are $V_{TP} = -0.6\text{ V}$, $k'_p = 40\text{ }\mu\text{A/V}^2$, and $\lambda = 0$. The diode cut-in voltage is $V_\gamma = 1.6\text{ V}$. Design R_D and the transistor width-to-length ratio such that $I_D = 15\text{ mA}$ for $V_I = 0\text{ V}$ and $V_{SD} = 0.20\text{ V}$.
- 3.53 For the two-input NMOS NOR logic gate in Figure 3.46 in the text, the transistor parameters are $V_{TN1} = V_{TN2} = 0.6\text{ V}$, $\lambda_1 = \lambda_2 = 0$, and $k'_{n1} = k'_{n2} = 120\text{ }\mu\text{A/V}^2$. The drain resistor is $R_D = 50\text{ k}\Omega$. (a) Determine the width-to-length ratios of the transistors so that $V_O = 0.15\text{ V}$ when $V_1 = V_2 = 5\text{ V}$. Assume that $(W/L)_1 = (W/L)_2$. (b) Using the results of part (a), find V_O when $V_1 = 5\text{ V}$ and $V_2 = 0.2\text{ V}$.
- 3.54 All transistors in the current-source circuit shown in Figure 3.49(a) in the text have parameters $V_{TN} = 0.4\text{ V}$, $k'_n = 120\text{ }\mu\text{A/V}^2$, and $\lambda = 0$. Transistors M_1 and M_2 are matched. The bias sources are $V^+ = 2.5\text{ V}$ and $V^- = -2.5\text{ V}$. The currents are to be $I_{Q1} = 125\text{ }\mu\text{A}$ and $I_{REF1} = 225\text{ }\mu\text{A}$. For M_2 , we require $V_{DS2}(\text{sat}) = 0.5\text{ V}$, and for M_1 we require $V_{DS1} = 2\text{ V}$. (a) Find the W/L ratios of the transistors. (b) Find R_D .
- 3.55 All transistors in the current-source circuit shown in Figure 3.49(b) in the text have parameters $V_{TP} = -0.4\text{ V}$, $k'_p = 50\text{ }\mu\text{A/V}^2$, and $\lambda = 0$. The bias sources are $V^+ = 5\text{ V}$ and $V^- = -5\text{ V}$. The currents are to be $I_{Q2} = 200\text{ }\mu\text{A}$ and $I_{REF2} = 125\text{ }\mu\text{A}$. For M_B , we require $V_{SDB}(\text{sat}) = 0.8\text{ V}$, and for M_A , we require $V_{SDA} = 4\text{ V}$. Transistors M_A and M_B are matched. (a) Find the W/L ratios of the transistors. (b) Find the value of R_D .
- 3.56 Consider the circuit shown in Figure 3.50 in the text. The threshold voltage and process conduction parameter for each transistor is $V_{TN} = 0.6\text{ V}$ and $k'_n = 120\text{ }\mu\text{A/V}^2$. Let $\lambda = 0$ for all transistors. Assume that M_1 and M_2 are matched. Design width-to-length ratios such that $I_Q = 0.35\text{ mA}$, $I_{REF} = 0.15\text{ mA}$, and $V_{DS2}(\text{sat}) = 0.5\text{ V}$. Find R_D such that $V_{DS1} = 3.5\text{ V}$.

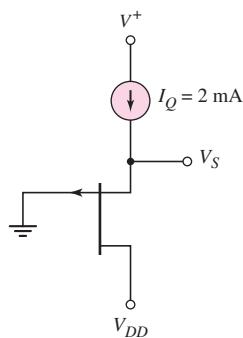


Figure P3.60

Section 3.6 Junction Field-Effect Transistor

- 3.57 The gate and source of an n-channel depletion-mode JFET are connected together. What value of V_{DS} will ensure that this two-terminal device is biased in the saturation region. What is the drain current for this bias condition?
- 3.58 For an n-channel JFET, the parameters are $I_{DSS} = 6 \text{ mA}$ and $V_P = -3 \text{ V}$. Calculate $V_{DS(\text{sat})}$. If $V_{DS} > V_{DS(\text{sat})}$, determine I_D for: (a) $V_{GS} = 0$, (b) $V_{GS} = -1 \text{ V}$, (c) $V_{GS} = -2 \text{ V}$, and (d) $V_{GS} = -3 \text{ V}$.
- 3.59 A p-channel JFET biased in the saturation region with $V_{SD} = 5 \text{ V}$ has a drain current of $I_D = 2.8 \text{ mA}$ at $V_{GS} = 1 \text{ V}$ and $I_D = 0.30 \text{ mA}$ at $V_{GS} = 3 \text{ V}$. Determine I_{DSS} and V_P .
- 3.60 Consider the p-channel JFET in Figure P3.60. Determine the range of V_{DD} that will bias the transistor in the saturation region. If $I_{DSS} = 6 \text{ mA}$ and $V_P = 2.5 \text{ V}$, find V_S .
- 3.61 Consider a GaAs MESFET. When the device is biased in the saturation region, we find that $I_D = 18.5 \mu\text{A}$ at $V_{GS} = 0.35 \text{ V}$ and $I_D = 86.2 \mu\text{A}$ at $V_{GS} = 0.50 \text{ V}$. Determine the conduction parameter k and the threshold voltage V_{TN} .
- 3.62 The threshold voltage of a GaAs MESFET is $V_{TN} = 0.24 \text{ V}$. The maximum allowable gate-to-source voltage is $V_{GS} = 0.75 \text{ V}$. When the transistor is biased in the saturation region, the maximum drain current is $I_D = 250 \mu\text{A}$. What is the value of the conduction parameter k ?
- *3.63 For the transistor in the circuit in Figure P3.63, the parameters are: $I_{DSS} = 10 \text{ mA}$ and $V_P = -5 \text{ V}$. Determine I_{DQ} , V_{GSQ} , and V_{DSQ} .

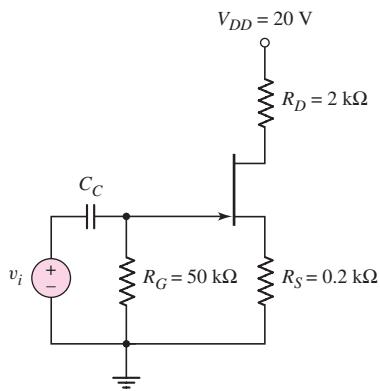


Figure P3.63

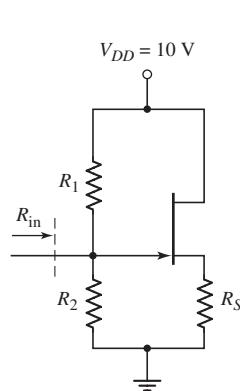


Figure P3.64

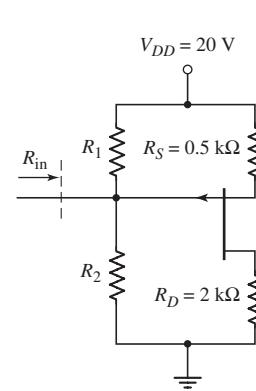


Figure P3.65

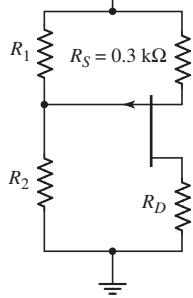


Figure P3.66

- 3.64 Consider the source follower with the n-channel JFET in Figure P3.64. The input resistance is to be $R_{in} = 500 \text{ k}\Omega$. We wish to have $I_{DQ} = 5 \text{ mA}$, $V_{DSQ} = 8 \text{ V}$, and $V_{GSQ} = -1 \text{ V}$. Determine R_S , R_1 , and R_2 , and the required transistor values of I_{DSS} and V_P .
- 3.65 The transistor in the circuit in Figure P3.65 has parameters $I_{DSS} = 8 \text{ mA}$ and $V_P = 4 \text{ V}$. Design the circuit such that $I_D = 5 \text{ mA}$. Assume $R_{in} = 100 \text{ k}\Omega$. Determine V_{GS} and V_{SD} .
- 3.66 For the circuit in Figure P3.66, the transistor parameters are $I_{DSS} = 7 \text{ mA}$ and $V_P = 3 \text{ V}$. Let $R_1 + R_2 = 100 \text{ k}\Omega$. Design the circuit such that $I_{DQ} = 5.0 \text{ mA}$ and $V_{SDQ} = 6 \text{ V}$.

- 3.67 The transistor in the circuit in Figure P3.67 has parameters $I_{DSS} = 8 \text{ mA}$ and $V_P = -4 \text{ V}$. Determine V_G , I_{DQ} , V_{GSQ} , and V_{DSQ} .
- 3.68 Consider the circuit in Figure P3.68. The quiescent value of V_{DS} is found to be $V_{DSQ} = 5 \text{ V}$. If $I_{DSS} = 10 \text{ mA}$, determine I_{DQ} , V_{GSQ} , and V_P .
- 3.69 For the circuit in Figure P3.69, the transistor parameters are $I_{DSS} = 4 \text{ mA}$ and $V_P = -3 \text{ V}$. Design R_D such that $V_{DS} = |V_P|$. What is the value of I_D ?

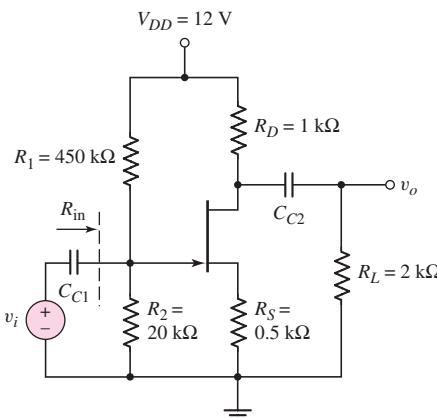


Figure P3.68

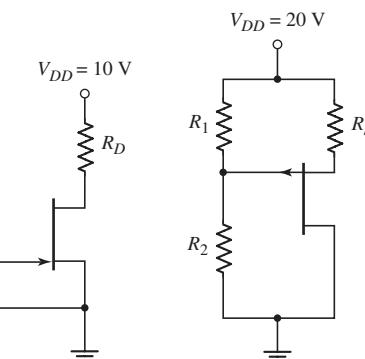


Figure P3.69

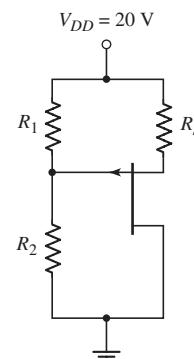


Figure P3.70

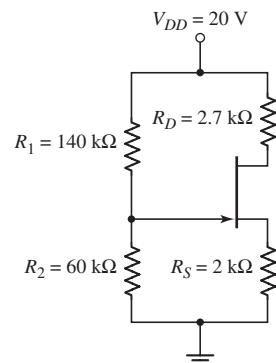


Figure P3.67

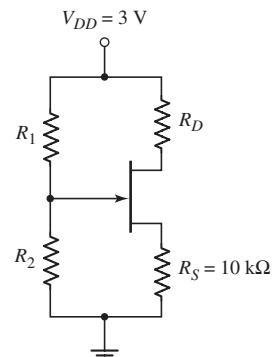


Figure P3.71

- 3.70 Consider the source-follower circuit in Figure P3.70. The transistor parameters are $I_{DSS} = 2 \text{ mA}$ and $V_P = 2 \text{ V}$. Design the circuit such that $I_{DQ} = 1 \text{ mA}$, $V_{SDQ} = 10 \text{ V}$, and the current through R_1 and R_2 is 0.1 mA .
- 3.71 The GaAs MESFET in the circuit in Figure P3.71 has parameters $k = 250 \mu\text{A/V}^2$ and $V_{TN} = 0.20 \text{ V}$. Let $R_1 + R_2 = 150 \text{ k}\Omega$. Design the circuit such that $I_D = 40 \mu\text{A}$ and $V_{DS} = 2 \text{ V}$.
- 3.72 For the circuit in Figure P3.72, the GaAs MESFET threshold voltage is $V_{TN} = 0.15 \text{ V}$. Let $R_D = 50 \text{ k}\Omega$. Determine the value of the conduction parameter required so that $V_O = 0.70 \text{ V}$ when $V_I = 0.75 \text{ V}$.

COMPUTER SIMULATION PROBLEMS

- 3.73 Using a computer simulation, verify the results of Exercise Ex 3.5.
- 3.74 (a) Using a computer simulation, plot the voltage transfer characteristics of the CMOS circuit shown in Figure 3.41. Use the parameters given in Example 3.11. (b) Repeat part (a) for the case when the width-to-length ratio of M_N is doubled.
- 3.75 (a) Using a computer simulation, plot the voltage transfer characteristics of the NMOS circuit shown in Figure 3.46 for $V_2 = 0$ and $0 \leq V_1 \leq 5 \text{ V}$. Use the circuit and transistor parameters given in Example 3.13. (b) Repeat part (a) for $0 \leq V_1 = V_2 \leq 5 \text{ V}$.
- 3.76 Using a computer simulation, verify the results of Example 3.17 for the multitransistor circuit shown in Figure 3.52.

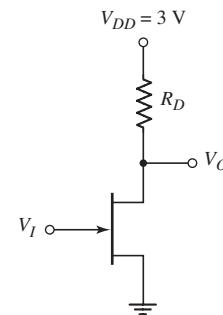


Figure P3.72

 DESIGN PROBLEMS

[Note: All design should be correlated with a computer simulation.]

- *D3.77 Consider the PMOS circuit shown in Figure 3.30. The circuit is to be redesigned such that $I_{DQ} = 100 \mu\text{A}$ and the Q-point is in the center of the saturation region of the load line. Assume $R_1 + R_2 = 500 \text{ k}\Omega$ and assume the same transistor parameters as given in Exercise Ex 3.6.
- *D3.78 Consider the circuit in Figure 3.39 with a depletion load. Assume the circuit is biased at $V_{DD} = 3.3 \text{ V}$, and assume transistor threshold voltages of $V_{TND} = 0.4 \text{ V}$ and $V_{TNL} = -0.75 \text{ V}$. Also assume $k'_n = 80 \mu\text{A/V}^2$. Design the circuit such that $V_O = 0.05 \text{ V}$ when $V_I = 3.3 \text{ V}$ and that the maximum power dissipation is $150 \mu\text{W}$.
- *D3.79 The constant-current source in Figure 3.50 is to be redesigned. The bias voltages are $V^+ = 3.3 \text{ V}$ and $V^- = -3.3 \text{ V}$. The parameters of all transistors are $V_{TN} = 0.4 \text{ V}$ and $k'_n = 100 \mu\text{A/V}^2$. The currents are to be $I_{REF} = 100 \mu\text{A}$ and $I_Q = 60 \mu\text{A}$. We will also specify that $V_{DS2(\text{sat})} = 0.6 \text{ V}$, $V_{GS1} = V_{GS2}$, and $V_{DS1} = 2.5 \text{ V}$. Determine all width-to-length ratios and the value of R_D . (Note: the minimum width-to-length ratio is to be greater than one.)
- *D3.80 Consider the multitransistor circuit in Figure 3.52. The bias voltages are changed to $V^+ = 3.3 \text{ V}$ and $V^- = -3.3 \text{ V}$. The transistor parameters are $V_{TN} = 0.4 \text{ V}$ and $k'_n = 100 \mu\text{A/V}^2$. Design the circuit such that $I_{DQ1} = 100 \mu\text{A}$, $I_{DQ2} = 250 \mu\text{A}$, $V_{DSQ1} = V_{DSQ2} = 3.3 \text{ V}$, and $R_i = 200 \text{ k}\Omega$.

Basic FET Amplifiers

In Chapter 3, we described the structure and operation of the FET, in particular the MOSFET, and analyzed and designed the dc response of circuits containing these devices. In this chapter, we emphasize the use of the FETs in linear amplifier applications. Linear amplifiers imply that, for the most part, we are dealing with analog signals. The magnitude of an analog signal may have any value, within limits, and may vary continuously with respect to time. Although a major use of MOSFETs is in digital applications, they are also used in linear amplifier circuits.

We will begin to see how all-transistor circuits, that is, circuits with no resistors, can be designed. Since MOS transistors are small devices, high-density all-transistor circuits can be fabricated as integrated circuits.

PREVIEW

In this chapter, we will:

- Investigate the process by which a single-MOS transistor circuit can amplify a small, time-varying input signal.
- Develop the small-signal models of the transistor that are used in the analysis of linear amplifiers.
- Discuss the three basic transistor amplifier configurations.
- Analyze the common-source, source-follower, and common-gate amplifiers, and become familiar with the general characteristics of these circuits.
- Compare the general characteristics of the three basic amplifier configurations.
- Analyze all-MOS transistor circuits that become the foundation of integrated circuits.
- Analyze multitransistor or multistage amplifiers and understand the advantages of these circuits over single-transistor amplifiers.
- Develop the small-signal model of JFET devices and analyze basic JFET amplifiers.
- As an application, incorporate MOS transistors in a design of a two-stage amplifier.

4.1 THE MOSFET AMPLIFIER

Objective: • Investigate the process by which a single-transistor circuit can amplify a small, time-varying input signal and develop the small-signal models of the transistor that are used in the analysis of linear amplifiers.

In this chapter, we will be considering **signals**, **analog circuits**, and **amplifiers**. A signal contains some type of information. For example, sound waves produced by a speaking human contain the information the person is conveying to another person. A sound wave is an analog signal. The magnitude of an **analog signal** can take on any value, within limits, and may vary continuously with time. Electronic circuits that process analog signals are called analog circuits. One example of an analog circuit is a linear amplifier. A **linear amplifier** magnifies an input signal and produces an output signal whose magnitude is larger and directly proportional to the input signal.

In this chapter, we analyze and design linear amplifiers that use field-effect transistors as the amplifying device. The term **small signal** means that we can linearize the ac equivalent circuit. We will define what is meant by small signal in the case of MOSFET circuits. The term linear amplifiers means that we can use superposition so that the dc analysis and ac analysis of the circuits can be performed separately and the total response is the sum of the two individual responses.

The mechanism with which MOSFET circuits amplify small time-varying signals was introduced in the last chapter. In this section, we will expand that discussion using the graphical technique, dc load line, and ac load line. In the process, we will develop the various small-signal parameters of linear circuits and the corresponding equivalent circuits.

4.1.1 Graphical Analysis, Load Lines, and Small-Signal Parameters

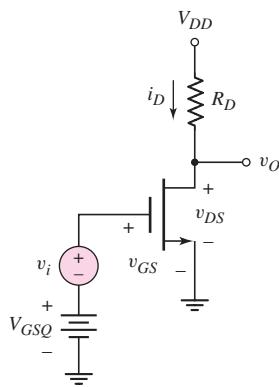


Figure 4.1 NMOS common-source circuit with time-varying signal source in series with gate dc source

Figure 4.1 shows an NMOS common-source circuit with a time-varying voltage source in series with the dc source. We assume the time-varying input signal is sinusoidal. Figure 4.2 shows the transistor characteristics, dc load line, and *Q*-point, where the dc load line and *Q*-point are functions of v_{GS} , V_{DD} , R_D , and the transistor parameters. For the output voltage to be a linear function of the input voltage, the transistor must be biased in the saturation region. (Note that, although we primarily use n-channel, enhancement-mode MOSFETs in our discussions, the same results apply to the other MOSFETs.)

Also shown in Figure 4.2 are the sinusoidal variations in the gate-to-source voltage, drain current, and drain-to-source voltage, as a result of the sinusoidal source v_i . The total gate-to-source voltage is the sum of V_{GSQ} and v_i . As v_i increases, the instantaneous value of v_{GS} increases, and the bias point moves up the load line. A larger value of v_{GS} means a larger drain current and a smaller value of v_{DS} . For a negative v_i (the negative portion of the sine wave), the instantaneous value of v_{GS} decreases below the quiescent value, and the bias point moves down the load line. A smaller v_{GS} value means a smaller drain current and increased value of v_{DS} . Once the

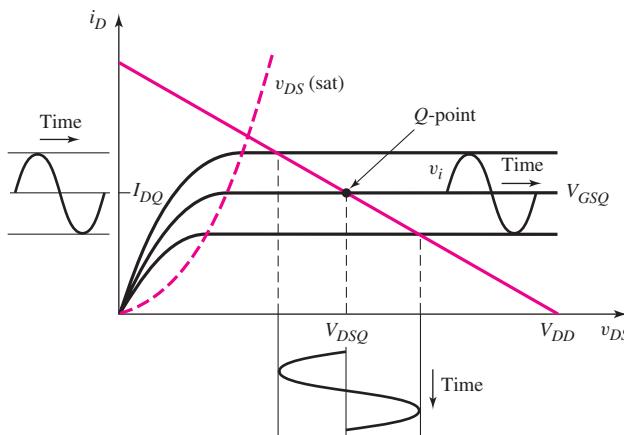


Figure 4.2 Common-source transistor characteristics, dc load line, and sinusoidal variation in gate-to-source voltage, drain current, and drain-to-source voltage

Q-point is established, we can develop a mathematical model for the sinusoidal, or small-signal, variations in gate-to-source voltage, drain-to-source voltage, and drain current.

The time-varying signal source v_i in Figure 4.1 generates a time-varying component of the gate-to-source voltage. In this case, $v_{gs} = v_i$, where v_{gs} is the time-varying component of the gate-to-source voltage. For the FET to operate as a linear amplifier, the transistor must be biased in the saturation region, and the instantaneous drain current and drain-to-source voltage must also be confined to the saturation region.

When symmetrical sinusoidal signals are applied to the input of an amplifier, symmetrical sinusoidal signals are generated at the output, as long as the amplifier operation remains linear. We can use the load line to determine the maximum output symmetrical swing. If the output exceeds this limit, a portion of the output signal will be clipped and signal distortion will occur.

In the case of FET amplifiers, the output signal must avoid cutoff ($i_D = 0$) and must stay in the saturation region ($v_{DS} > v_{DS}(\text{sat})$). This maximum range of output signal can be determined from the load line in Figure 4.2.

Transistor Parameters

We will be dealing with time-varying as well as dc currents and voltages in this chapter. Table 4.1 gives a summary of notation that will be used. This notation was

Table 4.1 Summary of notation

Variable	Meaning
i_D, v_{GS}	Total instantaneous values
I_D, V_{GS}	DC values
i_d, v_{gs}	Instantaneous ac values
I_d, V_{gs}	Phasor values

discussed in the Prologue, but is repeated here for convenience. A lowercase letter with an upper case subscript, such as i_D or v_{GS} , indicates a *total instantaneous value*. An uppercase letter with an uppercase subscript, such as I_D or V_{GS} , indicates a *dc quantity*. A lowercase letter with a lowercase subscript, such as i_d and v_{gs} , indicates an instantaneous value of an *ac signal*. Finally, an uppercase letter with a lowercase subscript, such as I_d or V_{gs} , indicates a *phasor quantity*. The phasor notation, which is also reviewed in the Prologue, becomes especially important in Chapter 7 during the discussion of frequency response. However, the phasor notation will generally be used in this chapter in order to be consistent with the overall ac analysis.

From Figure 4.1, we see that the instantaneous gate-to-source voltage is

$$v_{GS} = V_{GSQ} + v_i = V_{GSQ} + v_{gs} \quad (4.1)$$

where V_{GSQ} is the dc component and v_{gs} is the ac component. The instantaneous drain current is

$$i_D = K_n(v_{GS} - V_{TN})^2 \quad (4.2)$$

Substituting Equation (4.1) into (4.2) produces

$$i_D = K_n[V_{GSQ} + v_{gs} - V_{TN}]^2 = K_n[(V_{GSQ} - V_{TN}) + v_{gs}]^2 \quad (4.3(a))$$

or

$$i_D = K_n(V_{GSQ} - V_{TN})^2 + 2K_n(V_{GSQ} - V_{TN})v_{gs} + K_n v_{gs}^2 \quad (4.3(b))$$

The first term in Equation (4.3(b)) is the dc or quiescent drain current I_{DQ} , the second term is the time-varying drain current component that is linearly related to the signal v_{gs} , and the third term is proportional to the square of the signal voltage. For a sinusoidal input signal, the squared term produces undesirable harmonics, or non-linear distortion, in the output voltage. To minimize these harmonics, we require

$$v_{gs} \ll 2(V_{GSQ} - V_{TN}) \quad (4.4)$$

which means that the third term in Equation (4.3(b)) will be much smaller than the second term. *Equation (4.4) represents the small-signal condition that must be satisfied for linear amplifiers.*

Neglecting the v_{gs}^2 term, we can write Equation (4.3(b)) as

$$i_D = I_{DQ} + i_d \quad (4.5)$$

Again, small-signal implies linearity so that the total current can be separated into a dc component and an ac component. The ac component of the drain current is given by

$$i_d = 2K_n(V_{GSQ} - V_{TN})v_{gs} \quad (4.6)$$

The small-signal drain current is related to the small-signal gate-to-source voltage by the transconductance g_m . The relationship is

$$g_m = \frac{i_d}{v_{gs}} = 2K_n(V_{GSQ} - V_{TN}) \quad (4.7)$$

The transconductance is a transfer coefficient relating output current to input voltage and can be thought of as representing the gain of the transistor.

The transconductance can also be obtained from the derivative

$$g_m = \left. \frac{\partial i_D}{\partial v_{GS}} \right|_{v_{GS}=V_{GSQ}=\text{const.}} = 2K_n(V_{GSQ} - V_{TN}) \quad (4.8(a))$$

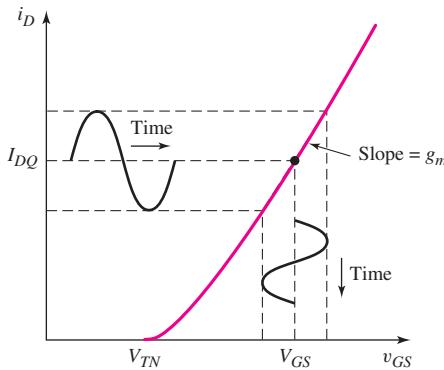


Figure 4.3 Drain current versus gate-to-source voltage characteristics, with superimposed sinusoidal signals

which can be written as

$$g_m = 2\sqrt{K_n I_{DQ}} \quad (4.8(b))$$

The drain current versus gate-to-source voltage for the transistor biased in the saturation region is given in Equation (4.2) and is shown in Figure 4.3. The transconductance g_m is the slope of the curve. If the time-varying signal v_{gs} is sufficiently small, the transconductance g_m is a constant. With the Q -point in the saturation region, the transistor operates as a current source that is linearly controlled by v_{gs} . If the Q -point moves into the nonsaturation region, the transistor no longer operates as a linearly controlled current source.

As shown in Equation (4.8(a)), the transconductance is directly proportional to the conduction parameter K_n , which in turn is a function of the width-to-length ratio. Therefore, increasing the width of the transistor increases the transconductance, or gain, of the transistor.

EXAMPLE 4.1

Objective: Calculate the transconductance of an n-channel MOSFET.

Consider an n-channel MOSFET with parameters $V_{TN} = 0.4$ V, $k'_n = 100 \mu\text{A/V}^2$, and $W/L = 25$. Assume the drain current is $I_D = 0.40$ mA.

Solution: The conduction parameter is

$$K_n = \frac{k'_n}{2} \cdot \frac{W}{L} = \left(\frac{0.1}{2}\right)(25) = 1.25 \text{ mA/V}^2$$

Assuming the transistor is biased in the saturation region, the transconductance is determined from Equation (4.8(b)) as

$$g_m = 2\sqrt{K_n I_{DQ}} = 2\sqrt{(1.25)(0.4)} = 1.41 \text{ mA/V}$$

Comment: The value of the transconductance can be increased by increasing the transistor W/L ratio and also by increasing the quiescent drain current.

EXERCISE PROBLEM

Ex 4.1: For an n-channel MOSFET biased in the saturation region, the parameters are $k'_n = 100 \mu\text{A/V}^2$, $V_{TN} = 0.6 \text{ V}$, and $I_{DQ} = 0.8 \text{ mA}$. Determine the transistor width-to-length ratio such that the transconductance is $g_m = 1.8 \text{ mA/V}$. (Ans. 20.25)

AC Equivalent Circuit

From Figure 4.1, we see that the output voltage is

$$v_{DS} = v_O = V_{DD} - i_D R_D \quad (4.9)$$

Using Equation (4.5), we obtain

$$v_O = V_{DD} - (I_{DQ} + i_d)R_D = (V_{DD} - I_{DQ}R_D) - i_d R_D \quad (4.10)$$

The output voltage is also a combination of dc and ac values. The time-varying output signal is the time-varying drain-to-source voltage, or

$$v_o = v_{ds} = -i_d R_D \quad (4.11)$$

Also, from Equations (4.6) and (4.7), we have

$$i_d = g_m v_{gs} \quad (4.12)$$

In summary, the following relationships exist between the time-varying signals for the circuit in Figure 4.1. The equations are given in terms of the instantaneous ac values, as well as the phasors. We have

$$v_{gs} = v_i \quad \text{or} \quad V_{gs} = V_i \quad (4.13)$$

and

$$i_d = g_m v_{gs} \quad \text{or} \quad I_d = g_m V_{gs} \quad (4.14)$$

and

$$v_{ds} = -i_d R_D \quad \text{or} \quad V_{ds} = -I_d R_D \quad (4.15)$$

The ac equivalent circuit in Figure 4.4 is developed by setting the dc sources in Figure 4.1 equal to zero. The small-signal relationships are given in Equations (4.13), (4.14), and (4.15). As shown in Figure 4.1, the drain current, which is composed of ac signals superimposed on the quiescent value, flows through the voltage source V_{DD} . Since the voltage across this source is assumed to be constant, the sinusoidal current produces no sinusoidal voltage component across this element. The equivalent ac impedance is therefore zero, or a short circuit. Consequently, in the ac equivalent circuit, the dc voltage sources are equal to zero. We say that the node connecting R_D and V_{DD} is at signal ground.

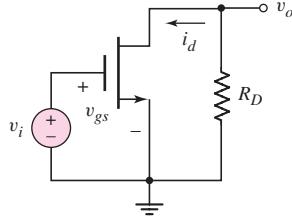


Figure 4.4 AC equivalent circuit of common-source amplifier with NMOS transistor

4.1.2 Small-Signal Equivalent Circuit

Now that we have the ac equivalent circuit for the NMOS amplifier circuit, (Figure 4.4), we must develop a small-signal equivalent circuit for the transistor.

Initially, we assume that the signal frequency is sufficiently low so that any capacitance at the gate terminal can be neglected. The input to the gate thus appears as an open circuit, or an infinite resistance. Equation (4.14) relates the small-signal

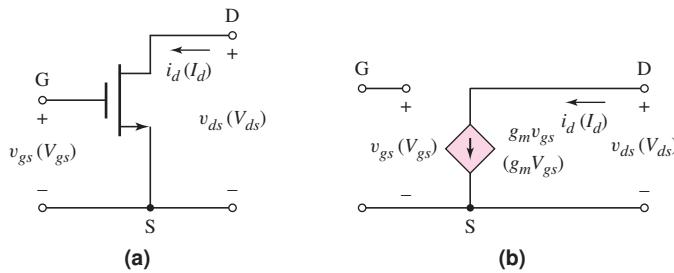


Figure 4.5 (a) Common-source NMOS transistor with small-signal parameters and (b) simplified small-signal equivalent circuit for NMOS transistor

drain current to the small-signal input voltage, and Equation (4.7) shows that the transconductance g_m is a function of the Q -point. The resulting simplified small-signal equivalent circuit for the NMOS device is shown in Figure 4.5. (The phasor components are in parentheses.)

This small-signal equivalent circuit can also be expanded to take into account the finite output resistance of a MOSFET biased in the saturation region. This effect, discussed in the last chapter, is a result of the nonzero slope in the i_D versus v_{DS} curve.

We know that

$$i_D = K_n[(v_{GS} - V_{TN})^2(1 + \lambda v_{DS})] \quad (4.16)$$

where λ is the **channel-length modulation parameter** and is a positive quantity. The small-signal output resistance, as previously defined, is

$$r_o = \left(\frac{\partial i_D}{\partial v_{DS}} \right)^{-1} \Big|_{v_{GS}=V_{GSQ}=\text{const.}} \quad (4.17)$$

or

$$r_o = [\lambda K_n(V_{GSQ} - V_{TN})^2]^{-1} \cong [\lambda I_{DQ}]^{-1} \quad (4.18)$$

This small-signal output resistance is also a function of the Q -point parameters.

The expanded small-signal equivalent circuit of the n-channel MOSFET is shown in Figure 4.6 in phasor notation. Note that this equivalent circuit is a transconductance amplifier in that the input signal is a voltage and the output signal is a current. This equivalent circuit can now be inserted into the amplifier ac equivalent circuit in Figure 4.4 to produce the circuit in Figure 4.7.

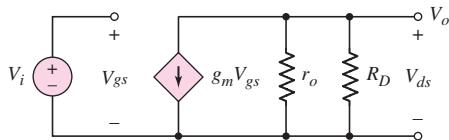


Figure 4.7 Small-signal equivalent circuit of common-source circuit with NMOS transistor model

EXAMPLE 4.2

Objective: Determine the small-signal voltage gain of a MOSFET circuit.

For the circuit in Figure 4.1, assume parameters are: $V_{GSQ} = 2.12$ V, $V_{DD} = 5$ V, and $R_D = 2.5$ k Ω . Assume transistor parameters are: $V_{TN} = 1$ V, $K_n = 0.80$ mA/V 2 , and $\lambda = 0.02$ V $^{-1}$. Assume the transistor is biased in the saturation region.

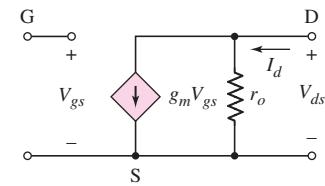


Figure 4.6 Expanded small-signal equivalent circuit, including output resistance, for NMOS transistor

Solution: The quiescent values are

$$I_{DQ} \cong K_n(V_{GSQ} - V_{TN})^2 = (0.8)(2.12 - 1)^2 = 1.0 \text{ mA}$$

and

$$V_{DSQ} = V_{DD} - I_{DQ}R_D = 5 - (1)(2.5) = 2.5 \text{ V}$$

Therefore,

$$V_{DSQ} = 2.5 \text{ V} > V_{DS(\text{sat})} = V_{GS} - V_{TN} = 1.82 - 1 = 0.82 \text{ V}$$

which means that the transistor is biased in the saturation region, as initially assumed, and as required for a linear amplifier. The transconductance is

$$g_m = 2K_n(V_{GSQ} - V_{TN}) = 2(0.8)(2.12 - 1) = 1.79 \text{ mA/V}$$

and the output resistance is

$$r_o = [\lambda I_{DQ}]^{-1} = [(0.02)(1)]^{-1} = 50 \text{ k}\Omega$$

From Figure 4.7, the output voltage is

$$V_o = -g_m V_{gs} (r_o \parallel R_D)$$

Since $V_{gs} = V_i$, the small-signal voltage gain is

$$A_v = \frac{V_o}{V_i} = -g_m (r_o \parallel R_d) = -(1.79)(50 \parallel 2.5) = -4.26$$

Comment: The magnitude of the ac output voltage is 4.26 times larger than the magnitude of the input voltage. Hence, we have an amplifier. Note that the small-signal voltage gain contains a minus sign, which means that the sinusoidal output voltage is 180 degrees out of phase with respect to the input sinusoidal signal.

EXERCISE PROBLEM

Ex 4.2: For the circuit shown in Figure 4.1, $V_{DD} = 3.3 \text{ V}$ and $R_D = 10 \text{ k}\Omega$. The transistor parameters are $V_{TN} = 0.4 \text{ V}$, $k'_n = 100 \mu\text{A/V}^2$, $W/L = 50$, and $\lambda = 0.025 \text{ V}^{-1}$. Assume the transistor is biased such that $I_{DQ} = 0.25 \text{ mA}$. (a) Verify that the transistor is biased in the saturation region. (b) Determine the small-signal parameters g_m and r_o . (c) Determine the small-signal voltage gain. (Ans. (a) $V_{GSQ} = 0.716 \text{ V}$ and $V_{DSQ} = 0.8 \text{ V}$ so that $V_{DS} > V_{DS(\text{sat})}$; (b) $g_m = 1.58 \text{ mA/V}$, $r_o = 160 \text{ k}\Omega$; (c) -14.9)

Problem-Solving Technique: MOSFET AC Analysis

Since we are dealing with linear amplifiers, superposition applies, which means that we can perform the dc and ac analyses separately. The analysis of the MOSFET amplifier proceeds as follows:

1. Analyze the circuit with only the dc sources present. This solution is the dc or quiescent solution. The transistor must be biased in the saturation region in order to produce a linear amplifier.
2. Replace each element in the circuit with its small-signal model, which means replacing the transistor by its small-signal equivalent circuit.
3. Analyze the small-signal equivalent circuit, setting the dc source components equal to zero, to produce the response of the circuit to the time-varying input signals only.

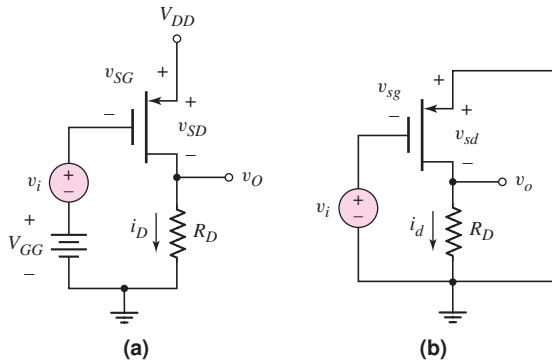


Figure 4.8 (a) Common-source circuit with PMOS transistor and (b) corresponding ac equivalent circuit

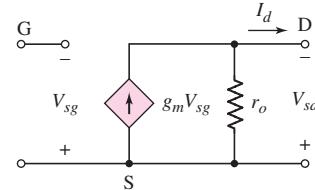


Figure 4.9 Small-signal equivalent circuit of PMOS transistor

The previous discussion was for an n-channel MOSFET amplifier. The same basic analysis and equivalent circuit also applies to the p-channel transistor. Figure 4.8(a) shows a circuit containing a p-channel MOSFET. Note that the power supply voltage \$V_{DD}\$ is connected to the source. (The subscript \$DD\$ can be used to indicate that the supply is connected to the drain terminal. Here, however, \$V_{DD}\$ is simply the usual notation for the power supply voltage in MOSFET circuits.) Also note the change in current directions and voltage polarities compared to the circuit containing the NMOS transistor. Figure 4.8(b) shows the ac equivalent circuit, with the dc voltage sources replaced by ac short circuits, and all currents and voltages shown are the time-varying components.

In the circuit of Figure 4.8(b), the transistor can be replaced by the equivalent circuit in Figure 4.9. The equivalent circuit of the p-channel MOSFET is the same as that of the n-channel device, except that all current directions and voltage polarities are reversed.

The final small-signal equivalent circuit of the p-channel MOSFET amplifier is shown in Figure 4.10. The output voltage is

$$V_o = g_m V_{sg} (r_o \parallel R_D) \quad (4.19)$$

The control voltage \$V_{sg}\$, given in terms of the input signal voltage, is

$$V_{sg} = -V_i \quad (4.20)$$

and the small-signal voltage gain is

$$A_v = \frac{V_o}{V_i} = -g_m (r_o \parallel R_D) \quad (4.21)$$

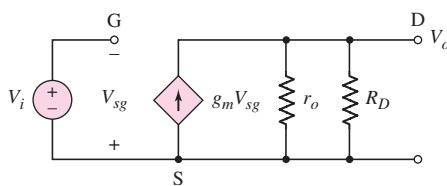


Figure 4.10 Small-signal equivalent circuit of common-source amplifier with PMOS transistor model

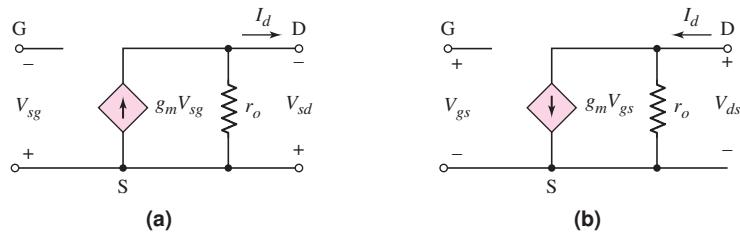


Figure 4.11 Small signal equivalent circuit of a p-channel MOSFET showing (a) the conventional voltage polarities and current directions and (b) the case when the voltage polarities and current directions are reversed.

This expression for the small-signal voltage gain of the p-channel MOSFET amplifier is exactly the same as that for the n-channel MOSFET amplifier. The negative sign indicates that a 180-degree phase reversal exists between the output and input signals, for both the PMOS and the NMOS circuit.

We may note that if the polarity of the small-signal gate-to-source voltage is reversed, then the small-signal drain current direction is reversed. This change of polarity is shown in Figure 4.11. Figure 4.11(a) shows the conventional voltage polarity and current directions in a PMOS transistor. If the control voltage polarity is reversed as shown in Figure 4.11(b), then the dependent current direction is also reversed. The equivalent circuit shown in Figure 4.11(b) is then the same as that of the NMOS transistor. However, the author prefers to use the small-signal equivalent circuit in Figure 4.9 to be consistent with the voltage polarities and current directions of the PMOS transistor.

4.1.3 Modeling the Body Effect

As mentioned in Section 3.1.9, Chapter 3, the body effect occurs in a MOSFET in which the substrate, or body, is not directly connected to the source. For an NMOS device, the body is connected to the most negative potential in the circuit and will be at signal ground. Figure 4.12(a) shows the four-terminal MOSFET with dc voltages and Figure 4.12(b) shows the device with ac voltages. Keep in mind that v_{SB} must be greater than or equal to zero. The simplified current-voltage relation is

$$i_D = K_n(v_{GS} - V_{TN})^2 \quad (4.22)$$

and the threshold voltage is given by

$$V_{TN} = V_{TNO} + \gamma [\sqrt{2\phi_f + v_{SB}} - \sqrt{2\phi_f}] \quad (4.23)$$

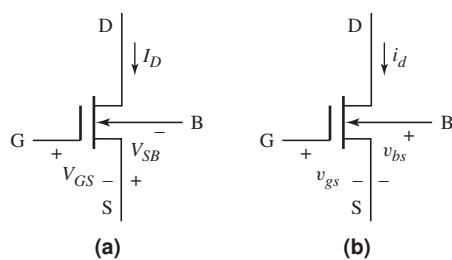


Figure 4.12 The four-terminal NMOS device with (a) dc voltages and (b) ac voltages

If an ac component exists in the source-to-body voltage, v_{SB} , there will be an ac component induced in the threshold voltage, which causes an ac component in the drain current. Thus, a back-gate transconductance can be defined as

$$g_{mb} = \frac{\partial i_D}{\partial v_{BS}} \Big|_{Q-pt} = \frac{-\partial i_D}{\partial v_{SB}} \Big|_{Q-pt} = -\left(\frac{\partial i_D}{\partial V_{TN}} \right) \cdot \left(\frac{\partial V_{TN}}{\partial v_{SB}} \right) \Big|_{Q-pt} \quad (4.24)$$

Using Equation (4.22), we find

$$\frac{\partial i_D}{\partial V_{TN}} = -2K_n(v_{GS} - V_{TN}) = -g_m \quad (4.25(a))$$

and using Equation (4.23), we find

$$\frac{\partial V_{TN}}{\partial v_{SB}} = \frac{\gamma}{2\sqrt{2\phi_f + v_{SB}}} \equiv \eta \quad (4.25(b))$$

The back-gate transconductance is then

$$g_{mb} = -(-g_m) \cdot (\eta) = g_m \eta \quad (4.26)$$

Including the body effect, the small-signal equivalent circuit of the MOSFET is shown in Figure 4.13. We note the direction of the current and the polarity of the small-signal source-to-body voltage. If $v_{bs} > 0$, then v_{SB} decreases, V_{TN} decreases, and i_D increases. The current direction and voltage polarity are thus consistent.

For $\phi_f = 0.35$ V and $\gamma = 0.35$ V $^{1/2}$, the value of η from Equation (4.25(b)) is $\eta \cong 0.23$. Therefore, η will be in the range $0 \leq \eta \leq 0.23$. The value of v_{bs} will depend on the particular circuit.

In general, we will neglect g_{mb} in our hand analyses and designs, but will investigate the body effect in PSpice analyses.

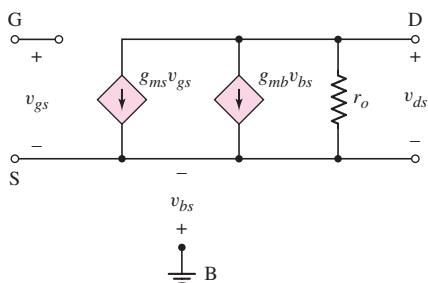


Figure 4.13 Small-signal equivalent circuit of NMOS device including body effect

Test Your Understanding

TYU 4.1 The parameters of an n-channel MOSFET are: $V_{TN} = 0.6$ V, $k'_n = 100 \mu\text{A/V}^2$, and $\lambda = 0.015 \text{ V}^{-1}$. The transistor is biased in the saturation region with $I_{DQ} = 1.2$ mA. (a) Design the width-to-length ratio such that the transconductance is $g_m = 2.5$ mA/V. (b) Determine the small-signal output resistance r_o . (Ans. (a) 26.0, (b) 55.6 k Ω).

TYU 4.2 For the circuit shown in Figure 4.1, $V_{DD} = 3.3$ V and $R_D = 8$ k Ω . The transistor parameters are $V_{TN} = 0.4$ V, $K_n = 0.5$ mA/V 2 , and $\lambda = 0.02 \text{ V}^{-1}$. (a) Determine V_{GSQ} and V_{DSQ} for $I_{DQ} = 0.15$ mA. (b) Calculate g_m , r_o , and the small-signal voltage gain. (Ans. (a) $V_{GSQ} = 0.948$ V, $V_{DSQ} = 2.1$ V; (b) $g_m = 0.548$ mA/V, $r_o = 333$ k Ω , $A_v = -4.28$).

TYU 4.3 For the circuit in Figure 4.1, the circuit and transistor parameters are given in Exercise TYU 4.2. If $v_i = 25 \sin \omega t$ (mV), determine i_D and v_{DS} . (Ans. $i_D = 0.15 + 0.0137 \sin \omega t$ (mA), $v_{DS} = 2.1 - 0.11 \sin \omega t$ (V)).

TYU 4.4 The parameters for the circuit in Figure 4.8 are $V_{DD} = 5$ V and $R_D = 5$ k Ω . The transistor parameters are $V_{TP} = -0.4$ V, $K_p = 0.4$ mA/V 2 , and $\lambda = 0$. (a) Determine V_{SGQ} and I_{DQ} such that $V_{SDQ} = 3$ V. (b) Calculate g_m and the small-signal voltage gain. (Ans. (a) $I_{DQ} = 0.4$ mA, $V_{SGQ} = 1.4$ V; (b) $g_m = 0.8$ mA/V, $A_v = -4$).

TYU 4.5 A transistor has the same parameters as those given in Exercise Ex4.1. In addition, the body effect coefficient is $\gamma = 0.40$ V $^{1/2}$ and $\phi_f = 0.35$ V. Determine the value of η and the back-gate transconductance g_{mb} for (a) $v_{SB} = 1$ V and (b) $v_{SB} = 3$ V. (Ans. (a) $\eta = 0.153$, (b) $\eta = 0.104$).



4.2 BASIC TRANSISTOR AMPLIFIER CONFIGURATIONS

Objective: • Discuss the three basic transistor amplifier configurations.

As we have seen, the MOSFET is a three-terminal device. Three basic single-transistor amplifier configurations can be formed, depending on which of the three transistor terminals is used as signal ground. These three basic configurations are appropriately called **common source**, **common drain (source follower)**, and **common gate**.

The input and output resistance characteristics of amplifiers are important in determining loading effects. These parameters, as well as voltage gain, for the three basic MOSFET circuit configurations will be determined in the following sections. The characteristics of the three types of amplifiers will then allow us to understand under what condition each amplifier is most useful.

Initially, we will consider MOSFET amplifier circuits that emphasize discrete designs, in that resistor biasing will be used. The purpose is to become familiar with basic MOSFET amplifier designs and their characteristics. In Section 4.7, we will begin to consider integrated circuit MOSFET designs that involve all-transistor circuits and current source biasing. These initial designs provide an introduction to more advanced MOS amplifier designs that will be considered in Part 2 of the text.



4.3 THE COMMON-SOURCE AMPLIFIER

Objective: • Analyze the common-source amplifier and become familiar with the general characteristics of this circuit.

In this section, we consider the first of the three basic circuits—the common-source amplifier. We will analyze several basic common-source circuits, and will determine small-signal voltage gain and input and output impedances.

4.3.1 A Basic Common-Source Configuration

Figure 4.14 shows the basic common-source circuit with voltage-divider biasing. We see that the source is at ground potential—hence the name common source. The signal from the signal source is coupled into the gate of the transistor through the coupling capacitor C_C , which provides dc isolation between the amplifier and the signal source. The dc transistor biasing is established by R_1 and R_2 , and is not disturbed when the signal source is capacitively coupled to the amplifier.

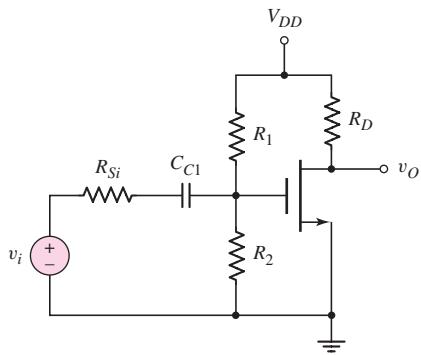


Figure 4.14 Common-source circuit with voltage divider biasing and coupling capacitor

If the signal source is a sinusoidal voltage at frequency f , then the magnitude of the capacitor impedance is $|Z_C| = [1/(2\pi f C_C)]$. For example, assume that $C_C = 10 \mu\text{F}$ and $f = 2 \text{ kHz}$. The magnitude of the capacitor impedance is then

$$|Z_C| = \frac{1}{2\pi f C_C} = \frac{1}{2\pi(2 \times 10^3)(10 \times 10^{-6})} \cong 8 \Omega$$

The magnitude of this impedance is generally much less than the Thevenin resistance at the capacitor terminals. We can therefore assume that the capacitor is essentially a short circuit to signals with frequencies greater than 2 kHz. We will also neglect, in this chapter, any capacitance effects within the transistor.

For the circuit shown in Figure 4.14, assume that the transistor is biased in the saturation region by resistors R_1 and R_2 , and that the signal frequency is sufficiently large for the coupling capacitor to act essentially as a short circuit. The signal source is represented by a Thevenin equivalent circuit, in which the signal voltage source v_i is in series with an equivalent source resistance R_{Si} . As we will see, R_{Si} should be much less than the amplifier input resistance, $R_i = R_1 \parallel R_2$, in order to minimize loading effects.

Figure 4.15 shows the resulting small-signal equivalent circuit. The small-signal variables, such as the input signal voltage V_i , are given in phasor form.

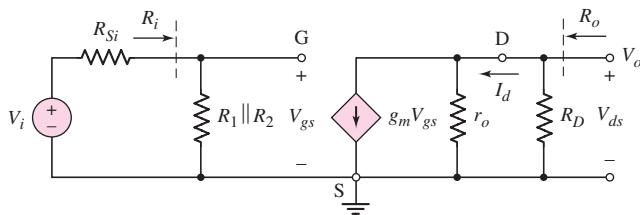


Figure 4.15 Small-signal equivalent circuit, assuming coupling capacitor acts as a short circuit

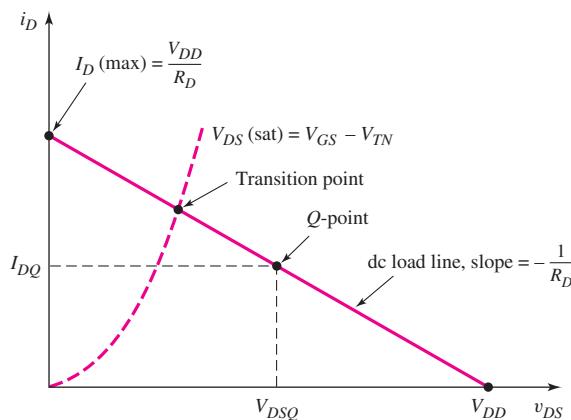


Figure 4.16 DC load line and transition point separating saturation and nonsaturation regions

Since the source is at ground potential, there is no body effect. The output voltage is

$$V_o = -g_m V_{gs} (r_o \parallel R_D) \quad (4.27)$$

The input gate-to-source voltage is

$$V_{gs} = \left(\frac{R_i}{R_i + R_{Si}} \right) \cdot V_i \quad (4.28)$$

so the small-signal voltage gain is

$$A_v = \frac{V_o}{V_i} = -g_m (r_o \parallel R_D) \cdot \left(\frac{R_i}{R_i + R_{Si}} \right) \quad (4.29)$$

We can also relate the ac drain current to the ac drain-to-source voltage, as $V_{ds} = -I_d(R_D)$.

Figure 4.16 shows the dc load line, the transition point (that separates the saturation bias region and nonsaturation bias region), and the *Q*-point, which is in the saturation region. In order to provide the maximum symmetrical output voltage swing and keep the transistor biased in the saturation region, the *Q*-point must be near the middle of the saturation region. At the same time, the input signal must be small enough for the amplifier to remain linear.

The input and output resistances of the amplifier can be determined from Figure 4.15. The input resistance to the amplifier is $R_i = R_1 \parallel R_2$. Since the low-frequency input resistance looking into the gate of the MOSFET is essentially infinite, the input resistance is only a function of the bias resistors. The output resistance looking back into the output terminals is found by setting the independent input source V_i equal to zero, which means that $V_{gs} = 0$. The output resistance is therefore $R_o = R_D \parallel r_o$.

EXAMPLE 4.3

Objective: Determine the small-signal voltage gain and input and output resistances of a common-source amplifier.

For the circuit shown in Figure 4.14, the parameters are: $V_{DD} = 3.3$ V, $R_D = 10$ k Ω , $R_1 = 140$ k Ω , $R_2 = 60$ k Ω , and $R_{Si} = 4$ k Ω . The transistor parameters are: $V_{TN} = 0.4$ V, $K_n = 0.5$ mA/V 2 , and $\lambda = 0.02$ V $^{-1}$.

Solution (dc calculations): The dc or quiescent gate-to-source voltage is

$$V_{GSQ} = \left(\frac{R_2}{R_1 + R_2} \right) (V_{DD}) = \left(\frac{60}{140 + 60} \right) (3.3) = 0.99 \text{ V}$$

The quiescent drain current is

$$I_{DQ} = K_n(V_{GSQ} - V_{TN})^2 = (0.5)(0.99 - 0.4)^2 = 0.174 \text{ mA}$$

and the quiescent drain-to-source voltage is

$$V_{DSQ} = V_{DD} - I_{DQ}R_D = 3.3 - (0.174)(10) = 1.56 \text{ V}$$

Since $V_{DSQ} > V_{GSQ} - V_{TN}$, the transistor is biased in the saturation region.

Small-signal Voltage Gain: The small-signal transconductance g_m is then

$$g_m = 2\sqrt{K_n I_{DQ}} = 2\sqrt{(0.5)(0.174)} = 0.590 \text{ mA/V}$$

and the small-signal output resistance is

$$r_o = \frac{1}{\lambda I_Q} = \frac{1}{(0.02)(0.174)} = 287 \text{ k}\Omega$$

The input resistance to the amplifier is

$$R_i = R_1 \| R_2 = 140 \| 60 = 42 \text{ k}\Omega$$

From Figure 4.15 and Equation (4.29), the small-signal voltage gain is

$$A_v = -g_m(r_o \| R_D) \left(\frac{R_i}{R_i + R_{Si}} \right) = -(0.59)(287 \| 10) \left(\frac{42}{42 + 4} \right)$$

or

$$A_v = -5.21$$

Input and Output Resistances: As already calculated, the amplifier input resistance is

$$R_i = R_1 \| R_2 = 140 \| 60 = 42 \text{ k}\Omega$$

and the amplifier output resistance is

$$R_o = R_D \| r_o = 10 \| 287 = 9.66 \text{ k}\Omega$$

Comment: The resulting Q -point is not in the center of the saturation region. Therefore, this circuit does not achieve the maximum symmetrical output voltage swing in this case.

Discussion: The small-signal input gate-to-source voltage is

$$V_{gs} = \left(\frac{R_i}{R_i + R_{Si}} \right) \cdot V_i = \left(\frac{42}{42 + 4} \right) \cdot V_i = (0.913) \cdot V_i$$

Since R_{Si} is not zero, the amplifier input signal V_{gs} is approximately 91 percent of the signal voltage. This is called a loading effect. Even though the input resistance to the gate of the transistor is essentially infinite, the bias resistors greatly influence the

amplifier input resistance and loading effect. This loading effect can be eliminated or minimized when current source biasing is considered.

EXERCISE PROBLEM

Ex 4.3: The parameters of the circuit shown in Figure 4.14 are $V_{DD} = 5$ V, $R_1 = 520 \text{ k}\Omega$, $R_2 = 320 \text{ k}\Omega$, $R_D = 10 \text{ k}\Omega$, and $R_{Si} = 0$. Assume transistor parameters of $V_{TN} = 0.8$ V, $K_n = 0.20 \text{ mA/V}^2$, and $\lambda = 0$. (a) Determine the small-signal transistor parameters g_m and r_o . (b) Find the small-signal voltage gain. (c) Calculate the input and output resistances R_i and R_o (see Figure 4.15). (Ans. (a) $g_m = 0.442 \text{ mA/V}$, $r_o = \infty$; (b) $A_v = -4.42$; (c) $R_i = 198 \text{ k}\Omega$, $R_o = R_D = 10 \text{ k}\Omega$)

DESIGN EXAMPLE 4.4

Objective: Design the bias of a MOSFET circuit such that the Q -point is in the middle of the saturation region. Determine the resulting small-signal voltage gain.

Specifications: The circuit to be designed has the configuration shown in Figure 4.17. Let $R_1 \parallel R_2 = 100 \text{ k}\Omega$. Design the circuit such that the Q -point is $I_{DQ} = 2 \text{ mA}$ and the Q -point is in the middle of the saturation region.

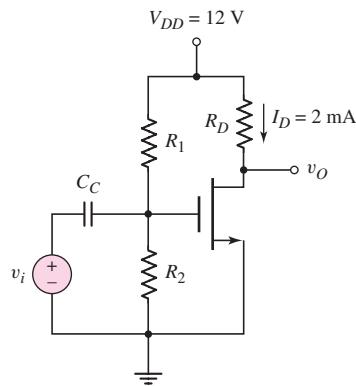


Figure 4.17 Common-source NMOS transistor circuit

Choices: A transistor with nominal parameters $V_{TN} = 1$ V, $k'_n = 80 \mu\text{A/V}^2$, $W/L = 25$, and $\lambda = 0.015 \text{ V}^{-1}$ is available.

Solution (dc design): The load line and the desired Q -point are given in Figure 4.18. If the Q -point is to be in the middle of the saturation region, the current at the transition point must be 4 mA.

The conductivity parameter is

$$K_n = \frac{k'_n}{2} \cdot \frac{W}{L} = \left(\frac{0.080}{2} \right) (25) = 1 \text{ mA/V}^2$$

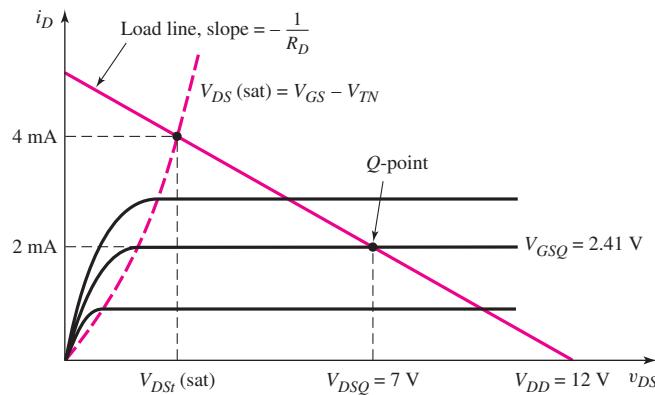


Figure 4.18 DC load line and transition point for NMOS circuit shown in Figure 4.17

We can now calculate $V_{DS}(\text{sat})$ at the transition point. The subscript t indicates transition point values. To determine V_{GS_t} , we use

$$I_{Dt} = 4 = K_n(V_{GS_t} - V_{TN})^2 = 1(V_{GS_t} - 1)^2$$

which yields

$$V_{GS_t} = 3 \text{ V}$$

Therefore

$$V_{DS_t} = V_{GS_t} - V_{TN} = 3 - 1 = 2 \text{ V}$$

If the Q -point is in the middle of the saturation region, then $V_{DSQ} = 7 \text{ V}$, which would yield a 10 V peak-to-peak symmetrical output voltage. From Figure 4.17, we can write

$$V_{DSQ} = V_{DD} - I_{DQ}R_D$$

or

$$R_D = \frac{V_{DD} - V_{DSQ}}{I_{DQ}} = \frac{12 - 7}{2} = 2.5 \text{ k}\Omega$$

We can determine the required quiescent gate-to-source voltage from the current equation, as follows:

$$I_{DQ} = 2 = K_n(V_{GSQ} - V_{TN})^2 = (1)(V_{GSQ} - 1)^2$$

or

$$V_{GSQ} = 2.41 \text{ V}$$

Then

$$\begin{aligned} V_{GSQ} &= 2.41 = \left(\frac{R_2}{R_1 + R_2} \right) (V_{DD}) = \left(\frac{1}{R_1} \right) \left(\frac{R_1 R_2}{R_1 + R_2} \right) (V_{DD}) \\ &= \frac{R_i}{R_1} \cdot V_{DD} = \frac{(100)(12)}{R_1} \end{aligned}$$

which yields

$$R_1 = 498 \text{ k}\Omega \quad \text{and} \quad R_2 = 125 \text{ k}\Omega$$

Solution (ac analysis): The small-signal transistor parameters are

$$g_m = 2\sqrt{K_n I_{DQ}} = 2\sqrt{(1)(2)} = 2.83 \text{ mA/V}$$

and

$$r_o = \frac{1}{\lambda I_{DQ}} = \frac{1}{(0.015)(2)} = 33.3 \text{ k}\Omega$$

The small-signal equivalent circuit is the same as shown in Figure 4.7. The small-signal voltage gain is

$$A_v = \frac{V_o}{V_i} = -g_m (r_o \parallel R_D) = -(2.83)(33.3 \parallel 2.5)$$

or

$$A_v = -6.58$$

Comment: Establishing the *Q*-point in the middle of the saturation region allows the maximum symmetrical swing in the output voltage, while keeping the transistor biased in the saturation region.

EXERCISE PROBLEM

Ex 4.4: Consider the circuit shown in Figure 4.14. Assume transistor parameters of $V_{TN} = 0.8 \text{ V}$, $K_n = 0.20 \text{ mA/V}^2$, and $\lambda = 0$. Let $V_{DD} = 5 \text{ V}$, $R_i = R_1 \parallel R_2 = 200 \text{ k}\Omega$, and $R_{Si} = 0$. Design the circuit such that $I_{DQ} = 0.5 \text{ mA}$ and the *Q*-point is in the center of the saturation region. Find the small-signal voltage gain. (Ans. $R_D = 2.76 \text{ k}\Omega$, $R_1 = 420 \text{ k}\Omega$, $R_2 = 382 \text{ k}\Omega$, $A_v = -1.75$)

4.3.2 Common-Source Amplifier with Source Resistor

A source resistor R_S tends to stabilize the *Q*-point against variations in transistor parameters (Figure 4.19). If, for example, the value of the conduction parameter varies from one transistor to another, the *Q*-point will not vary as much if a source

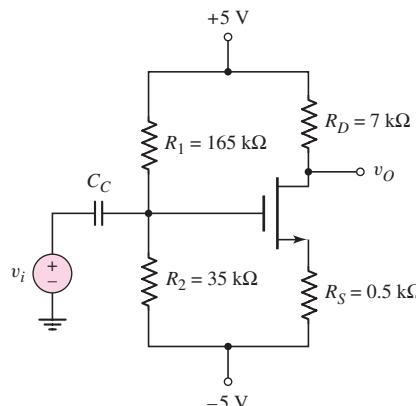


Figure 4.19 Common-source circuit with source resistor and positive and negative supply voltages

resistor is included in the circuit. However, as shown in the following example, a source resistor also reduces the signal gain.

The circuit in Figure 4.19 is an example of a situation in which the body effect should be taken into account. The substrate (not shown) would normally be connected to the -5 V supply, so that the body and substrate terminals are not at the same potential. However, in the following example, we will neglect this effect.

The circuit shown in Figure 4.20(a) is a PMOS version of the common-source amplifier with a source resistor R_S included.

EXAMPLE 4.5

Objective: Determine the small-signal voltage gain of a PMOS transistor circuit.

Consider the circuit shown in Figure 4.20(a). The transistor parameters are $K_p = 0.80 \text{ mA/V}^2$, $V_{TP} = -0.5 \text{ V}$, and $\lambda = 0$. The quiescent drain current is found to be $I_{DQ} = 0.297 \text{ mA}$.

The small-signal equivalent circuit is shown in Figure 4.20(b). To sketch the small-signal equivalent circuit, start with the three terminals of the transistor, draw in the transistor equivalent circuit between these three terminals, and then sketch in the other circuit elements around the transistor.

Solution: The small-signal output voltage is

$$V_o = +g_m V_{sg} R_D$$

Writing a KVL equation from the input around the gate–source loop, we find

$$V_i = -V_{sg} - g_m V_{sg} R_S$$

or

$$V_{sg} = \frac{-V_i}{1 + g_m R_S}$$

Substituting this expression for V_{sg} into the output voltage equation, we find the small-signal voltage gain as

$$A_v = \frac{V_o}{V_i} = \frac{-g_m R_D}{1 + g_m R_S}$$

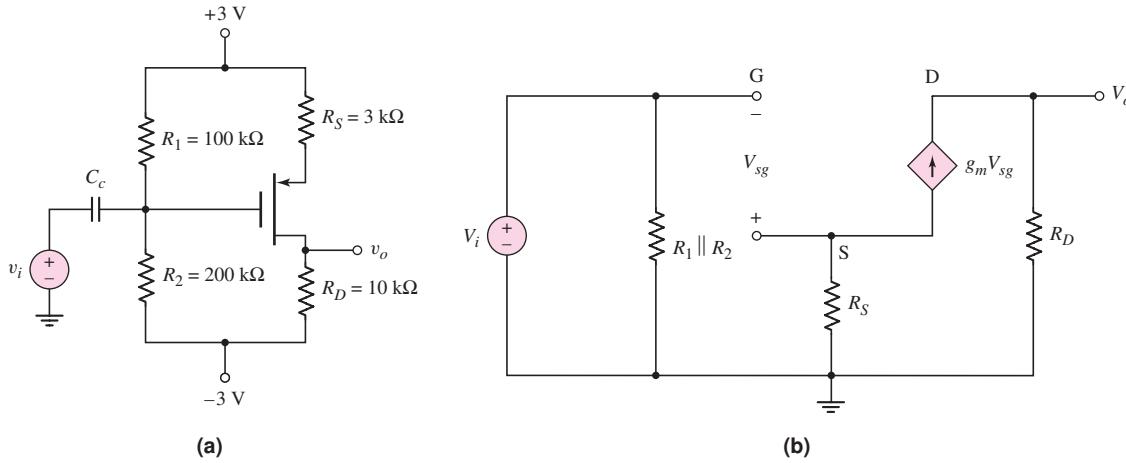


Figure 4.20 (a) PMOS circuit for Example 4.5, and (b) small-signal equivalent circuit

The small-signal transconductance is

$$g_m = 2\sqrt{K_p I_{DQ}} = 2\sqrt{(0.80)(0.297)} = 0.975 \text{ mA/V}$$

We then find the small-signal voltage gain as

$$A_v = \frac{-(0.975)(10)}{1 + (0.975)(3)} \quad \text{or} \quad A_v = -2.48$$

Comment: The analysis of a PMOS transistor circuit is essentially the same as that of an NMOS transistor circuit. The voltage gain of a MOS transistor circuit that contains a source resistor is degraded compared to a circuit without a source resistor. However, the Q -point tends to be stabilized.

Discussion: We mentioned that including a source resistor tends to stabilize the circuit characteristics against any changes in transistor parameters. If, for example, the conduction parameter K_p varies by ± 10 percent, we find the following results.

$K_p(\text{mA/V}^2)$	$g_m(\text{mA/V})$	$A_v(\text{V/V})$
0.72	0.9121	-2.441
0.80	0.9749	-2.484
0.88	1.035	-2.521

With a ± 10 percent variation in K_p , there is less than a ± 1.8 percent variation in the voltage gain.

EXERCISE PROBLEM

Ex 4.5: For the circuit shown in Figure 4.19, the transistor parameters are $V_{TN} = 0.8 \text{ V}$, $K_n = 1 \text{ mA/V}^2$, and $\lambda = 0$. (a) From the dc analysis, find I_{DQ} and V_{DSQ} . (b) Determine the small-signal voltage gain. (Ans. (a) $I_{DQ} = 0.494 \text{ mA}$, $V_{DSQ} = 6.30 \text{ V}$; (b) $A_v = -5.78$).

4.3.3

Common-Source Circuit with Source Bypass Capacitor

A source bypass capacitor added to the common-source circuit with a source resistor will minimize the loss in the small-signal voltage gain, while maintaining the Q -point stability. The Q -point stability can be further increased by replacing the source resistor with a constant-current source. The resulting circuit is shown in Figure 4.21, assuming an ideal signal source. If the signal frequency is sufficiently large so that the bypass capacitor acts essentially as an ac short-circuit, the source will be held at signal ground.

EXAMPLE 4.6

Objective: Determine the small-signal voltage gain of a circuit biased with a constant-current source and incorporating a source bypass capacitor.

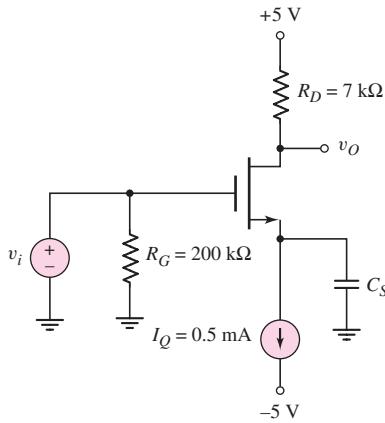


Figure 4.21 NMOS common-source circuit with source bypass capacitor

For the circuit shown in Figure 4.21, the transistor parameters are: $V_{TN} = 0.8 \text{ V}$, $K_n = 1 \text{ mA/V}^2$, and $\lambda = 0$.

Solution: Since the dc gate current is zero, the dc voltage at the source terminal is $V_S = -V_{GSQ}$, and the gate-to-source voltage is determined from

$$I_{DQ} = I_Q = K_n(V_{GSQ} - V_{TN})^2$$

or

$$0.5 = (1)(V_{GSQ} - 0.8)^2$$

which yields

$$V_{GSQ} = -V_S = 1.51 \text{ V}$$

The quiescent drain-to-source voltage is

$$V_{DSQ} = V_{DD} - I_{DQ}R_D - V_S = 5 - (0.5)(7) - (-1.51) = 3.01 \text{ V}$$

The transistor is therefore biased in the saturation region.

The small-signal equivalent circuit is shown in Figure 4.22. The output voltage is

$$V_o = -g_m V_{gs} R_D$$

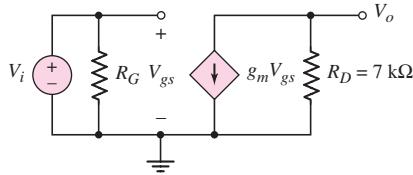


Figure 4.22 Small-signal equivalent circuit, assuming the source bypass capacitor acts as a short circuit

Since $V_{gs} = V_i$, the small-signal voltage gain is

$$A_v = \frac{V_o}{V_i} = -g_m R_D = -(1.414)(7) = -9.9$$

Comment: Comparing the small-signal voltage gain of 9.9 in this example to the 2.48 calculated in Example 4.5, we see that the magnitude of the gain increases when a source bypass capacitor is included.

EXERCISE PROBLEM

Ex 4.6: The common-source amplifier in Figure 4.23 has transistor parameters $k'_p = 40 \mu\text{A/V}^2$, $W/L = 40$, $V_{TP} = -0.4 \text{ V}$, and $\lambda = 0.02 \text{ V}^{-1}$. (a) Determine I_{DQ} and V_{SDQ} . (b) Find the small-signal voltage gain. (Ans. (a) $I_{DQ} = 1.16 \text{ mA}$, $V_{SDQ} = 2.29 \text{ V}$; (b) $A_v = -3.68$)

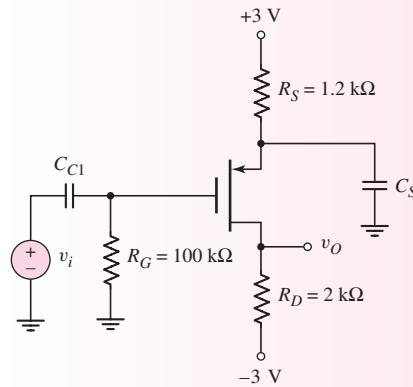


Figure 4.23 Figure for Exercise Ex 4.6

Test Your Understanding

TYU 4.6 Consider the common-source amplifier in Figure 4.24 with transistor parameters $V_{TN} = 1.8 \text{ V}$, $K_n = 0.15 \text{ mA/V}^2$, and $\lambda = 0$. (a) Calculate I_{DQ} and V_{DSQ} . (b) Determine the small-signal voltage gain. (c) Discuss the purpose of R_G and its effect on the small-signal operation of the amplifier. (Ans. (a) $I_{DQ} = 1.05 \text{ mA}$, $V_{DSQ} = 4.45 \text{ V}$; (b) $A_v = -2.65$)

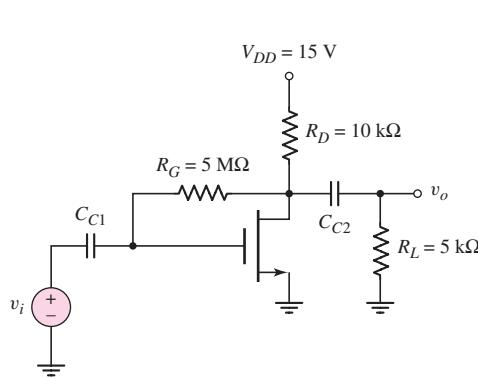


Figure 4.24 Figure for Exercise TYU 4.6

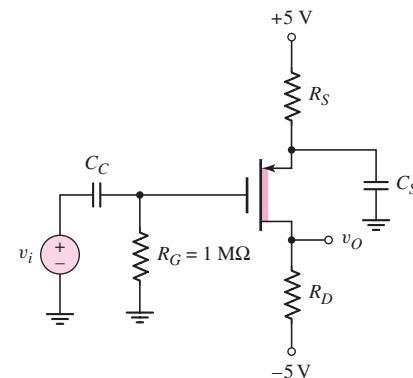


Figure 4.25 Figure for Exercise TYU 4.7

TYU 4.7 The parameters of the transistor shown in Figure 4.25 are: $V_{TP} = +0.8 \text{ V}$, $K_p = 0.5 \text{ mA/V}^2$, and $\lambda = 0.02 \text{ V}^{-1}$. (a) Determine R_S and R_D such that $I_{DQ} = 0.8 \text{ mA}$ and $V_{SDQ} = 3 \text{ V}$. (b) Find the small-signal voltage gain. (Ans. (a) $R_S = 5.67 \text{ k}\Omega$, $R_D = 3.08 \text{ k}\Omega$; (b) $A_v = -3.71$)



4.4 THE COMMON-DRAIN (SOURCE-FOLLOWER) AMPLIFIER

Objective: • Analyze the common-drain (source-follower) amplifier and become familiar with the general characteristics of this circuit.

The second type of MOSFET amplifier to be considered is the **common-drain circuit**. An example of this circuit configuration is shown in Figure 4.26. As seen in the figure, the output signal is taken off the source with respect to ground and the drain is connected directly to V_{DD} . Since V_{DD} becomes signal ground in the ac equivalent circuit, we have the name common drain. The more common name is **source follower**. The reason for this name will become apparent as we proceed through the analysis.

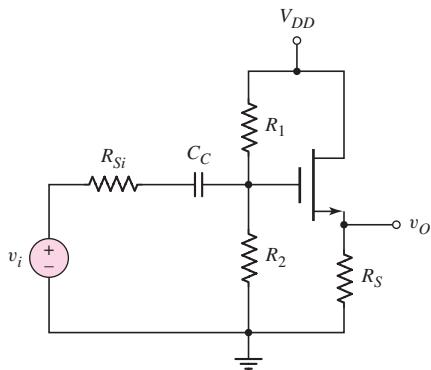


Figure 4.26 NMOS source-follower or common-drain amplifier

4.4.1 Small-Signal Voltage Gain

The dc analysis of the circuit is exactly the same as we have already seen, so we will concentrate on the small-signal analysis. The small-signal equivalent circuit, assuming the coupling capacitor acts as a short circuit, is shown in Figure 4.27(a). The drain is at signal ground, and the small-signal resistance r_o of the transistor is in

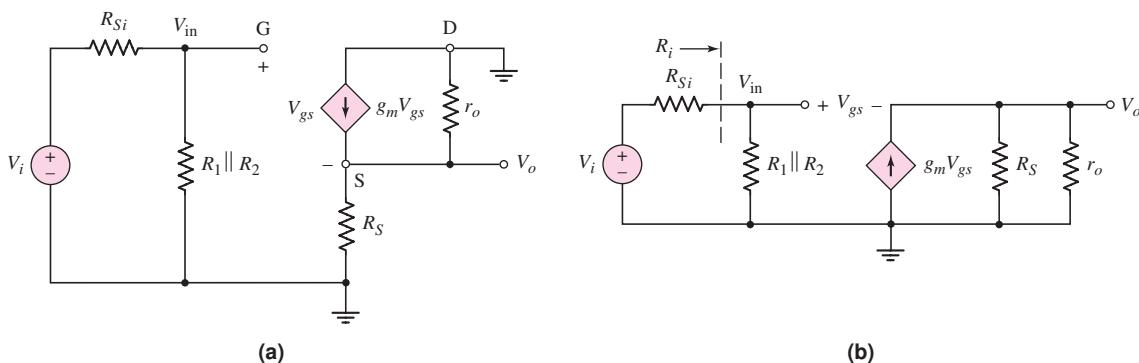


Figure 4.27 (a) Small-signal equivalent circuit of NMOS source follower and (b) small-signal equivalent circuit of NMOS source follower with all signal grounds at a common point

parallel with the dependent current source. Figure 4.27(b) is the same equivalent circuit, but with all signal grounds at a common point. We are again neglecting the body effect. The output voltage is

$$V_o = (g_m V_{gs})(R_S \| r_o) \quad (4.30)$$

Writing a KVL equation from input to output results in the following:

$$V_{in} = V_{gs} + V_o = V_{gs} + g_m V_{gs}(R_S \| r_o) \quad (4.31(a))$$

Therefore, the gate-to-source voltage is

$$V_{gs} = \frac{V_{in}}{1 + g_m(R_S \| r_o)} = \left[\frac{\frac{1}{g_m}}{\frac{1}{g_m} + (R_S \| r_o)} \right] \cdot V_{in} \quad (4.31(b))$$

Equation (4.31(b)) is written in the form of a voltage-divider equation, in which the gate-to-source of the NMOS device looks like a resistance with a value of $1/g_m$. More accurately, the effective resistance looking into the source terminal (ignoring r_o) is $1/g_m$. The voltage V_{in} is related to the source input voltage V_i by

$$V_{in} = \left(\frac{R_i}{R_i + R_{Si}} \right) \cdot V_i \quad (4.32)$$

where $R_i = R_1 \| R_2$ is the input resistance to the amplifier.

Substituting Equations (4.31(b)) and (4.32) into (4.30), we have the small-signal voltage gain:

$$A_v = \frac{V_o}{V_i} = \frac{g_m(R_S \| r_o)}{1 + g_m(R_S \| r_o)} \cdot \left(\frac{R_i}{R_i + R_{Si}} \right) \quad (4.33(a))$$

or

$$A_v = \frac{R_S \| r_o}{\frac{1}{g_m} + R_S \| r_o} \cdot \left(\frac{R_i}{R_i + R_{Si}} \right) \quad (4.33(b))$$

which again is written in the form of a voltage-divider equation. An inspection of Equation 4.33(b) shows that the magnitude of the voltage gain is always less than unity.

EXAMPLE 4.7

Objective: Calculate the small-signal voltage gain of the source-follower circuit in Figure 4.26.

Assume the circuit parameters are $V_{DD} = 12$ V, $R_1 = 162$ k Ω , $R_2 = 463$ k Ω , and $R_S = 0.75$ k Ω , and the transistor parameters are $V_{TN} = 1.5$ V, $K_n = 4$ mA/V 2 , and $\lambda = 0.01$ V $^{-1}$. Also assume $R_{Si} = 4$ k Ω .

Solution: The dc analysis results are $I_{DQ} = 7.97$ mA and $V_{GSQ} = 2.91$ V. The small-signal transconductance is therefore

$$g_m = 2K_n(V_{GSQ} - V_{TN}) = 2(4)(2.91 - 1.5) = 11.3 \text{ mA/V}$$

and the small-signal transistor resistance is

$$r_o \cong [\lambda I_{DQ}]^{-1} = [(0.01)(7.97)]^{-1} = 12.5 \text{ k}\Omega$$

The amplifier input resistance is

$$R_i = R_1 \| R_2 = 162 \| 463 = 120 \text{ k}\Omega$$

The small-signal voltage gain then becomes

$$A_v = \frac{g_m(R_S \| r_o)}{1 + g_m(R_S \| r_o)} \cdot \frac{R_i}{R_i + R_{Si}}$$

$$= \frac{(11.3)(0.75 \| 12.5)}{1 + (11.3)(0.75 \| 12.5)} \cdot \frac{120}{120 + 4} = +0.860$$

Comment: The magnitude of the small-signal voltage gain is less than 1. An examination of Equation (4.33(b)) shows that this is always true. Also, the voltage gain is positive, which means that the output signal voltage is in phase with the input signal voltage. Since the output signal is essentially equal to the input signal, the circuit is called a source follower.

EXERCISE PROBLEM

Ex 4.7: The source-follower circuit in Figure 4.26 has transistor parameters $V_{TN} = +0.8$ V, $K_n = 1$ mA/V², and $\lambda = 0.015$ V⁻¹. Let $V_{DD} = 10$ V, $R_{Si} = 200$ Ω, and $R_1 + R_2 = 400$ kΩ. Design the circuit such that $I_{DQ} = 1.5$ mA and $V_{DSQ} = 5$ V. Determine the small-signal voltage gain. (Ans. $R_S = 3.33$ kΩ, $R_1 = 119$ kΩ, $R_2 = 281$, kΩ, $A_v = 0.884$)

Although the voltage gain is slightly less than 1, the source follower is an extremely useful circuit because the output resistance is less than that of a common-source circuit, as we will show in the next section. A small output resistance is desirable when the circuit is to act as an ideal voltage source and drive a load circuit without suffering any loading effects.

DESIGN EXAMPLE 4.8

Objective: Design a source-follower amplifier with a p-channel enhancement-mode MOSFET to meet a set of specifications.

Specifications: The circuit to be designed has the configuration shown in Figure 4.28 with circuit parameters $V_{DD} = 20$ V and $R_{Si} = 4$ kΩ. The Q -point values are to be in the center of the load line with $I_{DQ} = 2.5$ mA. The input resistance is to

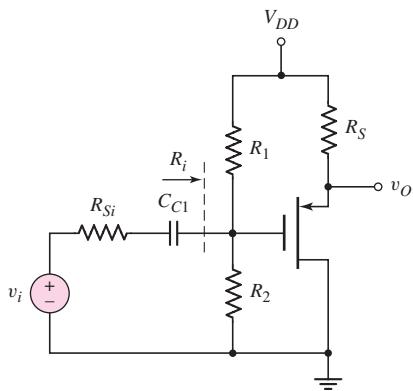


Figure 4.28 PMOS source follower

be $R_i = 200 \text{ k}\Omega$. The transistor W/L ratio is to be designed such that the small signal voltage gain is $A_v = 0.90$.

Choices: A transistor with nominal parameters $V_{TP} = -2 \text{ V}$, $k'_p = 40 \mu\text{A/V}^2$, and $\lambda = 0$ is available.

Solution (dc analysis): From a KVL equation around the source-to-drain loop, we have

$$V_{DD} = V_{SDQ} + I_{DQ}R_S$$

or

$$20 = 10 + (2.5)R_S$$

which yields the required source resistor to be $R_S = 4 \text{ k}\Omega$.

Solution (ac design): The small-signal voltage gain of this circuit is the same as that of a source follower with an NMOS device. From Equation (4.33(a)), we have

$$A_v = \frac{V_o}{V_i} = \frac{g_m R_S}{1 + g_m R_S} \cdot \frac{R_i}{R_i + R_{Si}}$$

which yields

$$0.90 = \frac{g_m(4)}{1 + g_m(4)} \cdot \frac{200}{200 + 4}$$

We find that the required transconductance must be $g_m = 2.80 \text{ mA/V}$. The transconductance can be written as

$$g_m = 2\sqrt{K_p I_{DQ}}$$

We have

$$2.80 \times 10^{-3} = 2\sqrt{K_p(2.5 \times 10^{-3})}$$

which yields

$$K_p = 0.784 \times 10^{-3} \text{ A/V}^2$$

The conduction parameter, as a function of width-to-length ratio, is

$$K_p = 0.784 \times 10^{-3} = \frac{k'_p}{2} \cdot \frac{W}{L} = \left(\frac{40 \times 10^{-6}}{2}\right) \cdot \left(\frac{W}{L}\right)$$

which means that the required width-to-length ratio must be

$$\frac{W}{L} = 39.2$$

Solution (dc design): Completing the dc analysis and design, we have

$$I_{DQ} = K_p(V_{SGQ} + V_{TP})^2$$

or

$$2.5 = 0.784(V_{SGQ} - 2)^2$$

which yields a quiescent source-to-gate voltage of $V_{SGQ} = 3.79 \text{ V}$. The quiescent source-to-gate voltage can also be written as

$$V_{SGQ} = (V_{DD} - I_{DQ}R_S) - \left(\frac{R_2}{R_1 + R_2}\right)(V_{DD})$$

Since

$$\left(\frac{R_2}{R_1 + R_2} \right) = \left(\frac{1}{R_1} \right) \left(\frac{R_1 R_2}{R_1 + R_2} \right) = \left(\frac{1}{R_1} \right) \cdot R_i$$

we have

$$3.79 = [20 - (2.5)(4)] - \left(\frac{1}{R_1} \right) (200)(20)$$

The bias resistor R_1 is then found to be

$$R_1 = 644 \text{ k}\Omega$$

Since $R_i = R_1 \| R_2 = 200 \text{ k}\Omega$, we find

$$R_2 = 290 \text{ k}\Omega$$

Comment: In order to achieve the desired specifications, a relatively large transconductance is required, which means that a relatively large transistor is needed. A large value of input resistance R_i has minimized the effect of loading due to the output resistance, R_{Si} , of the signal source.

EXERCISE PROBLEM

Ex 4.8: The circuit and transistor parameters for the source-follower amplifier shown in Figure 4.29 are $R_S = 2 \text{ k}\Omega$, $V_{TP} = -1.2 \text{ V}$, $k'_p = 40 \mu\text{A/V}^2$, and $\lambda = 0$. (a) Design the transistor width-to-length ratio such that $I_{DQ} = 1.5 \text{ mA}$. (b) Find the small-signal voltage gain. (c) Using the results of part (a), determine the value of R_L that will result in a 10 percent reduction in voltage gain. (Ans. (a) $W/L = 117$, (b) $A_v = 0.882$, (c) $R_L = 2.12 \text{ k}\Omega$)

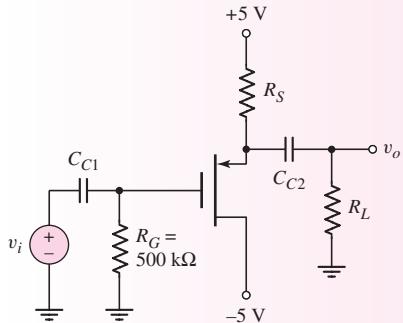


Figure 4.29 Figure for Exercise Ex 4.8

4.4.2 Input and Output Impedance

The small-signal input resistance R_i as defined in Figure 4.27(b), for example, is the Thevenin equivalent resistance of the bias resistors. Even though the input resistance to the gate of the MOSFET is essentially infinite, the input bias resistances do provide a loading effect. This same effect was seen in the common-source circuits.

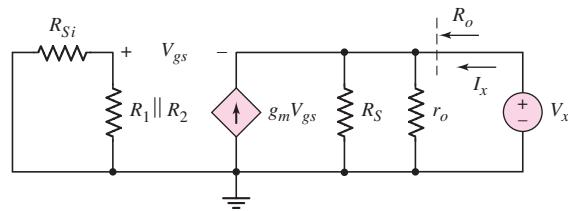


Figure 4.30 Equivalent circuit of NMOS source follower, for determining output resistance

To calculate the small-signal output resistance, we set all independent small-signal sources equal to zero, apply a test voltage to the output terminals, and measure a test current. Figure 4.30 shows the circuit we will use to determine the output resistance of the source follower shown in Figure 4.26. We set $V_i = 0$ and apply a test voltage V_x . Since there are no capacitances in the circuit, the output impedance is simply an output resistance, which is defined as

$$R_o = \frac{V_x}{I_x} \quad (4.34)$$

Writing a KCL equation at the output source terminal produces

$$I_x + g_m V_{gs} = \frac{V_x}{R_S} + \frac{V_x}{r_o} \quad (4.35)$$

Since there is no current in the input portion of the circuit, we see that $V_{gs} = -V_x$. Therefore, Equation (4.35) becomes

$$I_x = V_x \left(g_m + \frac{1}{R_S} + \frac{1}{r_o} \right) \quad (4.36(a))$$

or

$$\frac{I_x}{V_x} = \frac{1}{R_o} = g_m + \frac{1}{R_S} + \frac{1}{r_o} \quad (4.36(b))$$

The output resistance is then

$$R_o = \frac{1}{g_m} \| R_S \| r_o \quad (4.37)$$

From Figure 4.30, we see that the voltage V_{gs} is directly across the current source $g_m V_{gs}$. This means that the effective resistance of the device is $1/g_m$. The output resistance given by Equation (4.37) can therefore be written directly. This result also means that the resistance looking into the source terminal (ignoring r_o) is $1/g_m$, as previously noted.

EXAMPLE 4.9

Objective: Calculate the output resistance of a source-follower circuit.

Consider the circuit shown in Figure 4.26 with circuit and transistor parameters given in Example 4.7.

Solution: The results of Example 4.7 are: $R_S = 0.75 \text{ k}\Omega$, $r_o = 12.5 \text{ k}\Omega$, and $g_m = 11.3 \text{ mA/V}$. Using Figure 4.30 and Equation (4.37), we find

$$R_o = \frac{1}{g_m} \| R_S \| r_o = \frac{1}{11.3} \| 0.75 \| 12.5$$

or

$$R_o = 0.0787 \text{ k}\Omega = 78.7 \Omega$$

Comment: The output resistance of a source-follower circuit is dominated by the transconductance parameter. Also, because the output resistance is very low, the source follower tends to act like an ideal voltage source, which means that the output can drive another circuit without significant loading effects.

EXERCISE PROBLEM

Ex 4.9: Consider the circuit shown in Figure 4.28 with circuit parameters $V_{DD} = 5 \text{ V}$, $R_S = 5 \text{ k}\Omega$, $R_1 = 70.7 \text{ k}\Omega$, $R_2 = 9.3 \text{ k}\Omega$, and $R_{Si} = 500 \Omega$. The transistor parameters are: $V_{TP} = -0.8 \text{ V}$, $K_p = 0.4 \text{ mA/V}^2$, and $\lambda = 0$. Calculate the small-signal voltage gain $A_v = v_o/v_i$ and the output resistance R_o seen looking back into the circuit. (Ans. $A_v = 0.817$, $R_o = 0.915 \text{ k}\Omega$)

Test Your Understanding

TYU 4.8 For an NMOS source-follower circuit, the parameters are $g_m = 4 \text{ mA/V}$ and $r_o = 50 \text{ k}\Omega$. (a) Find the no load ($R_S = \infty$) small-signal voltage gain and the output resistance. (b) Determine the small-signal voltage gain when a $4 \text{ k}\Omega$ load is connected to the output. (Ans. (a) $A_v = 0.995$, $R_o \cong 0.25 \text{ k}\Omega$; (b) $A_v = 0.937$)

TYU 4.9 The transistor in the source-follower circuit shown in Figure 4.31 is biased with a constant current source. The transistor parameters are: $V_{TN} = 2 \text{ V}$, $k'_n = 40 \mu\text{A/V}^2$, and $\lambda = 0.01 \text{ V}^{-1}$. The load resistor is $R_L = 4 \text{ k}\Omega$. (a) Design the transistor width-to-length ratio such that $g_m = 2 \text{ mA/V}$ when $I = 0.8 \text{ mA}$. What is the corresponding value for V_{GS} ? (b) Determine the small-signal voltage gain and the output resistance R_o . (Ans. (a) $W/L = 62.5$, $V_{GS} = 2.8 \text{ V}$; (b) $A_v = 0.886$, $R_o \cong 0.5 \text{ k}\Omega$)

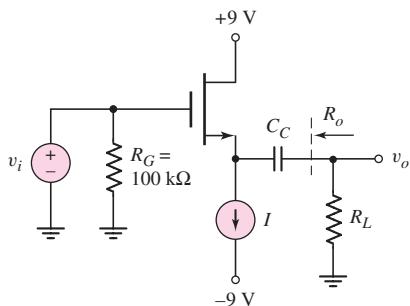


Figure 4.31 Figure for Exercise TYU 4.9

4.5 THE COMMON-GATE CONFIGURATION

Objective: • Analyze the common-gate amplifier and become familiar with the general characteristics of this circuit.

The third amplifier configuration is the **common-gate circuit**. To determine the small-signal voltage and current gains, and the input and output impedances, we will use the same small-signal equivalent circuit for the transistor that was used previously. The dc analysis of the common-gate circuit is the same as that of previous MOSFET circuits.

4.5.1 Small-Signal Voltage and Current Gains

In the common-gate configuration, the input signal is applied to the source terminal and the gate is at signal ground. The common-gate configuration shown in Figure 4.32 is biased with a **constant-current source** I_Q . The gate resistor R_G prevents the buildup of static charge on the gate terminal, and the capacitor C_G ensures that the gate is at signal ground. The coupling capacitor C_{C1} couples the signal to the source, and coupling capacitor C_{C2} couples the output voltage to load resistance R_L .

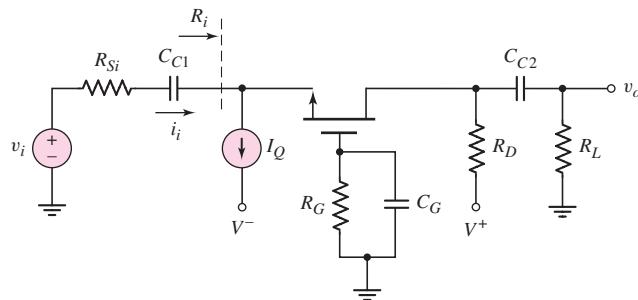


Figure 4.32 Common-gate circuit

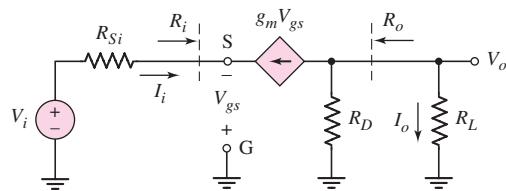


Figure 4.33 Small-signal equivalent circuit of common-gate amplifier

The small-signal equivalent circuit is shown in Figure 4.33. The small-signal transistor resistance r_o is assumed to be infinite. Since the source is the input terminal, the small-signal equivalent circuit shown in Figure 4.33 may appear to be different from those considered previously. However, to sketch the equivalent circuit, we can use the same technique as used previously. Sketch in the three terminals of the

transistor with the source at the input for this case. Then draw in the transistor equivalent circuit between the three terminals and then sketch in the remaining circuit elements around the transistor.

The output voltage is

$$V_o = -(g_m V_{gs})(R_D \parallel R_L) \quad (4.38)$$

Writing the KVL equation around the input, we find

$$V_i = I_i R_{Si} - V_{gs} \quad (4.39)$$

where $I_i = -g_m V_{gs}$. The gate-to-source voltage can then be written as

$$V_{gs} = \frac{-V_i}{1 + g_m R_{Si}} \quad (4.40)$$

The small-signal voltage gain is found to be

$$A_v = \frac{V_o}{V_i} = \frac{g_m (R_D \parallel R_L)}{1 + g_m R_{Si}} \quad (4.41)$$

Also, since the voltage gain is positive, the output and input signals are in phase.

In many cases, the signal input to a common-gate circuit is a current. Figure 4.34 shows the small-signal equivalent common-gate circuit with a Norton equivalent circuit as the signal source. We can calculate a current gain. The output current I_o can be written

$$I_o = \left(\frac{R_D}{R_D + R_L} \right) (-g_m V_{gs}) \quad (4.42)$$

At the input we have

$$I_i + g_m V_{gs} + \frac{V_{gs}}{R_{Si}} = 0 \quad (4.43)$$

or

$$V_{gs} = -I_i \left(\frac{R_{Si}}{1 + g_m R_{Si}} \right) \quad (4.44)$$

The small-signal current gain is then

$$A_i = \frac{I_o}{I_i} = \left(\frac{R_D}{R_D + R_L} \right) \cdot \left(\frac{g_m R_{Si}}{1 + g_m R_{Si}} \right) \quad (4.45)$$

We may note that if $R_D \gg R_L$ and $g_m R_{Si} \gg 1$, then the current gain is essentially unity.

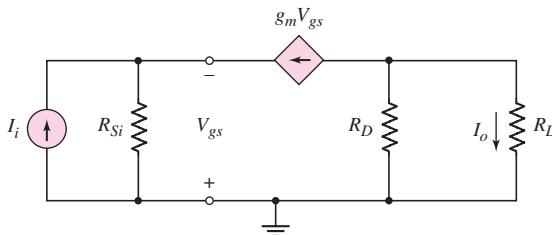


Figure 4.34 Small-signal equivalent circuit of common-gate amplifier with a Norton equivalent signal source

4.5.2 Input and Output Impedance

In contrast to the common-source and source-follower amplifiers, the common-gate circuit has a low input resistance because of the transistor. However, if the input signal is a current, a low input resistance is an advantage. The input resistance is defined, using Figure 4.33, as

$$R_i = \frac{-V_{gs}}{I_i} \quad (4.46)$$

Since $I_i = -g_m V_{gs}$, the input resistance is

$$R_i = \frac{1}{g_m} \quad (4.47)$$

This result has been obtained previously.

We can find the output resistance by setting the input signal voltage equal to zero. From Figure 4.33, we see that $V_{gs} = -g_m V_{gs} R_{Si}$, which means that $V_{gs} = 0$. Consequently, $g_m V_{gs} = 0$. The output resistance, looking back from the load resistance, is therefore

$$R_o = R_D \quad (4.48)$$

EXAMPLE 4.10

Objective: For the common-gate circuit, determine the output voltage for a given input current.

For the circuits shown in Figures 4.32 and 4.34, the circuit parameters are: $I_Q = 1 \text{ mA}$, $V^+ = 5 \text{ V}$, $V^- = -5 \text{ V}$, $R_G = 100 \text{ k}\Omega$, $R_D = 4 \text{ k}\Omega$, and $R_L = 10 \text{ k}\Omega$. The transistor parameters are: $V_{TN} = 1 \text{ V}$, $K_n = 1 \text{ mA/V}^2$, and $\lambda = 0$. Assume the input current in Figure 4.34 is $100 \sin \omega t \mu\text{A}$ and assume $R_{Si} = 50 \text{ k}\Omega$.

Solution: The quiescent gate-to-source voltage is determined from

$$I_Q = I_{DQ} = K_n(V_{GSQ} - V_{TN})^2$$

or

$$1 = 1(V_{GSQ} - 1)^2$$

which yields

$$V_{GSQ} = 2 \text{ V}$$

The small-signal transconductance is

$$g_m = 2K_n(V_{GSQ} - V_{TN}) = 2(1)(2 - 1) = 2 \text{ mA/V}$$

From Equation (4.45), we can write the output current as

$$I_o = I_i \left(\frac{R_D}{R_D + R_L} \right) \cdot \left(\frac{g_m R_{Si}}{1 + g_m R_{Si}} \right)$$

The output voltage is $V_o = I_o R_L$, so we find

$$V_o = I_i \left(\frac{R_L R_D}{R_D + R_L} \right) \cdot \left(\frac{g_m R_{Si}}{1 + g_m R_{Si}} \right)$$

$$= \left[\frac{(10)(4)}{4+10} \right] \cdot \left[\frac{(2)(50)}{1+(2)(50)} \right] \cdot (0.1) \sin \omega t$$

or

$$V_o = 0.283 \sin \omega t \text{ V}$$

Comment: The MOSFET common-gate amplifier is useful if the input signal is a current.

EXERCISE PROBLEM

Ex 4.10: Consider the circuit shown in Figure 4.35 with circuit parameters $V^+ = 5 \text{ V}$, $V^- = -5 \text{ V}$, $R_S = 4 \text{ k}\Omega$, $R_D = 2 \text{ k}\Omega$, $R_L = 4 \text{ k}\Omega$ and $R_G = 50 \text{ k}\Omega$. The transistor parameters are: $K_p = 1 \text{ mA/V}^2$, $V_{TP} = -0.8 \text{ V}$, and $\lambda = 0$. Draw the small-signal equivalent circuit, determine the small-signal voltage gain $A_v = V_o / V_i$, and find the input resistance R_i . (Ans. $A_v = 2.41$, $R_i = 0.485 \text{ k}\Omega$)

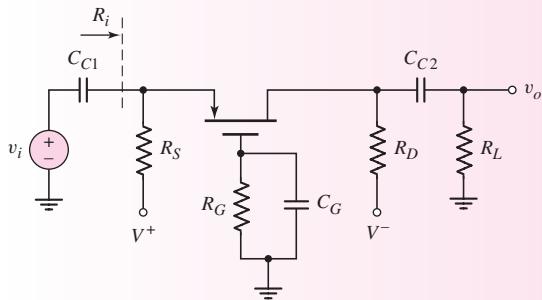


Figure 4.35 Figure for Exercise Ex 4.10

Test Your Understanding

TYU 4.10 For the circuit shown in Figure 4.32, the circuit parameters are: $V^+ = 5 \text{ V}$, $V^- = -5 \text{ V}$, $R_G = 100 \text{ k}\Omega$, $R_L = 4 \text{ k}\Omega$, and $I_Q = 0.5 \text{ mA}$. The transistor parameters are $V_{TN} = 1 \text{ V}$ and $\lambda = 0$. The circuit is driven by a signal current source I_i . Redesign R_D and g_m such that the transfer function V_o/I_i is $2.4 \text{ k}\Omega$ and the input resistance is $R_i = 350 \Omega$. Determine V_{GSQ} and show that the transistor is biased in the saturation region. (Ans. $g_m = 2.86 \text{ mA/V}$, $R_D = 6 \text{ k}\Omega$, $V_{GSQ} = 1.35 \text{ V}$)



4.6

THE THREE BASIC AMPLIFIER CONFIGURATIONS: SUMMARY AND COMPARISON

Objective: • Compare the general characteristics of the three basic amplifier configurations.

Table 4.2 Characteristics of the three MOSFET amplifier configurations

Configuration	Voltage gain	Current gain	Input resistance	Output resistance
Common source	$A_v > 1$	—	R_{TH}	Moderate to high
Source follower	$A_v \approx 1$	—	R_{TH}	Low
Common gate	$A_v > 1$	$A_i \approx 1$	Low	Moderate to high

Table 4.2 is a summary of the small-signal characteristics of the three amplifier configurations.

The common-source amplifier voltage gain magnitude is generally greater than 1. The voltage gain of the source follower is slightly less than 1, and that of the common-gate circuit is generally greater than 1.

The input resistance looking directly into the gate of the common-source and source-follower circuits is essentially infinite at low to moderate signal frequencies. However, the input resistance of these discrete amplifiers is the Thevenin equivalent resistance R_{TH} of the bias resistors. In contrast, the input resistance to the common-gate circuit is generally in the range of only a few hundred ohms.

The output resistance of the source follower is generally in the range of a few hundred ohms or less. The output resistance of the common-source and common-gate configurations is dominated by the resistance R_D . In Chapters 10 and 11, we will see that the output resistance of these configurations is dominated by the resistance r_o when transistors are used as load devices in ICs.

The specific characteristics of these single-stage amplifiers are used in the design of multistage amplifiers.

4.7

SINGLE-STAGE INTEGRATED CIRCUIT MOSFET AMPLIFIERS

Objective: • Analyze all-MOS transistor circuits that become the foundation of integrated circuits.

In the last chapter, we considered three all-MOSFET inverters and plotted the voltage transfer characteristics. All three inverters use an n-channel enhancement-mode driver transistor. The three types of load devices are an n-channel enhancement-mode device, an n-channel depletion-mode device, and a p-channel enhancement-mode device. The MOS transistor used as a load device is referred to as an **active load**. We mentioned that these three circuits can be used as amplifiers.

In this section, we revisit these three circuits and consider their amplifier characteristics. We will emphasize the small-signal equivalent circuits. This section serves as an introduction to more advanced MOS integrated circuit amplifier designs considered in Part 2 of the text.

4.7.1 Load Line Revisited

In dealing with all-transistor circuits, it will be instructive to consider the equivalent load lines that we have considered previously in circuits with resistive loads. Before we deal with the nonlinear load lines or load curves, it may be worthwhile to revisit the load line concept of a single transistor with a resistive load.

Figure 4.36 shows a single MOSFET with a resistive load. The current–voltage characteristic of the resistive load device is given by Ohm's law, or $V_R = I_D R_D$. This curve is plotted in the top portion of Figure 4.37. The load line is given by the KVL equation around the drain-source loop, or $V_{DS} = V_{DD} - I_D R_D$, and is superimposed on the transistor characteristics in the lower portion of Figure 4.37. We may note that the last term in the load line equation, $I_D R_D$, is the voltage across the load device.

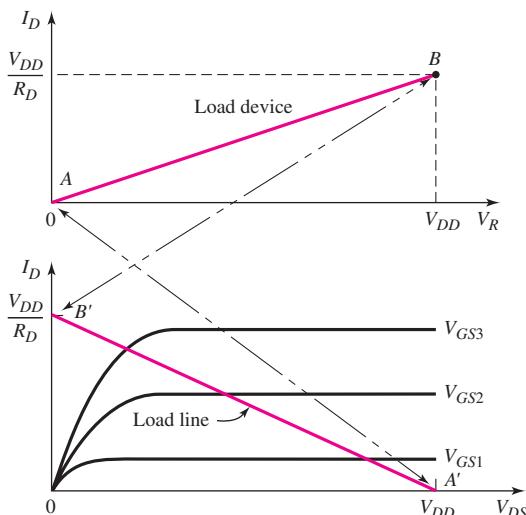


Figure 4.37 The I – V curve for the resistor load device (top) and the load line superimposed on the transistor characteristics (bottom)

We may compare two points on the load device characteristic to the load line. When $I_D = 0$, $V_R = 0$ on the load characteristic curve denoted by point A . On the load line, the $I_D = 0$ point corresponds to $V_{DS} = V_{DD}$, denoted by the point A' . The maximum current on the load characteristic curve occurs when $V_R = V_{DD}$ and is denoted by point B . On the load line, the maximum current point corresponds to $V_{DS} = 0$, denoted by point B' . The load line can be created by taking the mirror image of the load characteristic curve and superimposing this curve on the plot of transistor characteristics. We will see this same effect in the following sections.

4.7.2 NMOS Amplifiers with Enhancement Load

The characteristics of an n-channel enhancement load device were presented in the last chapter. Figure 4.38(a) shows an NMOS enhancement load transistor, and Figure 4.38(b) shows the current–voltage characteristics. The threshold voltage is V_{TNL} .

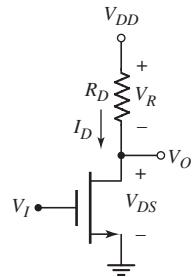


Figure 4.36 Single MOSFET circuit with resistive load

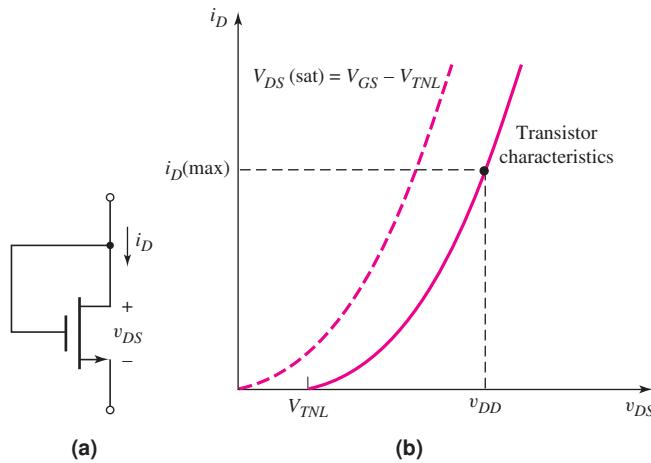


Figure 4.38 (a) NMOS enhancement-mode transistor with gate and drain connected in a load device configuration and (b) current–voltage characteristics of NMOS enhancement load transistor

Figure 4.39(a) shows an NMOS amplifier with enhancement load. The driver transistor is \$M_D\$ and the load transistor is \$M_L\$. The characteristics of transistor \$M_D\$ and the load curve are shown in Figure 4.39(b). The load curve is essentially the mirror image of the \$i\$–\$v\$ characteristic of the load device, as we discussed in the last section. Since the \$i\$–\$v\$ characteristics of the load device are nonlinear, the load curve is also nonlinear. The load curve intersects the voltage axis at \$V_{DD} - V_{TNL}\$, which is the point where the current in the enhancement load device goes to zero. The transition point is also shown on the curve.

The voltage transfer characteristic is also useful in visualizing the operation of the amplifier. This curve is shown in Figure 4.39(c). When the enhancement-mode driver first begins to conduct, it is biased in the saturation region. For use as an amplifier, the circuit \$Q\$-point should be in this region, as shown in both Figures 4.39(b) and (c).

We can now apply the small-signal equivalent circuits to find the voltage gain. In the discussion of the source follower, we found that the equivalent resistance looking into the source terminal (with \$R_S = \infty\$) was \$R_o = (1/g_m) \parallel r_o\$. The small-signal equivalent circuit of the inverter is given in Figure 4.40, where the subscripts \$D\$ and \$L\$ refer to the driver and load transistors, respectively. We are again neglecting the body effect of the load transistor.

The small-signal voltage gain is then

$$A_v = \frac{V_o}{V_i} = -g_{mD} \left(r_{oD} \parallel \frac{1}{g_{mL}} \parallel r_{oL} \right) \quad (4.49)$$

Since, generally, \$1/g_{mL} \ll r_{oL}\$ and \$1/g_{mD} \ll r_{oD}\$, the voltage gain, to a good approximation is given by

$$A_v = \frac{-g_{mD}}{g_{mL}} = -\sqrt{\frac{K_{nD}}{K_{nL}}} = -\sqrt{\frac{(W/L)_D}{(W/L)_L}} \quad (4.50)$$

The voltage gain, then, is related to the size of the two transistors.

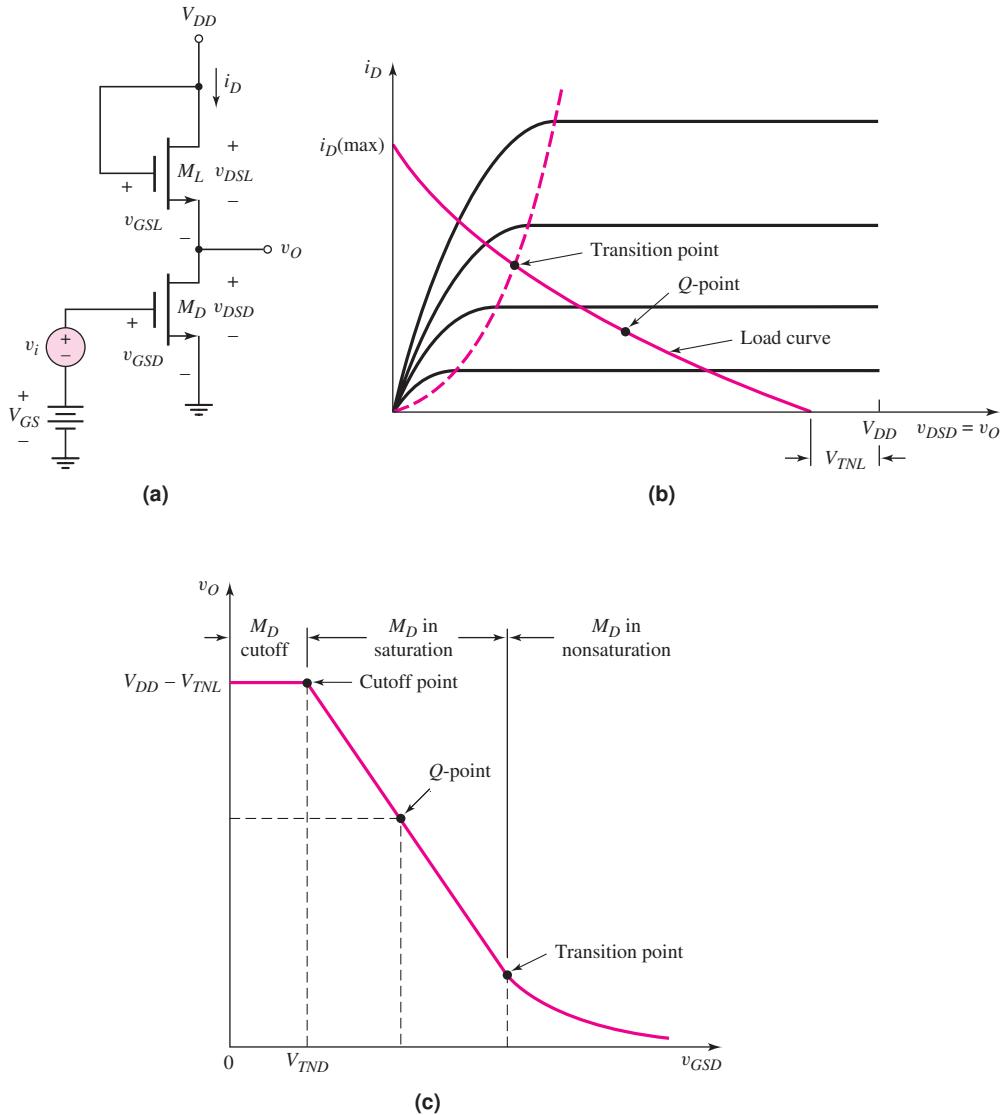


Figure 4.39 (a) NMOS amplifier with enhancement load device; (b) driver transistor characteristics and enhancement load curve with transition point; and (c) voltage transfer characteristics of NMOS amplifier with enhancement load device

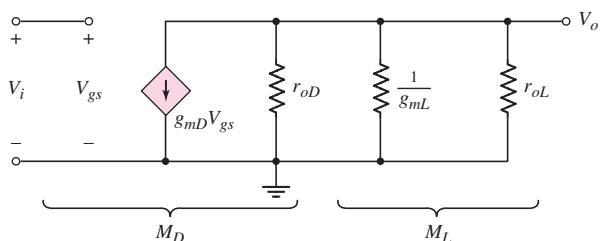


Figure 4.40 Small-signal equivalent circuit of NMOS inverter with enhancement load device

DESIGN EXAMPLE 4.11

Objective: Design an NMOS amplifier with an enhancement load to meet a set of specifications.

Specifications: An NMOS amplifier with the configuration shown in Figure 4.39(a) is to be designed to provide a small-signal voltage gain of $|A_v| = 10$. The Q-point is to be in the center of the saturation region. The circuit is to be biased at $V_{DD} = 5$ V.

Choices: NMOS transistors with parameters $V_{TN} = 1$ V, $k'_n = 60 \mu\text{A/V}^2$, and $\lambda = 0$ are available. The minimum width-to-length ratio is $(W/L)_{\min} = 1$. Tolerances of ± 5 percent in the k'_n and V_{TN} parameters must be considered.

Solution (ac design): From Equation (4.50), we have

$$|A_v| = 10 = \sqrt{\frac{(W/L)_D}{(W/L)_L}}$$

which can be written as

$$\left(\frac{W}{L}\right)_D = 100 \left(\frac{W}{L}\right)_L$$

If we set $(W/L)_L = 1$, then $(W/L)_D = 100$.

Solution (dc design): Setting the currents in the two transistors equal to each other (both transistors biased in saturation region), we have

$$i_{DD} = K_{nD}(v_{GSD} - V_{TND})^2 = i_{DL} = K_{nL}(v_{GSL} - V_{TNL})^2$$

From Figure 4.39(a), we see that $v_{GSL} = V_{DD} - v_O$. Substituting, we have

$$K_{nD}(v_{GSD} - V_{TND})^2 = K_{nL}(V_{DD} - v_O - V_{TNL})^2$$

Solving for v_O , we have

$$v_O = (V_{DD} - V_{TNL}) - \sqrt{\frac{K_{nD}}{K_{nL}}}(v_{GSD} - V_{TND})$$

At the transition point,

$$v_{O_t} = v_{DS(t)}(\text{sat}) = v_{GSDt} - V_{TND}$$

where v_{GSDt} is the gate-to-source voltage of the driver at the transition point. Then

$$v_{GSDt} - V_{TND} = (V_{DD} - V_{TNL}) - \sqrt{\frac{K_{nD}}{K_{nL}}}(v_{GSDt} - V_{TND})$$

Solving for v_{GSDt} , we obtain

$$v_{GSDt} = \frac{(V_{DD} - V_{TNL}) + V_{TND} \left(1 + \sqrt{\frac{K_{nD}}{K_{nL}}} \right)}{1 + \sqrt{\frac{K_{nD}}{K_{nL}}}}$$

Noting that

$$\sqrt{\frac{K_{nD}}{K_{nL}}} = \sqrt{\frac{(W/L)_D}{(W/L)_L}} = 10$$

we find

$$v_{GSDt} = \frac{(5 - 1) + (1)(1 + 10)}{1 + 10} = 1.36 \text{ V}$$

and

$$v_{Ot} = v_{DSDt} = v_{GSDt} - V_{TND} = 1.36 - 1 = 0.36 \text{ V}$$

Considering the transfer characteristics shown in Figure 4.41, we see that the center of the saturation region is halfway between the cutoff point ($v_{GSD} = V_{TND} = 1 \text{ V}$) and the transition point ($v_{GSDt} = 1.36 \text{ V}$), or

$$V_{GSQ} = \frac{1.36 - 1.0}{2} + 1.0 = 1.18 \text{ V}$$

Also

$$V_{DSDQ} = \frac{4 - 0.36}{2} + 0.36 = 2.18 \text{ V}$$

Trade-offs: Considering the tolerances in the k'_n parameter, we find the range in the small-signal voltage gain to be

$$|A_v|_{\max} = \sqrt{\frac{k'_{nD}}{k'_{nL}} \cdot \frac{(W/L)_D}{(W/L)_L}} = \sqrt{\frac{1.05}{0.95} \cdot (100)} = 10.5$$

and

$$|A_v|_{\min} = \sqrt{\frac{k'_{nD}}{k'_{nL}} \cdot \frac{(W/L)_D}{(W/L)_L}} = \sqrt{\frac{0.95}{1.05} \cdot (100)} = 9.51$$

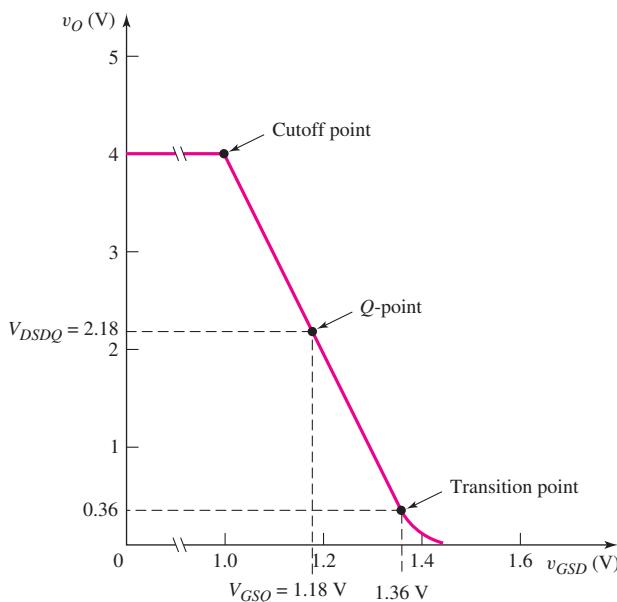


Figure 4.41 Voltage transfer characteristics and Q -point of NMOS amplifier with enhancement load, for Example 4.11

The tolerances in the k'_n and V_{TN} parameters will also affect the Q -point. This analysis is left as an end-of-chapter problem.

Comment: These results show that a very large difference is required in the sizes of the two transistors to produce a gain of 10. In fact, a gain of 10 is about the largest practical gain that can be produced by an enhancement load device. A larger small-signal gain can be obtained by using a depletion-mode MOSFET as a load device, as shown in the next section.

Design Pointer: The body effect of the load transistor was neglected in this analysis. The body effect will actually lower the small-signal voltage gain from that determined in the example.

EXERCISE PROBLEM

Ex 4.11: The bias voltage for the enhancement-load amplifier shown in Figure 4.39(a) is $V_{DD} = 3.3$ V. The transistor parameters are $V_{TND} = V_{TNL} = 0.4$ V, $k'_n = 100 \mu\text{A}/\text{V}^2$, $(W/L)_L = 1.2$, and $\lambda = 0$. (a) Design the circuit such that the small-signal voltage gain is $|A_v| = 8$. (b) Determine V_{GSDQ} such that the Q -point is in the center of the saturation region. (Ans. (a) $(W/L)_D = 76.8$, (b) $V_{GSDQ} = 0.561$ V).

4.7.3 NMOS Amplifier with Depletion Load

Figure 4.42(a) shows the NMOS depletion-mode transistor connected as a load device and Figure 4.42(b) shows the current–voltage characteristics. The transition point is also indicated. The threshold voltage V_{TNL} of this device is negative, which means that the v_{DS} value at the transition point is positive. Also, the slope of the curve in the saturation region is not zero; therefore, a finite resistance r_o exists in this region.

Figure 4.43(a) shows an **NMOS depletion load amplifier**. The transistor characteristics of M_D and the load curve for the circuit are shown in Figure 4.43(b). The load curve, again, is the mirror image of the i – v characteristic of the load device. Since the i – v characteristics of the load device are nonlinear, the load curve is also nonlinear. The transition points for both M_D and M_L are also indicated. Point A is the transition point for M_D , and point B is the transition point for M_L . The Q -point should be approximately midway between the two transition points.

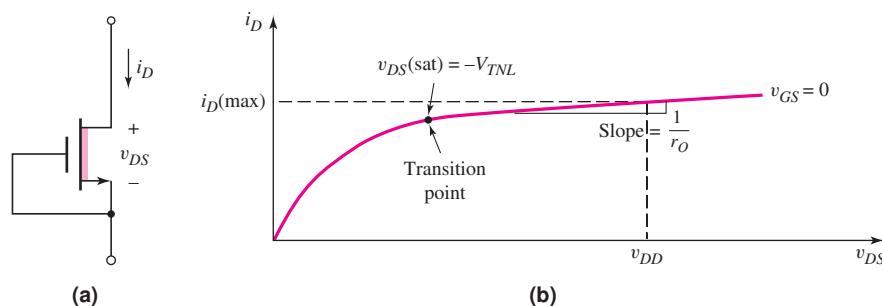


Figure 4.42 (a) NMOS depletion-mode transistor with gate and source connected in a load device configuration and (b) current–voltage characteristic of NMOS depletion load transistor

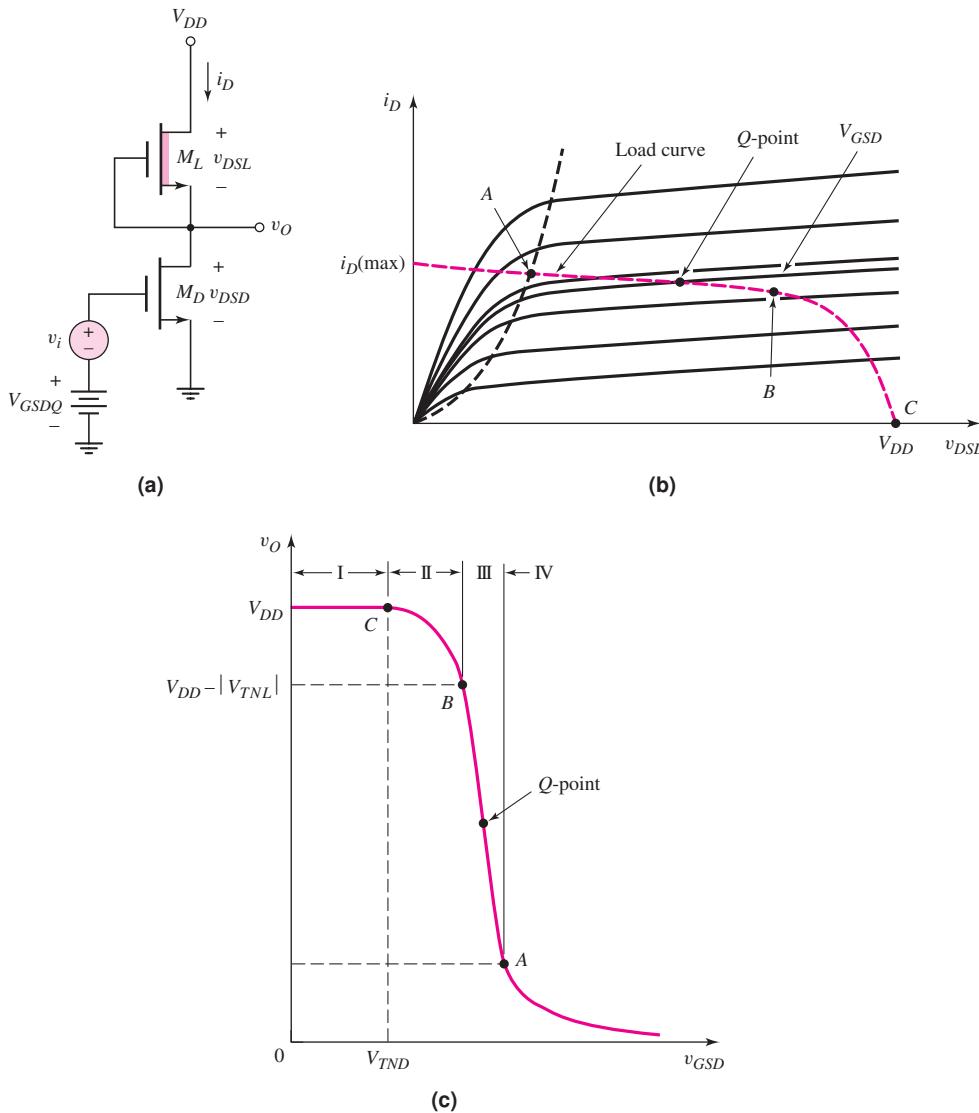


Figure 4.43 (a) NMOS amplifier with depletion load device; (b) driver transistor characteristics and depletion load curve, with transition points between saturation and nonsaturation regions; (c) voltage transfer characteristics

The dc voltage V_{GSDQ} biases transistor M_D in the saturation region at the Q -point. The signal voltage v_i superimposes a time-varying gate-to-source voltage on the dc value, and the bias point moves along the load curve about the Q -point. Again, both M_D and M_L must be biased in their saturation regions at all times.

The voltage transfer characteristic of this circuit is shown in Figure 4.43(c). Region III corresponds to the condition in which both transistors are biased in the saturation region. The desired Q -point is indicated.

We can again apply the small-signal equivalent circuit to find the small-signal voltage gain. Since the gate-to-source voltage of the depletion-load device is held at zero, the equivalent resistance looking into the source terminal is $R_o = r_o$.

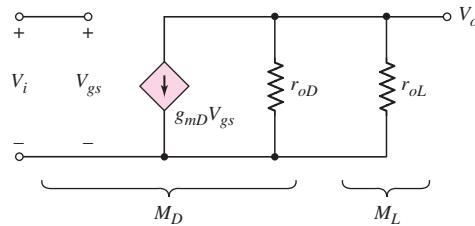


Figure 4.44 Small-signal equivalent circuit of NMOS inverter with depletion load device

The small-signal equivalent circuit of the inverter is given in Figure 4.44, where the subscripts *D* and *L* refer to the driver and load transistors, respectively. We are again neglecting the body effect of the load device.

The small-signal voltage gain is then

$$A_v = \frac{V_o}{V_i} = -g_{mD}(r_{oD} \parallel r_{oL}) \quad (4.51)$$

In this circuit, the voltage gain is directly proportional to the output resistances of the two transistors.

EXAMPLE 4.12

Objective: Determine the small-signal voltage gain of the NMOS amplifier with depletion load.

For the circuit shown in Figure 4.43(a), assume transistor parameters of $V_{TND} = +0.8$ V, $V_{TNL} = -1.5$ V, $K_{nD} = 1$ mA/V², $K_{nL} = 0.2$ mA/V², and $\lambda_D = \lambda_L = 0.01$ V⁻¹. Assume the transistors are biased at $I_{DQ} = 0.2$ mA.

Solution: The transconductance of the driver is

$$g_{mD} = 2\sqrt{K_{nD}I_{DQ}} = 2\sqrt{(1)(0.2)} = 0.894 \text{ mA/V}$$

Since $\lambda_D = \lambda_L$, the output resistances are

$$r_{oD} = r_{oL} = \frac{1}{\lambda I_{DQ}} = \frac{1}{(0.01)(0.2)} = 500 \text{ k}\Omega$$

The small-signal voltage gain is then

$$A_v = -g_{mD}(r_{oD} \parallel r_{oL}) = -(0.894)(500 \parallel 500) = -224$$

Comment: The voltage gain of the NMOS amplifier with depletion load is, in general, significantly larger than that with the enhancement load device. The body effect will lower the ideal gain factor.

Discussion: One aspect of this circuit design that we have not emphasized is the dc biasing. We mentioned that both transistors need to be biased in their saturation regions. From Figure 4.43(a), this dc biasing is accomplished with the dc source V_{GSDQ} . However, because of the steep slope of the transfer characteristics (Figure 4.43(c)), applying the “correct” voltage becomes difficult. As we will see in the next section, dc biasing is generally accomplished with current source biasing.

EXERCISE PROBLEM

Ex 4.12: Assume the depletion-load amplifier in Figure 4.43(a) is biased at $I_{DQ} = 0.1 \text{ mA}$. The transistor parameters are $K_{nD} = 250 \mu\text{A/V}^2$, $K_{nL} = 25 \mu\text{A/V}^2$, $V_{TND} = 0.4 \text{ V}$, $V_{TNL} = -0.8 \text{ V}$, and $\lambda_1 = \lambda_2 = 0.02 \text{ V}^{-1}$. Determine the small-signal voltage gain. (Ans. $A_v = -79.1$)

4.7.4 NMOS Amplifier with Active Loads

CMOS Common-Source Amplifier

An amplifier using an n-channel enhancement-mode driver and a p-channel enhancement mode active load is shown in Figure 4.45(a) in a common-source configuration. The p-channel active load transistor M_2 is biased from M_3 and I_{Bias} . This configuration is similar to the MOSFET current source shown in Figure 3.49 in Chapter 3. With both n- and p-channel transistors in the same circuit, this circuit is now referred to as a CMOS amplifier. The CMOS configuration is used almost exclusively rather than the NMOS enhancement load or depletion load devices.

The $i-v$ characteristic curve for M_2 is shown in Figure 4.45(b). The source-to-gate voltage is a constant and is established by M_3 . The driver transistor characteristics and the load curve are shown in Figure 4.45(c). The transition points of both M_1

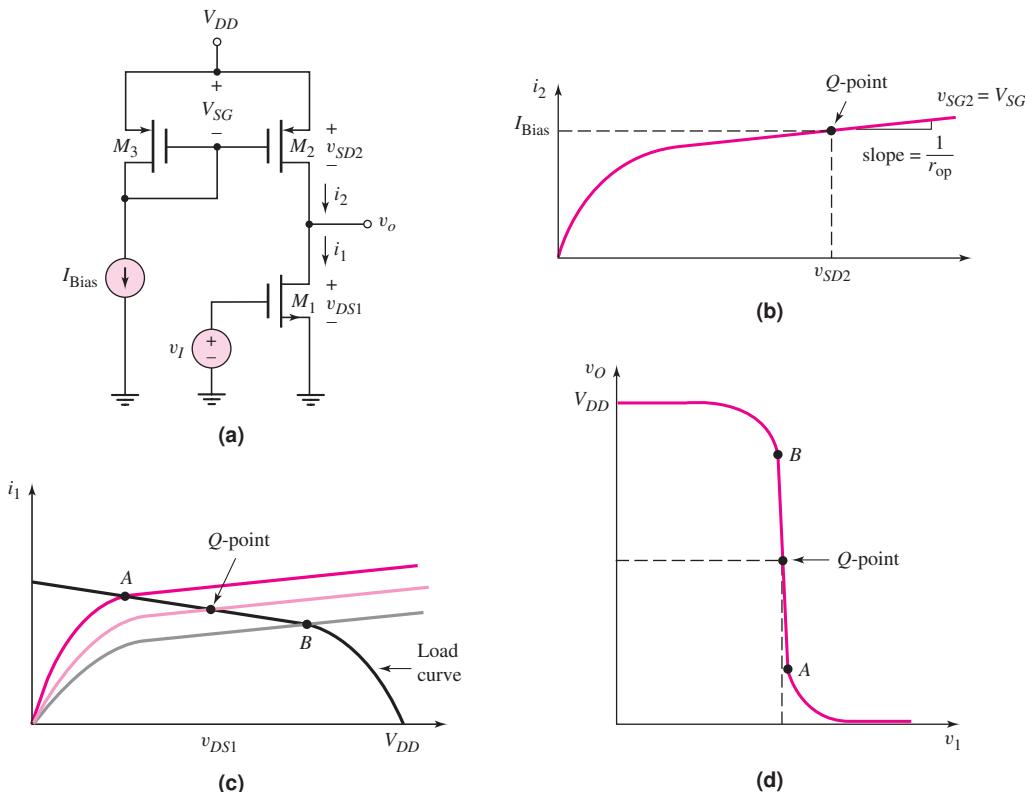


Figure 4.45 (a) CMOS common-source amplifier; (b) PMOS active load $i-v$ characteristic, (c) driver transistor characteristics with load curve, (d) voltage transfer characteristics

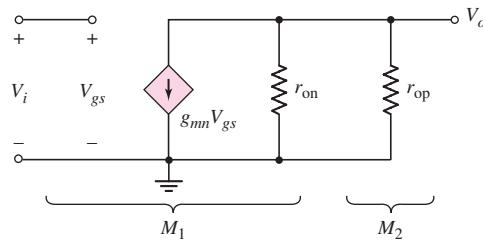


Figure 4.46 Small-signal equivalent circuit of the CMOS common-source amplifier

and M_2 are shown. Point A is the transition point for M_1 and point B is the transition point for M_2 . The Q -point, to establish an amplifier, should be approximately halfway between points A and B , so that both transistors are biased in their saturation regions. The voltage transfer characteristics are shown in Figure 4.45(d). Shown on the curve are the same transition points A and B and the desired Q -point.

We again apply the small-signal equivalent circuits to find the small-signal voltage gain. With v_{SG2} held constant, the equivalent resistance looking into the drain of M_2 is just $R_o = r_{op}$. The small-signal equivalent circuit of the inverter is then as given in Figure 4.46. The subscripts n and p refer to the n-channel and p-channel transistors, respectively. We may note that the body terminal of M_1 will be tied to ground, which is the same as the source of M_1 , and the body terminal of M_2 will be tied to V_{DD} , which is the same as the source of M_2 . Hence, there is no body effect in this circuit.

The small-signal voltage gain is

$$A_v = \frac{V_o}{V_i} = -g_{mn}(r_{on} \parallel r_{op}) \quad (4.52)$$

Again for this circuit, the small-signal voltage gain is directly proportional to the output resistances of the two transistors.

EXAMPLE 4.13

Objective: Determine the small-signal voltage gain of the CMOS amplifier.

For the circuit shown in Figure 4.45(a), assume transistor parameters of $V_{TN} = +0.8$ V, $V_{TP} = -0.8$ V, $k'_n = 80 \mu\text{A/V}^2$, $k'_p = 40 \mu\text{A/V}^2$, $(W/L)_n = 15$, $(W/L)_p = 30$, and $\lambda_n = \lambda_p = 0.01 \text{ V}^{-1}$. Also, assume $I_{\text{Bias}} = 0.2$ mA.

Solution: The transconductance of the NMOS driver is

$$\begin{aligned} g_{mn} &= 2\sqrt{K_n I_{DQ}} = 2\sqrt{\left(\frac{k'_n}{2}\right) \left(\frac{W}{L}\right)_n I_{\text{Bias}}} \\ &= 2\sqrt{\left(\frac{0.08}{2}\right) (15)(0.2)} = 0.693 \text{ mA/V} \end{aligned}$$

Since $\lambda_n = \lambda_p$, the output resistances are

$$r_{on} = r_{op} = \frac{1}{\lambda I_{DQ}} = \frac{1}{(0.01)(0.2)} = 500 \text{ k}\Omega$$

The small-signal voltage gain is then

$$A_v = -g_m(r_{on}\|r_{op}) = -(0.693)(500\|500) = -173$$

Comment: The voltage gain of the CMOS amplifier is on the same order of magnitude as the NMOS amplifier with depletion load. However, the CMOS amplifier does not suffer from the body effect.

Discussion: In the circuit configuration shown in Figure 4.45(a), we must again apply a dc voltage to the gate of M_1 to achieve the “proper” Q -point. We will show in later chapters using more sophisticated circuits how the Q -point is more easily established with current-source biasing. However, this circuit demonstrates the basic principles of the CMOS common-source amplifier.

EXERCISE PROBLEM

Ex 4.13: For the circuit shown in Figure 4.45(a), assume transistor parameters of $V_{TN} = +0.5$ V, $V_{TP} = -0.5$ V, $k'_n = 80 \mu\text{A}/\text{V}^2$, $k'_p = 40 \mu\text{A}/\text{V}^2$, and $\lambda_n = \lambda_p = 0.015 \text{ V}^{-1}$. Assume $I_{Bias} = 0.1$ mA. Assume M_2 and M_3 are matched. Find the width-to-length ratio of M_1 such that the small-signal voltage gain is $A_v = -250$. (Ans. $(W/L)_1 = 35.2$)

CMOS Source-Follower Amplifier

The same basic CMOS circuit configuration can be used to form a CMOS source-follower amplifier. Figure 4.47(a) shows a source-follower circuit. We see that for this source-follower circuit, the active load, which is M_2 , is an n-channel rather than a p-channel device. The input signal is applied to the gate of M_1 and the output is at the source of M_1 .

The small-signal equivalent circuit of this source-follower is shown in Figure 4.47(b). This circuit, with two signal grounds, is redrawn as shown in Figure 4.47(c) to combine the signal grounds.

EXAMPLE 4.14

Objective: Determine the small-signal voltage gain and output resistance of the source-follower amplifier shown in Figure 4.47(a).

Assume the reference bias current is $I_{Bias} = 0.20$ mA and the bias voltage is $V_{DD} = 3.3$ V. Assume that all transistors are matched (identical) with parameters $V_{TN} = 0.4$ V, $K_n = 0.20 \text{ mA/V}^2$, and $\lambda = 0.01 \text{ V}^{-1}$.

We may note that since M_3 and M_2 are matched transistors and have the same gate-to-source voltages, the drain current in M_1 is $I_{D1} = I_{Bias} = 0.2$ mA.

Solution (voltage gain): From Figure 4.47(c), we find the small-signal output voltage to be

$$V_o = g_{m1}V_{gs}(r_{o1}\|r_{o2})$$

A KVL equation around the outside loop produces

$$V_i = V_{gs} + V_o = V_{gs} + g_{m1}V_{gs}(r_{o1}\|r_{o2})$$

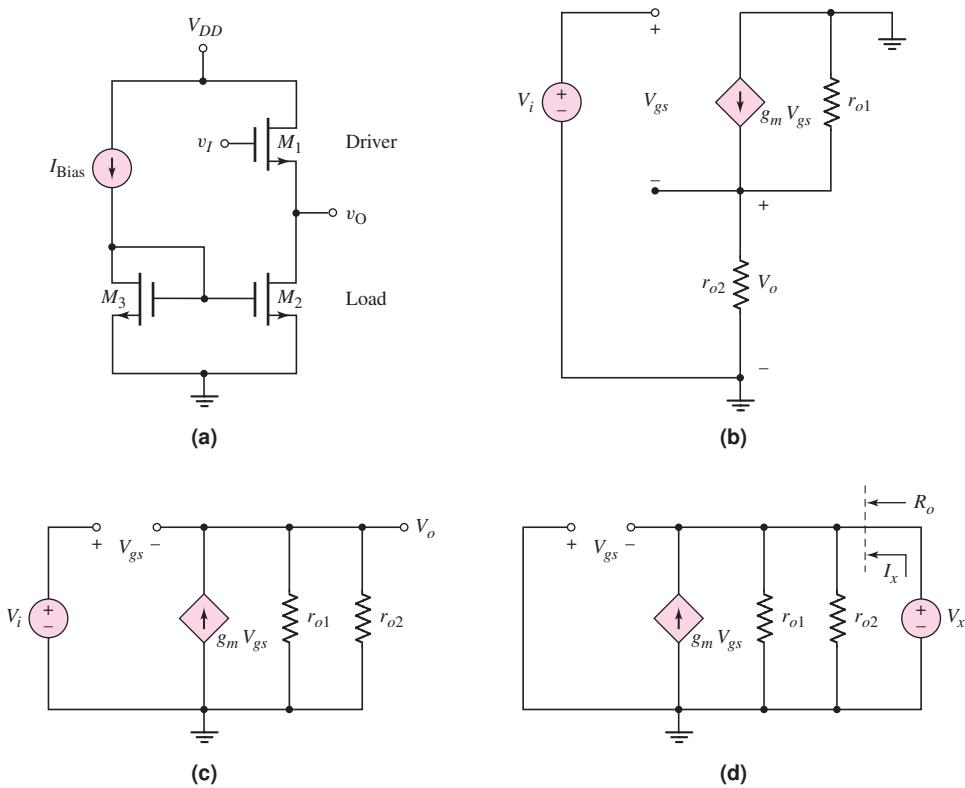


Figure 4.47 (a) All NMOS source-follower circuit, (b) small-signal equivalent circuit, (c) reconfiguration of small-signal equivalent circuit, and (d) small-signal equivalent circuit for determining the output resistance

or

$$V_{gs} = \frac{V_i}{1 + g_{m1}(r_{o1} \parallel r_{o2})}$$

Substituting this equation for \$V_{gs}\$ into the output voltage expression, we obtain the small-signal voltage gain as

$$A_v = \frac{V_o}{V_i} = \frac{g_{m1}(r_{o1} \parallel r_{o2})}{1 + g_{m1}(r_{o1} \parallel r_{o2})}$$

The small-signal equivalent circuit parameters are determined to be

$$g_{m1} = 2\sqrt{K_n I_{D1}} = 2\sqrt{(0.20)(0.20)} = 0.40 \text{ mA/V}$$

and

$$r_{o1} = r_{o2} = \frac{1}{\lambda I_D} = \frac{1}{(0.01)(0.20)} = 500 \text{ k}\Omega$$

The small-signal voltage gain is then

$$A_v = \frac{(0.40)(500 \parallel 500)}{1 + (0.40)(500 \parallel 500)}$$

or

$$A_v = 0.990$$

Solution (output resistance): The output resistance can be determined from the equivalent circuit shown in Figure 4.47(d). The independent source V_i is set equal to zero and a test voltage V_x is applied to the output.

Summing currents at the output node, we find

$$I_x + g_{m1}V_{gs} = \frac{V_x}{r_{o2}} + \frac{V_x}{r_{o1}}$$

From the circuit, we see that $V_{gs} = -V_x$. We then have

$$I_x = V_x \left(g_{m1} + \frac{1}{r_{o2}} + \frac{1}{r_{o1}} \right)$$

The output resistance is then given as

$$R_o = \frac{V_x}{I_x} = \frac{1}{g_{m1}} \| r_{o2} \| r_{o1}$$

We find

$$R_o = \frac{1}{0.40} \| 500 \| 500$$

or

$$R_o = 2.48 \text{ k}\Omega$$

Comment: A voltage gain of $A_v = 0.99$ is typical of a source-follower circuit. An output resistance of $R_o = 2.48 \text{ k}\Omega$ is relatively small for a MOSFET circuit and is also a characteristic of a source-follower circuit.

EXERCISE PROBLEM

Ex 4.14: The transconductance g_m of the transistor in the circuit of Figure 4.47 is to be changed by changing the bias current such that the output resistance of the circuit is $R_o = 2 \text{ k}\Omega$. Assume all other parameters are as given in Example 4.14. (a) What are the required values of g_m and I_{Bias} ? (b) Using the results of part (a), what is the small-signal voltage gain? (Ans. (a) $I_D = 0.3125 \text{ mA}$; (b) $A_v = 0.988$)

COMPUTER ANALYSIS EXERCISE

PS 4.1: Using a PSpice analysis, investigate the small-signal voltage gain and output resistance of the source-follower circuit shown in Figure 4.47 taking into account the body effect.

CMOS Common-Gate Amplifier

Figure 4.48(a) shows a common-gate circuit. We see that in this common-gate circuit, the active load is the PMOS device M_2 . The input signal is applied to the source of M_1 and the output is at the drain of M_1 .

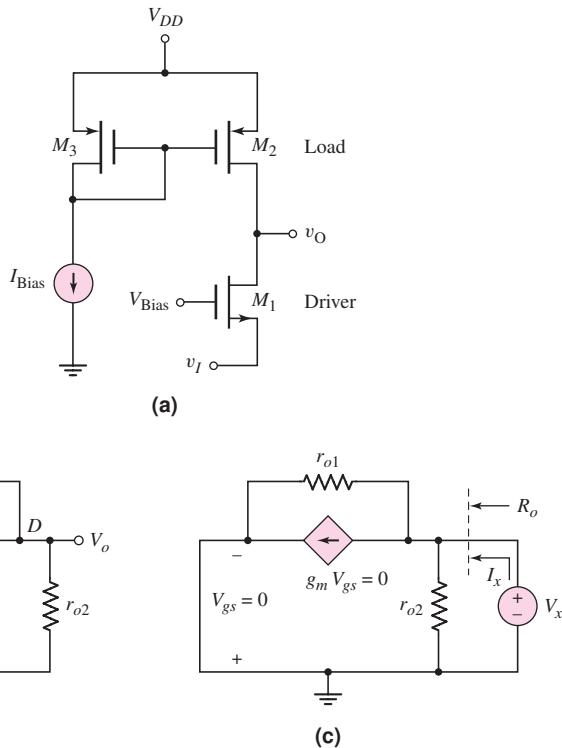


Figure 4.48 (a) CMOS common-gate amplifier, (b) small-signal equivalent circuit, and (c) small-signal equivalent circuit for determining the output resistance

The small-signal equivalent circuit of the common-gate circuit is shown in Figure 4.48(b).

EXAMPLE 4.15

Objective: Determine the small-signal voltage gain and output resistance of the common-gate circuit shown in Figure 4.48(a).

Assume the reference bias current is $I_{Bias} = 0.20$ mA and the bias voltage is $V_{DD} = 3.3$ V. Assume that the transistor parameters are $V_{TN} = +0.4$ V, $V_{TP} = -0.4$ V, $K_n = 0.20$ mA/V², $K_p = 0.20$ mA/V², and $\lambda_n = \lambda_p = 0.01$ V⁻¹.

We may note that, since M_2 and M_3 are matched transistors and have the same source-to-gate voltage, the bias current in M_1 is $I_{D1} = I_{Bias} = 0.20$ mA.

Solution (voltage gain): From Figure 4.48(b), we can sum currents at the output node and obtain

$$\frac{V_o}{r_{o2}} + g_{m1} V_{gs} + \frac{V_o - (-V_{gs})}{r_{o1}} = 0$$

or

$$V_o \left(\frac{1}{r_{o2}} + \frac{1}{r_{o1}} \right) + V_{gs} \left(g_{m1} + \frac{1}{r_{o1}} \right) = 0$$

From the circuit, we see that $V_{gs} = -V_i$. We then find the small-signal voltage gain to be

$$A_v = \frac{\left(g_{m1} + \frac{1}{r_{o1}}\right)}{\left(\frac{1}{r_{o2}} + \frac{1}{r_{o1}}\right)}$$

We find the small-signal equivalent circuit parameters to be

$$g_{m1} = 2\sqrt{K_n I_{D1}} = 2\sqrt{(0.20)(0.20)} = 0.40 \text{ mA/V}$$

and

$$r_{o1} = r_{o2} = \frac{1}{\lambda I_{D1}} = \frac{1}{(0.01)(0.20)} = 500 \text{ k}\Omega$$

We then find

$$A_v = \frac{\left(0.40 + \frac{1}{500}\right)}{\left(\frac{1}{500} + \frac{1}{500}\right)}$$

or

$$A_v = 101$$

Solution (output resistance): The output resistance can be found from Figure 4.48(c). Summing currents at the output node, we find

$$I_x = \frac{V_x}{r_{o2}} + g_{m1} V_{gs} + \frac{V_x - (-V_{gs})}{r_{o1}}$$

However, $V_{gs} = 0$ so that $g_{m1} V_{gs} = 0$. We then find

$$R_o = \frac{V_x}{I_x} = r_{o1} \| r_{o2} = 500 \| 500$$

or

$$R_o = 250 \text{ k}\Omega$$

Comment: A voltage gain of $A_v = +101$ is typical of a common-gate amplifier. The output signal is in phase with respect to the input signal and the gain is relatively large. Also, a large output resistance of $R_o = 250 \text{ k}\Omega$ is typical of a common-gate amplifier in that the circuit acts like a current source.

EXERCISE PROBLEM

Ex 4.15: The transconductance g_m of the transistor in the circuit of Figure 4.48 is to be changed by changing the bias current such that the small-signal voltage gain is $A_v = 120$. Assume all other parameters are as given in Example 4.15. (a) What are the required values of g_m and I_{Bias} ? (b) Using the results of part (a), what is the output resistance? (Ans. (a) $I_D = 0.14 \text{ mA}$, $g_m = 0.335 \text{ mA/V}$; (b) $R_o = 357 \text{ k}\Omega$)

COMPUTER ANALYSIS EXERCISE

PS 4.2: Using a PSpice analysis, investigate the small-signal voltage gain and output resistance of the common-gate amplifier shown in Figure 4.48 taking into account the body effect.

Test Your Understanding

TYU 4.11 For the enhancement load amplifier shown in Figure 4.39(a), the parameters are: $V_{TND} = V_{TNL} = 0.8$ V, $k'_n = 40 \mu\text{A}/\text{V}^2$, $(W/L)_D = 80$, $(W/L)_L = 1$, and $V_{DD} = 5$ V. Determine the small-signal voltage gain. Determine V_{GS} such that the Q -point is in the middle of the saturation region. (Ans. $A_v = -8.94$, $V_{GS} = 1.01$ V)

4.8 MULTISTAGE AMPLIFIERS

Objective: • Analyze multitransistor or multistage amplifiers and understand the advantages of these circuits over single-transistor amplifiers.

In most applications, a single-transistor amplifier will not be able to meet the combined specifications of a given amplification factor, input resistance, and output resistance. For example, the required voltage gain may exceed that which can be obtained in a single-transistor circuit. We will consider, here, the ac analysis of the two multitransistor circuits investigated in Chapter 3.

4.8.1 Multistage Amplifier: Cascade Circuit

The circuit shown in Figure 4.49 is a cascade of a common-source amplifier followed by a source-follower amplifier. As shown previously, the common-source amplifier

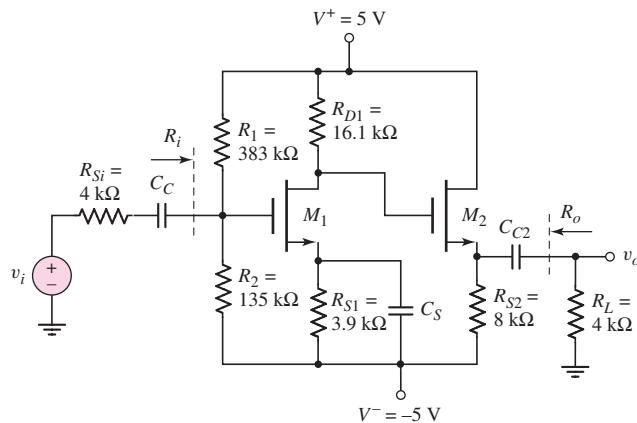


Figure 4.49 Common-source amplifier in cascade with source follower

provides a small-signal voltage gain and the source follower has a low output impedance and provides the required output current. The resistor values are those determined in Section 3.5.1 of the previous chapter.

The midband small-signal voltage gain of the multistage amplifier is determined by assuming that all external coupling capacitors act as short circuits and inserting the small-signal equivalent circuits for the transistors.

EXAMPLE 4.16

Objective: Determine the small-signal voltage gain of a multistage cascade circuit.

Consider the circuit shown in Figure 4.49. The transistor parameters are $K_{n1} = 0.5 \text{ mA/V}^2$, $K_{n2} = 0.2 \text{ mA/V}^2$, $V_{TN1} = V_{TN2} = 1.2 \text{ V}$, and $\lambda_1 = \lambda_2 = 0$. The quiescent drain currents are $I_{D1} = 0.2 \text{ mA}$ and $I_{D2} = 0.5 \text{ mA}$.

Solution: The small-signal equivalent circuit is shown in Figure 4.50. The small-signal transconductance parameters are

$$g_{m1} = 2\sqrt{K_{n1}I_{D1}} = 2\sqrt{(0.5)(0.2)} = 0.632 \text{ mA/V}$$

and

$$g_{m2} = 2\sqrt{K_{n2}I_{D2}} = 2\sqrt{(0.2)(0.5)} = 0.632 \text{ mA/V}$$

The output voltage is

$$V_o = g_{m2}V_{gs2}(R_{S2}\parallel R_L)$$

Also,

$$V_{gs2} + V_o = -g_{m1}V_{gs1}R_{D1}$$

where

$$V_{gs1} = \left(\frac{R_i}{R_i + R_{Si}} \right) \cdot V_i$$

Then

$$V_{gs2} = -g_{m1}R_{D1} \left(\frac{R_i}{R_i + R_{Si}} \right) \cdot V_i - V_o$$

Therefore

$$V_o = g_{m2} \left[-g_{m1}R_{D1} \left(\frac{R_i}{R_i + R_{Si}} \right) \cdot V_i - V_o \right] (R_{S2}\parallel R_L)$$

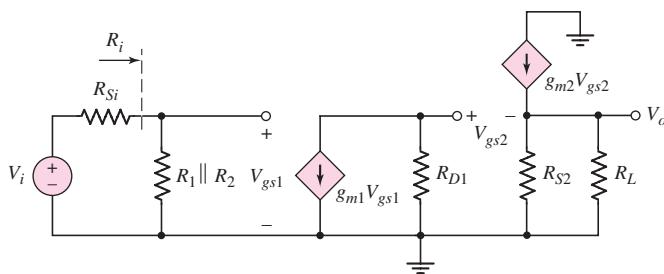


Figure 4.50 Small-signal equivalent circuit of NMOS cascade circuit

The small-signal voltage gain is then

$$A_v = \frac{V_o}{V_i} = \frac{-g_{m1}g_{m2}R_{D1}(R_{S2}\parallel R_L)}{1 + g_{m2}(R_{S2}\parallel R_L)} \cdot \left(\frac{R_i}{R_i + R_{Si}} \right)$$

or

$$A_v = \frac{-(0.632)(0.632)(16.1)(8\parallel 4)}{1 + (0.632)(8\parallel 4)} \cdot \left(\frac{100}{100 + 4} \right) = -6.14$$

Comment: Since the small-signal voltage gain of the source follower is slightly less than 1, the overall gain is due essentially to the common-source input stage. Also, as shown previously, the output resistance of the source follower is small, which is desirable in many applications.

EXERCISE PROBLEM

Ex 4.16: For the cascade circuit shown in Figure 4.49, the transistor and circuit parameters are given in Example 4.16. Calculate the small-signal output resistance R_o . (The small-signal equivalent circuit is shown in Figure 4.50.) (Ans. $R_o = 1.32 \text{ k}\Omega$)

4.8.2 Multistage Amplifier: Cascode Circuit

Figure 4.51 shows a cascode circuit with n-channel MOSFETs. Transistor M_1 is connected in a common-source configuration and M_2 is connected in a common-gate configuration. The advantage of this type of circuit is a higher frequency response, which will be discussed in Chapter 7. The resistor values are those determined in Section 3.5.2 of the previous chapter.

We will consider additional multistage and multitransistor circuits in Chapters 11 and 13.

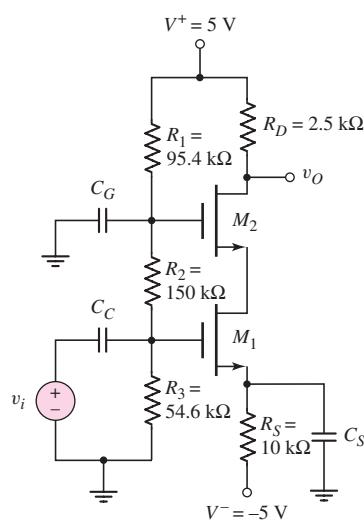


Figure 4.51 NMOS cascode circuit

EXAMPLE 4.17

Objective: Determine the small-signal voltage gain of a cascode circuit.

Consider the cascode circuit shown in Figure 4.51. The transistor parameters are $K_{n1} = K_{n2} = 0.8 \text{ mA/V}^2$, $V_{TN1} = V_{TN2} = 1.2 \text{ V}$, and $\lambda_1 = \lambda_2 = 0$. The quiescent drain current is $I_D = 0.4 \text{ mA}$ in each transistor. The input signal to the circuit is assumed to be an ideal voltage source.

Solution: Since the transistors are identical and since the current in the two transistors is the same, the small-signal transconductance parameters are

$$g_{m1} = g_{m2} = 2\sqrt{K_n I_D} = 2\sqrt{(0.8)(0.4)} = 1.13 \text{ mA/V}$$

The small-signal equivalent circuit is shown in Figure 4.52. Transistor M_1 supplies the source current of M_2 with the signal current ($g_{m1}V_i$). Transistor M_2 acts as a current follower and passes this current on to its drain terminal. The output voltage is therefore

$$V_o = -g_{m1}V_{gs1}R_D$$

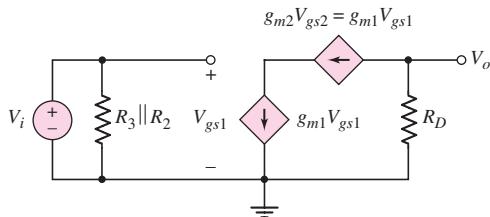


Figure 4.52 Small-signal equivalent circuit of NMOS cascode circuit

Since $V_{gs1} = V_i$, the small-signal voltage gain is

$$A_v = \frac{V_o}{V_i} = -g_{m1}R_D$$

or

$$A_v = -(1.13)(2.5) = -2.83$$

Comment: The small-signal voltage gain is essentially the same as that of a single common-source amplifier stage. The addition of a common-gate transistor will increase the frequency bandwidth, as we will see in a later chapter.

EXERCISE PROBLEM

Ex 4.17: The transistor parameters of the NMOS cascode circuit in Figure 4.51 are $V_{TN1} = V_{TN2} = 0.8 \text{ V}$, $K_{n1} = K_{n2} = 3 \text{ mA/V}^2$, and $\lambda_1 = \lambda_2 = 0$. (a) Determine I_{DQ} , V_{DSQ1} , and V_{DSQ2} . (b) Determine the small-signal voltage gain. (Ans. (a) $I_{DQ} = 0.471 \text{ mA}$, $V_{DSQ1} = 2.5 \text{ V}$, $V_{DSQ2} = 1.61 \text{ V}$; (b) $A_v = -5.94$)

*Test Your Understanding

TYU 4.12 The transistor parameters of the circuit in Figure 4.49 are $V_{TN1} = V_{TN2} = 0.6$ V, $K_{n1} = 1.5$ mA/V², $K_{n2} = 2$ mA/V², and $\lambda_1 = \lambda_2 = 0$. (a) Find I_{DQ1} , I_{DQ2} , V_{DSQ1} , and V_{DSQ2} . (b) Determine the small-signal voltage gain. (c) Find the output resistance R_o . (Ans. (a) $I_{DQ1} = 0.3845$ mA, $I_{DQ2} = 0.349$ mA, $V_{DSQ1} = 2.31$ V, $V_{DSQ2} = 7.21$ V; (b) $A_v = -20.3$; (c) $R_o = 402 \Omega$)

4.9 BASIC JFET AMPLIFIERS

Objective: • Develop the small-signal model of JFET devices and analyze basic JFET amplifiers.

Like MOSFETs, JFETs can be used to amplify small time-varying signals. Initially, we will develop the small-signal model and equivalent circuit of the JFET. We will then use the model in the analysis of JFET amplifiers.

4.9.1 Small-Signal Equivalent Circuit

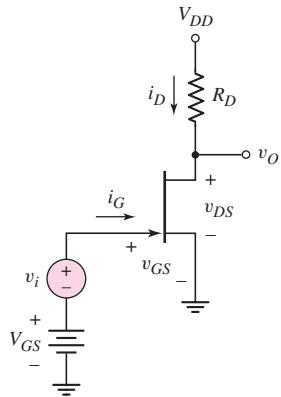


Figure 4.53 JFET common-source circuit with time-varying signal source in series with gate dc source

Figure 4.53 shows a JFET circuit with a time-varying signal applied to the gate. The instantaneous gate-to-source voltage is

$$v_{GS} = V_{GS} + v_i = V_{GS} + v_{gs} \quad (4.53)$$

where v_{gs} is the small-signal gate-to-source voltage. Assuming the transistor is biased in the saturation region, the instantaneous drain current is

$$i_D = I_{DSS} \left(1 - \frac{v_{GS}}{V_P} \right)^2 \quad (4.54)$$

where I_{DSS} is the saturation current and V_P is the pinchoff voltage. Substituting Equation (4.53) into (4.54), we obtain

$$i_D = I_{DSS} \left[\left(1 - \frac{V_{GS}}{V_P} \right) - \left(\frac{v_{gs}}{V_P} \right) \right]^2 \quad (4.55)$$

If we expand the squared term, we have

$$i_D = I_{DSS} \left(1 - \frac{V_{GS}}{V_P} \right)^2 - 2I_{DSS} \left(1 - \frac{V_{GS}}{V_P} \right) \left(\frac{v_{gs}}{V_P} \right) + I_{DSS} \left(\frac{v_{gs}}{V_P} \right)^2 \quad (4.56)$$

The first term in Equation (4.56) is the dc or quiescent drain current I_{DQ} , the second term is the time-varying drain current component, which is linearly related to the signal voltage v_{gs} , and the third term is proportional to the square of the signal voltage. As in the case of the MOSFET, the third term produces a nonlinear distortion in the output current. To minimize this distortion, we will usually impose the following condition:

$$\left| \frac{v_{gs}}{V_P} \right| \ll 2 \left(1 - \frac{V_{GS}}{V_P} \right) \quad (4.57)$$

Equation (4.57) represents the small-signal condition that must be satisfied for JFET amplifiers to be linear.

Neglecting the term v_{gs}^2 in Equation (4.56), we can write

$$i_D = I_{DQ} + i_d \quad (4.58)$$

where the time-varying signal current is

$$i_d = +\frac{2I_{DSS}}{(-V_P)} \left(1 - \frac{V_{GS}}{V_P}\right) v_{gs} \quad (4.59)$$

The constant relating the small-signal drain current and small-signal gate-to-source voltage is the transconductance g_m . We can write

$$i_d = g_m v_{gs} \quad (4.60)$$

where

$$g_m = +\frac{2I_{DSS}}{(-V_P)} \left(1 - \frac{V_{GS}}{V_P}\right) \quad (4.61)$$

Since V_P is negative for n-channel JFETs, the transconductance is positive. A relationship that applies to both n-channel and p-channel JFETs is

$$g_m = \frac{2I_{DSS}}{|V_P|} \left(1 - \frac{V_{GS}}{V_P}\right) \quad (4.62)$$

We can also obtain the transconductance from

$$g_m = \left. \frac{\partial i_D}{\partial v_{GS}} \right|_{v_{GS}=V_{GSQ}} \quad (4.63)$$

Since the transconductance is directly proportional to the saturation current I_{DSS} , the transconductance is also a function of the width-to-length ratio of the transistor.

Since we are looking into a reverse-biased pn junction, we assume that the input gate current i_g is zero, which means that the small-signal input resistance is infinite. Equation (4.54) can be expanded to take into account the finite output resistance of a JFET biased in the saturation region. The equation becomes

$$i_D = I_{DSS} \left(1 - \frac{V_{GS}}{V_P}\right)^2 (1 + \lambda v_{DS}) \quad (4.64)$$

The small-signal output resistance is

$$r_o = \left. \left(\frac{\partial i_D}{\partial v_{DS}} \right)^{-1} \right|_{v_{GS}=\text{const.}} \quad (4.65)$$

Using Equation (4.64), we obtain

$$r_o = \left[\lambda I_{DSS} \left(1 - \frac{V_{GS}}{V_P}\right)^2 \right]^{-1} \quad (4.66(a))$$

or

$$r_o \cong [\lambda I_{DQ}]^{-1} = \frac{1}{\lambda I_{DQ}} \quad (4.66(b))$$

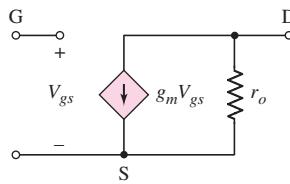


Figure 4.54 Small-signal equivalent circuit of n-channel JFET

The small-signal equivalent circuit of the n-channel JFET, shown in Figure 4.54, is exactly the same as that of the n-channel MOSFET. The small-signal equivalent circuit of the p-channel JFET is also the same as that of the p-channel MOSFET. However, the polarity of the controlling gate-to-source voltage and the direction of the dependent current source are reversed from those of the n-channel device.

4.9.2 Small-Signal Analysis

Since the small-signal equivalent circuit of the JFET is the same as that of the MOSFET, the small-signal analyses of the two types of circuits are identical. For illustration purposes, we will analyze two JFET circuits.

EXAMPLE 4.18

Objective: Determine the small-signal voltage gain of a JFET amplifier.

Consider the circuit shown in Figure 4.55 with transistor parameters $I_{DSS} = 12 \text{ mA}$, $V_P = -4 \text{ V}$, and $\lambda = 0.008 \text{ V}^{-1}$. Determine the small-signal voltage gain $A_v = v_o/v_i$.

Solution: The dc quiescent gate-to-source voltage is determined from

$$V_{GSQ} = \left(\frac{R_2}{R_1 + R_2} \right) V_{DD} - I_{DQ} R_S$$

where

$$I_{DQ} = I_{DSS} \left(1 - \frac{V_{GSQ}}{V_P} \right)^2$$

Combining these two equations produces

$$V_{GSQ} = \left(\frac{180}{180 + 420} \right) (20) - (12)(2.7) \left(1 - \frac{V_{GSQ}}{(-4)} \right)^2$$

which reduces to

$$2.025 V_{GSQ}^2 + 17.25 V_{GSQ} + 26.4 = 0$$

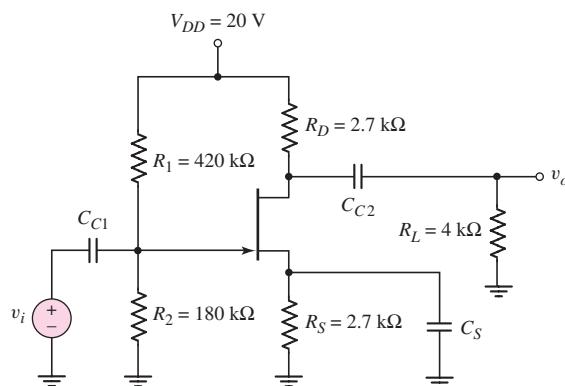


Figure 4.55 Common-source JFET circuit with source resistor and source bypass capacitor

The appropriate solution is

$$V_{GSQ} = -2.0 \text{ V}$$

The quiescent drain current is

$$I_{DQ} = I_{DSS} \left(1 - \frac{V_{GSQ}}{V_P} \right)^2 = (12) \left(1 - \frac{(-2.0)}{(-4)} \right)^2 = 3.00 \text{ mA}$$

The small-signal parameters are then

$$g_m = \frac{2I_{DSS}}{(-V_P)} \left(1 - \frac{V_{GS}}{V_P} \right) = \frac{2(12)}{(4)} \left(1 - \frac{(-2.0)}{(-4)} \right) = 3.00 \text{ mA/V}$$

and

$$r_o = \frac{1}{\lambda I_{DQ}} = \frac{1}{(0.008)(3.00)} = 41.7 \text{ k}\Omega$$

The small-signal equivalent circuit is shown in Figure 4.56.

Since $V_{gs} = V_i$, the small-signal voltage gain is

$$A_v = \frac{V_o}{V_i} = -g_m(r_o \parallel R_D \parallel R_L)$$

or

$$A_v = -(3.0)(41.7 \parallel 2.7 \parallel 4) = -4.66$$

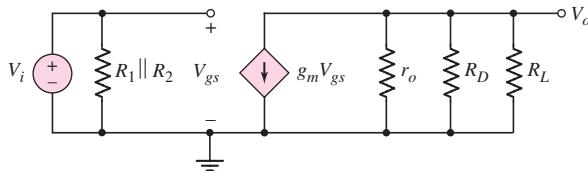


Figure 4.56 Small-signal equivalent circuit of common-source JFET, assuming bypass capacitor acts as a short circuit

Comment: The voltage gain of JFET amplifiers is the same order of magnitude as that of MOSFET amplifiers.

EXERCISE PROBLEM

Ex 4.18: For the JFET amplifier shown in Figure 4.55, the transistor parameters are: $I_{DSS} = 4 \text{ mA}$, $V_P = -3 \text{ V}$, and $\lambda = 0.005 \text{ V}^{-1}$. Let $R_L = 4 \text{ k}\Omega$, $R_S = 2.7 \text{ k}\Omega$, and $R_1 + R_2 = 500 \text{ k}\Omega$. Redesign the circuit such that $I_{DQ} = 1.2 \text{ mA}$ and $V_{DSQ} = 12 \text{ V}$. Calculate the small-signal voltage gain. (Ans. $R_D = 3.97 \text{ k}\Omega$, $R_1 = 453 \text{ k}\Omega$, $R_2 = 47 \text{ k}\Omega$, $A_v = -2.87$)

DESIGN EXAMPLE 4.19

Objective: Design a JFET source-follower circuit with a specified small-signal voltage gain.

For the source-follower circuit shown in Figure 4.57, the transistor parameters are: $I_{DSS} = 12 \text{ mA}$, $V_P = -4 \text{ V}$, and $\lambda = 0.01 \text{ V}^{-1}$. Determine R_S and I_{DQ} such that the small-signal voltage gain is at least $A_v = v_o/v_i = 0.90$.

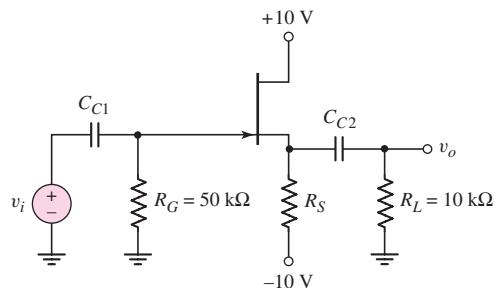


Figure 4.57 JFET source-follower circuit

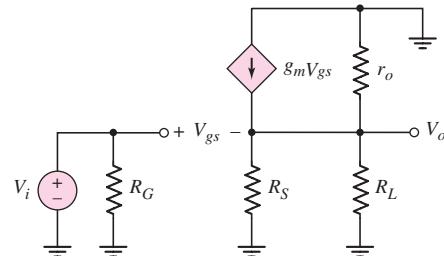


Figure 4.58 Small-signal equivalent circuit of JFET source-follower circuit

Solution: The small-signal equivalent circuit is shown in Figure 4.58. The output voltage is

$$V_o = g_m V_{gs} (R_S \parallel R_L \parallel r_o)$$

Also

$$V_i = V_{gs} + V_o$$

or

$$V_{gs} = V_i - V_o$$

Therefore, the output voltage is

$$V_o = g_m (V_i - V_o) (R_S \parallel R_L \parallel r_o)$$

The small-signal voltage gain becomes

$$A_v = \frac{V_o}{V_i} = \frac{g_m (R_S \parallel R_L \parallel r_o)}{1 + g_m (R_S \parallel R_L \parallel r_o)}$$

As a first approximation, assume r_o is sufficiently large for the effect of r_o to be neglected.

The transconductance is

$$g_m = \frac{2I_{DSS}}{(-V_P)} \left(1 - \frac{V_{GS}}{V_P}\right) = \frac{2(12)}{4} \left(1 - \frac{V_{GS}}{(-4)}\right)$$

If we pick a nominal transconductance value of $g_m = 2 \text{ mA/V}$, then $V_{GS} = -2.67 \text{ V}$ and the quiescent drain current is

$$I_{DQ} = I_{DSS} \left(1 - \frac{V_{GS}}{V_P}\right)^2 = (12) \left(1 - \frac{(-2.67)}{(-4)}\right)^2 = 1.335 \text{ mA}$$

The value of R_S is then determined from

$$R_S = \frac{-V_{GS} - (-10)}{I_{DQ}} = \frac{2.67 + 10}{1.335} = 9.49 \text{ k}\Omega$$

Also, the value of r_o is

$$r_o = \frac{1}{\lambda I_{DQ}} = \frac{1}{(0.01)(1.335)} = 74.9 \text{ k}\Omega$$

The small-signal voltage gain, including the effect of r_o , is

$$A_v = \frac{g_m(R_S \| R_L \| r_o)}{1 + g_m(R_S \| R_L \| r_o)} = \frac{(2)(9.49 \| 10 \| 74.9)}{1 + (2)(9.49 \| 10 \| 74.9)} = 0.902$$

Comment: This particular design meets the design criteria, but the solution is not unique.

EXERCISE PROBLEM

Ex 4.19: Reconsider the source-follower circuit shown in Figure 4.57 with transistor parameters $I_{DSS} = 8 \text{ mA}$, $V_P = -3.5 \text{ V}$, and $\lambda = 0.01 \text{ V}^{-1}$. (a) Design the circuit such that $I_{DQ} = 2 \text{ mA}$. (b) Calculate the small-signal voltage gain if R_L approaches infinity. (c) Determine the value of R_L at which the small-signal gain is reduced by 20 percent from its value for (b). (Ans. (a) $R_S = 5.88 \text{ k}\Omega$, (b) $A_v = 0.923$, $R_L = 1.61 \text{ k}\Omega$)

In Example 4.19, we chose a value of transconductance and continued through the design. A more detailed examination shows that both g_m and R_S depend upon the drain current I_{DQ} in such a way that the product $g_m R_S$ is approximately a constant. This means the small-signal voltage gain is insensitive to the initial value of the transconductance.

Test Your Understanding

TYU 4.13 Reconsider the JFET amplifier shown in Figure 4.55 with transistor parameters given in Example 4.18. Determine the small-signal voltage gain if a $20 \text{ k}\Omega$ resistor is in series with the signal source v_i . (Ans. $A_v = -3.98$)

***TYU 4.14** For the circuit shown in Figure 4.59, the transistor parameters are: $I_{DSS} = 6 \text{ mA}$, $|V_P| = 2 \text{ V}$, and $\lambda = 0$. (a) Calculate the quiescent drain current and drain-to-source voltage of each transistor. (b) Determine the overall small-signal voltage gain $A_v = v_o/v_i$. (Ans. (a) $I_{DQ1} = 1 \text{ mA}$, $V_{SDQ1} = 12 \text{ V}$, $I_{DQ2} = 1.27 \text{ mA}$, $V_{DSQ2} = 14.9 \text{ V}$; (b) $A_v = -2.05$)

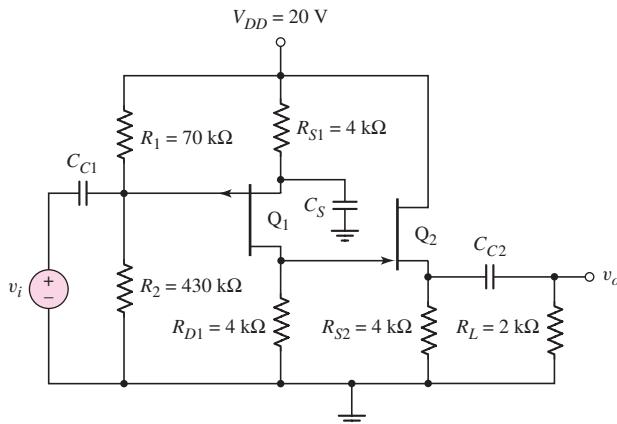


Figure 4.59 Figure for Exercise TYU 4.14

4.10 DESIGN APPLICATION: A TWO-STAGE AMPLIFIER

Objective: • Design a two-stage MOSFET circuit to amplify the output of a sensor.

Specifications: Assume the resistance R_2 in the voltage divider circuit in Figure 4.60 varies linearly as a function of temperature, pressure, or some other variable. The output of the amplifier is to be zero volts when $\delta = 0$.

Design Approach: The amplifier configuration to be designed is shown in Figure 4.60. A resistor R_1 will be chosen such that the voltage divider between R_1 and R_2 will produce a dc voltage v_I that is negative. A negative gate voltage to M_1 then means that the resistance R_{S1} does not need to be so large.

Choices: Assume NMOS and PMOS transistors are available with parameters $V_{TN} = 1$ V, $V_{TP} = -1$ V, $K_n = K_p = 2$ mA/V², and $\lambda_n = \lambda_p \approx 0$.

Solution (voltage divider analysis): The voltage v_I can be written as

$$v_I = \left[\frac{R(1 + \delta)}{R(1 + \delta) + 3R} \right] (10) - 5 = \frac{(1 + \delta)(10)}{4 + \delta} - 5$$

or

$$v_I = \frac{(1 + \delta)(10) - 5(4 + \delta)}{4 + \delta} = \frac{-10 + 5\delta}{4 + \delta}$$

Assuming that $\delta \ll 4$, we then have

$$v_I = -2.5 + 1.25\delta$$

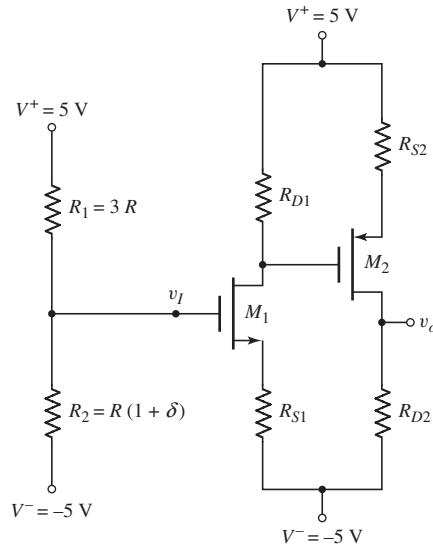


Figure 4.60 Two-stage MOSFET amplifier for design application

Solution (DC Design): We will choose $I_{D1} = 0.5 \text{ mA}$ and $I_{D2} = 1 \text{ mA}$. The gate-to-source voltages are determined to be:

$$0.5 = 2(V_{GS1} - 1)^2 \Rightarrow V_{GS1} = 1.5 \text{ V}$$

and

$$1 = 2(V_{SG2} - 1)^2 \Rightarrow V_{SG2} = 1.707 \text{ V}$$

We find $V_{S1} = V_I - V_{GS1} = -2.5 - 1.5 = -4 \text{ V}$. The resistor R_{S1} is then

$$R_{S1} = \frac{V_{S1} - V^-}{I_{D1}} = \frac{-4 - (-5)}{0.5} = 2 \text{ k}\Omega$$

Letting $V_{D1} = 1.5 \text{ V}$, we find the resistor R_{D1} to be

$$R_{D1} = \frac{V^+ - V_{D1}}{I_{D1}} = \frac{5 - 1.5}{0.5} = 7 \text{ k}\Omega$$

We have $V_{S2} = V_{D1} + V_{SG2} = 1.5 + 1.707 = 3.207 \text{ V}$. Then

$$R_{S2} = \frac{V^+ - V_{S2}}{I_{D2}} = \frac{5 - 3.207}{1} = 1.79 \text{ k}\Omega$$

For $V_O = 0$, we find

$$R_{D2} = \frac{V_O - V^-}{I_{D2}} = \frac{0 - (-5)}{1} = 5 \text{ k}\Omega$$

Solution (ac Analysis): The small-signal equivalent circuit is shown in Figure 4.61. We find $V_2 = -g_{m1}V_{gs1}R_{D1}$ and $V_{gs1} = V_i/(1 + g_{m1}R_{S1})$. We also find $V_o = g_{m2}V_{sg2}R_{D2}$ and $V_{sg2} = -V_2/(1 + g_{m2}R_{S2})$. Combining terms, we find

$$V_o = \frac{g_{m1}g_{m2}R_{D1}R_{D2}}{(1 + g_{m1}R_{S1})(1 + g_{m2}R_{S2})}V_i$$

The ac input signal is $V_i = 1.25 \delta$, so we have

$$V_o = \frac{(1.25)g_{m1}g_{m2}R_{D1}R_{D2}}{(1 + g_{m1}R_{S1})(1 + g_{m2}R_{S2})}\delta$$

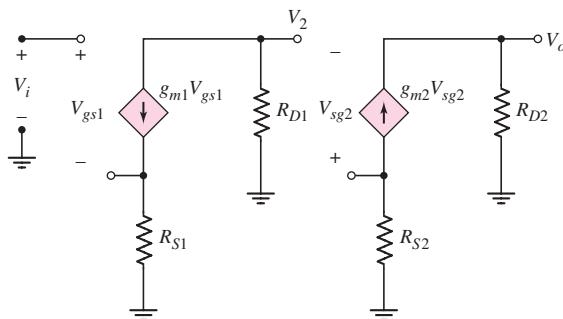


Figure 4.61 Small-signal equivalent circuit of two-stage MOSFET amplifier for design application

We find that

$$g_{m1} = 2\sqrt{K_n I_{D1}} = 2\sqrt{(2)(0.5)} = 2 \text{ mA/V}$$

and

$$g_{m2} = 2\sqrt{K_p I_{D2}} = 2\sqrt{(2)(1)} = 2.828 \text{ mA/V}$$

We then find

$$V_o = \frac{(1.25)(2)(2.828)(7)(5)}{[1 + (2)(2)][1 + (2.828)(1.79)]} \delta$$

or

$$V_o = 8.16\delta$$

Comment: Since the low-frequency input impedance to the gate of the NMOS is essentially infinite, there is no loading effect on the voltage divider circuit.

Design Pointer: As mentioned previously, by choosing the value of R_1 to be larger than R_2 , the dc voltage to the gate of M_1 is negative. A negative gate voltage implies that the required value of R_{S1} is reduced and can still establish the required current. Since the drain voltage at M_1 is positive, then by using a PMOS transistor in the second stage, the source resistor value of R_{S2} is also reduced. Smaller source resistances generate larger voltage gains.

4.11 SUMMARY

- The application of MOSFET transistors in linear amplifiers was emphasized in this chapter. The basic process by which a transistor circuit can amplify a small time-varying input signal was discussed.
- A small-signal equivalent circuit for the MOSFET transistor, which is used in the analysis and design of linear amplifiers, was developed.
- The three basic amplifier configurations were considered: the common-source, source-follower, and common-gate. These three circuits form the basic building blocks for more complex integrated circuits.
- The common-source circuit amplifies a time-varying voltage.
- The small-signal voltage gain of a source-follower circuit is approximately unity, but has a low output resistance.
- The common-gate circuit amplifies a time-varying voltage, and has a low input resistance and a large output resistance.
- Analysis of n-channel circuits with enhancement-load or depletion-load devices was performed. In addition, the analysis of CMOS circuits was carried out. These circuits are examples of all MOSFET circuits, which are developed throughout the remainder of the text.
- The small-signal equivalent circuit of a JFET was developed and used in the analysis of several configurations of JFET amplifiers.
- As an application, MOS transistors were incorporated in the design of a two-stage amplifier.

 **CHECKPOINT**

After studying this chapter, the reader should have the ability to:

- ✓ Explain graphically the amplification process in a simple MOSFET amplifier circuit.
- ✓ Describe the small-signal equivalent circuit of the MOSFET and to determine the values of the small-signal parameters.
- ✓ Apply the small-signal equivalent circuit to various MOSFET amplifier circuits to obtain the time-varying circuit characteristics.
- ✓ Characterize the small-signal voltage gain and output resistance of the common-source, source-follower, and common-gate amplifiers.
- ✓ Describe the operation of an NMOS amplifier with either an enhancement load, a depletion load, or a PMOS load.
- ✓ Apply the MOSFET small-signal equivalent circuit in the analysis of multistage amplifier circuits.
- ✓ Describe the operation and analyze basic JFET amplifier circuits.

 **REVIEW QUESTIONS**

1. Discuss, using the concept of a load line, how a simple common-source circuit can amplify a time-varying signal.
2. How does the transistor width-to-length ratio affect the small-signal voltage gain of a common-source amplifier?
3. Discuss the physical meaning of the small-signal circuit parameter r_o .
4. How does the body effect change the small-signal equivalent circuit of the MOSFET?
5. Sketch a simple common-source amplifier circuit and discuss the general ac circuit characteristics (voltage gain and output resistance).
6. Discuss the general conditions under which a common-source amplifier would be used.
7. Why, in general, is the magnitude of the voltage gain of a common-source amplifier relatively small?
8. What are the changes in dc and ac characteristics of a common-source amplifier when a source resistor and a source bypass capacitor are incorporated in the design?
9. Sketch a simple source-follower amplifier circuit and discuss the general ac circuit characteristics (voltage gain and output resistance).
10. Sketch a simple common-gate amplifier circuit and discuss the general ac circuit characteristics (voltage gain and output resistance).
11. Discuss the general conditions under which a source-follower or a common-gate amplifier would be used.
12. Compare the ac circuit characteristics of the common-source, source-follower, and common-gate circuits.
13. State the advantage of using transistors in place of resistors in MOSFET integrated circuits.
14. State at least two reasons why a multistage amplifier circuit would be required in a design compared to using a single-stage circuit.

 PROBLEMS

Section 4.1 The MOSFET Amplifier

- 4.1 An NMOS transistor has parameters $V_{TN} = 0.4$ V, $k'_n = 100 \mu\text{A}/\text{V}^2$, and $\lambda = 0.02 \text{ V}^{-1}$. (a) (i) Determine the width-to-length ratio W/L such that $g_m = 0.5$ mA/V at $I_{DQ} = 0.5$ mA when biased in the saturation region. (ii) Calculate the required value of V_{GSQ} . (b) Repeat part (a) for $I_{DQ} = 0.15$ mA.
- 4.2 A PMOS transistor has parameters $V_{TP} = -0.6$ V, $k'_p = 40 \mu\text{A}/\text{V}^2$, and $\lambda = 0.015 \text{ V}^{-1}$. (a) (i) Determine the width-to-length ratio (W/L) such that $g_m = 1.2$ mA/V at $I_{DQ} = 0.15$ mA. (ii) What is the required value of V_{SGQ} ? (b) Repeat part (a) for $I_{DQ} = 0.50$ mA.
- 4.3 An NMOS transistor is biased in the saturation region at a constant V_{GS} . The drain current is $I_D = 3$ mA at $V_{DS} = 5$ V and $I_D = 3.4$ mA at $V_{DS} = 10$ V. Determine λ and r_o .
- 4.4 The minimum value of small-signal resistance of a PMOS transistor is to be $r_o = 100$ k Ω . If $\lambda = 0.012 \text{ V}^{-1}$, calculate the maximum allowed value of I_D .
- 4.5 An n-channel MOSFET is biased in the saturation region at a constant V_{GS} . (a) The drain current is $I_D = 0.250$ mA at $V_{DS} = 1.5$ V and $I_D = 0.258$ mA at $V_{DS} = 3.3$ V. Determine the value of λ and r_o . (b) Using the results of part (a), determine I_D at $V_{DS} = 5$ V.
- 4.6 The value of λ for a MOSFET is 0.02 V^{-1} . (a) What is the value of r_o at (i) $I_D = 50 \mu\text{A}$ and at (ii) $I_D = 500 \mu\text{A}$? (b) If V_{DS} increases by 1 V, what is the percentage increase in I_D for the conditions given in part (a)?
- 4.7 A MOSFET with $\lambda = 0.01 \text{ V}^{-1}$ is biased in the saturation region at $I_D = 0.5$ mA. If V_{GS} and V_{DS} remain constant, what are the new values of I_D and r_o if the channel length L is doubled?
- 4.8 The parameters of the circuit in Figure 4.1 are $V_{DD} = 3.3$ V and $R_D = 5$ k Ω . The transistor parameters are $k'_n = 100 \mu\text{A}/\text{V}^2$, $W/L = 40$, $V_{TN} = 0.4$ V, and $\lambda = 0.025 \text{ V}^{-1}$. (a) Find I_{DQ} and V_{GSQ} such that $V_{DSQ} = 1.5$ V. (b) Determine the small-signal voltage gain.
- 4.9 The circuit shown in Figure 4.1 has parameters $V_{DD} = 2.5$ V and $R_D = 10$ k Ω . The transistor is biased at $I_{DQ} = 0.12$ mA. The transistor parameters are $V_{TN} = 0.3$ V, $k'_n = 100 \mu\text{A}/\text{V}^2$, and $\lambda = 0$. (a) Design the W/L ratio of the transistor such that the small-signal voltage gain is $A_v = -3.8$. (b) Repeat part (a) for $A_v = -5.0$.
- 4.10 For the circuit shown in Figure 4.1, the transistor parameters are $V_{TN} = 0.6$ V, $k'_n = 80 \mu\text{A}/\text{V}^2$, and $\lambda = 0.015 \text{ V}^{-1}$. Let $V_{DD} = 5$ V. (a) Design the transistor width-to-length ratio W/L and the resistance R_D such that $I_{DQ} = 0.5$ mA, $V_{GSQ} = 1.2$ V, and $V_{DSQ} = 3$ V. (b) Determine g_m and r_o . (c) Determine the small-signal voltage gain $A_v = v_o/v_i$.
- *4.11 In our analyses, we assumed the small-signal condition given by Equation (4.4). Now consider Equation (4.3(b)) and let $v_{gs} = V_{gs} \sin \omega t$. Show that the ratio of the signal at frequency 2ω to the signal at frequency ω is given by $V_{gs}/[4(V_{GS} - V_{TN})]$. This ratio, expressed in a percentage, is called the **second-harmonic distortion**. [Hint: Use the trigonometric identity $\sin^2 \theta = \frac{1}{2} - \frac{1}{2} \cos 2\theta$.]
- 4.12 Using the results of Problem 4.11, find the peak amplitude V_{gs} that produces a second-harmonic distortion of 1 percent if $V_{GS} = 3$ V and $V_{TN} = 1$ V.

Section 4.3 The Common-Source Amplifier

- 4.13 Consider the circuit in Figure 4.14 in the text. The circuit parameters are $V_{DD} = 3.3$ V, $R_D = 8$ k Ω , $R_1 = 240$ k Ω , $R_2 = 60$ k Ω , and $R_{Si} = 2$ k Ω . The transistor parameters are $V_{TN} = 0.4$ V, $k'_n = 100 \mu\text{A/V}^2$, $W/L = 80$, and $\lambda = 0.02 \text{ V}^{-1}$. (a) Determine the quiescent values I_{DQ} and V_{DSQ} . (b) Find the small-signal parameters g_m and r_o . (c) Determine the small-signal voltage gain.
- 4.14 A common-source amplifier, such as shown in Figure 4.14 in the text, has parameters $r_o = 100$ k Ω and $R_D = 5$ k Ω . Determine the transconductance of the transistor if the small-signal voltage gain is $A_v = -10$. Assume $R_{Si} = 0$.
- 4.15 For the NMOS common-source amplifier in Figure P4.15, the transistor parameters are: $V_{TN} = 0.8$ V, $K_n = 1 \text{ mA/V}^2$, and $\lambda = 0$. The circuit parameters are $V_{DD} = 5$ V, $R_S = 1$ k Ω , $R_D = 4$ k Ω , $R_1 = 225$ k Ω , and $R_2 = 175$ k Ω . (a) Calculate the quiescent values I_{DQ} and V_{DSQ} . (b) Determine the small-signal voltage gain for $R_L = \infty$. (c) Determine the value of R_L that will reduce the small-signal voltage gain to 75 percent of the value found in part (b).

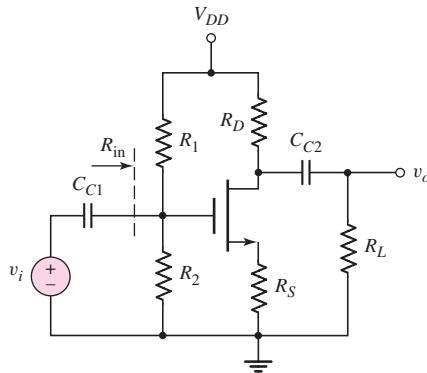


Figure P4.15

- 4.16 The parameters of the circuit shown in Figure P4.15 are $V_{DD} = 12$ V, $R_S = 0.5$ k Ω , $R_{in} = 250$ k Ω , and $R_L = 10$ k Ω . The transistor parameters are $V_{TN} = 1.2$ V, $K_n = 1.5 \text{ mA/V}^2$, and $\lambda = 0$. (a) Design the circuit such that $I_{DQ} = 2$ mA and $V_{DSQ} = 5$ V. (b) Determine the small-signal voltage gain.
- 4.17 Repeat Problem 4.15 if the source resistor is bypassed by a source capacitor C_S .
- 4.18 The ac equivalent circuit of a common-source amplifier is shown in Figure P4.18. The small-signal parameters of the transistor are $g_m = 2 \text{ mA/V}$ and $r_o = \infty$. (a) The voltage gain is found to be $A_v = V_o/V_i = -15$ with $R_S = 0$. What is the value of R_D ? (b) A source resistor R_S is inserted. Assuming the transistor parameters do not change, what is the value of R_S if the voltage gain is reduced to $A_v = -5$.
- 4.19 Consider the ac equivalent circuit shown in Figure P4.18. Assume $r_o = \infty$ for the transistor. The small-signal voltage gain is $A_v = -8$ for the case when $R_S = 1$ k Ω . (a) When R_S is shorted ($R_S = 0$), the magnitude of the voltage gain doubles. Assuming the small-signal transistor parameters do not change, what are the values of g_m and R_D ? (b) A new value of R_S is

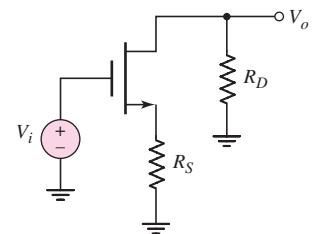


Figure P4.18

inserted into the circuit and the voltage gain becomes $A_v = -10$. Using the results of part (a), determine the value of R_S .

- 4.20 The transistor in the common-source amplifier in Figure P4.20 has parameters $V_{TN} = 0.8$ V, $k'_n = 100 \mu\text{A/V}^2$, $W/L = 50$, and $\lambda = 0.02 \text{ V}^{-1}$. The circuit parameters are $V^+ = 5$ V, $V^- = -5$ V, $I_Q = 0.5$ mA, and $R_D = 6 \text{ k}\Omega$. (a) Determine V_{GSQ} and V_{DSQ} . (b) Find the small-signal voltage gain for $R_L = \infty$. (c) Repeat part (b) for $R_L = 20 \text{ k}\Omega$. (d) Repeat part (b) for $R_L = 6 \text{ k}\Omega$.

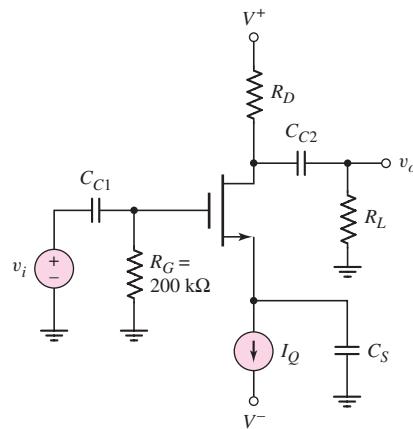


Figure P4.20

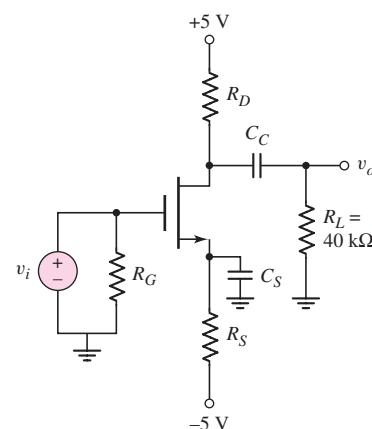


Figure P4.21

- 4.21 The parameters of the MOSFET in the circuit shown in Figure P4.21 are $V_{TN} = 0.8$ V, $K_n = 0.85 \text{ mA/V}^2$, and $\lambda = 0.02 \text{ V}^{-1}$. (a) Determine R_S and R_D such that $I_{DQ} = 0.1$ mA and $V_{DSQ} = 5.5$ V. (b) Find the small-signal transistor parameters. (c) Determine the small-signal voltage gain.

- 4.22 For the common-source amplifier in Figure P4.22, the transistor parameters are $V_{TN} = -0.8$ V, $K_n = 2 \text{ mA/V}^2$, and $\lambda = 0$. The circuit parameters are $V_{DD} = 3.3$ V and $R_L = 10 \text{ k}\Omega$. (a) Design the circuit such that $I_{DQ} = 0.5$ mA and $V_{DSQ} = 2$ V. (b) Determine the small-signal voltage gain.

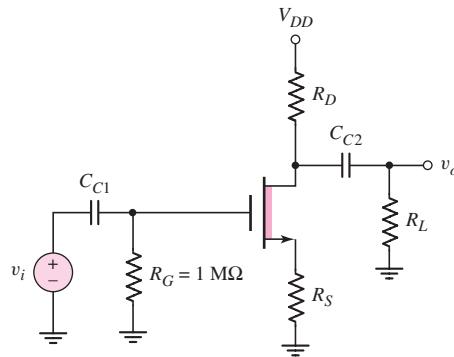


Figure P4.22

- *4.23 The transistor in the common-source circuit in Figure P4.22 has the same parameters as given in Problem 4.22. The circuit parameters are $V_{DD} = 5$ V and $R_D = R_L = 2 \text{ k}\Omega$. (a) Find R_S for $V_{DSQ} = 2.5$ V. (b) Determine the small-signal voltage gain.

- *4.24 Consider the PMOS common-source circuit in Figure P4.24 with transistor parameters $V_{TP} = -2$ V and $\lambda = 0$, and circuit parameters $R_D = R_L = 10$ k Ω . (a) Determine the values of K_p and R_S such that $V_{SDQ} = 6$ V. (b) Determine the resulting value of I_{DQ} and the small-signal voltage gain. (c) Can the values of K_p and R_S from part (a) be changed to achieve a larger voltage gain, while still meeting the requirements of part (a)?

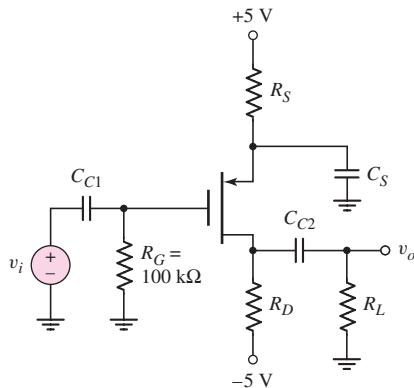


Figure P4.24

- D4.25 For the common-source circuit in Figure P4.24, the bias voltages are changed to $V^+ = 3$ V and $V^- = -3$ V. The PMOS transistor parameters are: $V_{TP} = -0.5$ V, $K_p = 0.8$ mA/V 2 , and $\lambda = 0$. The load resistor is $R_L = 2$ k Ω . (a) Design the circuit such that $I_{DQ} = 0.25$ mA and $V_{SDQ} = 1.5$ V. (b) Determine the small-signal voltage gain $A_v = v_o/v_i$.
- *D4.26 Design the common-source circuit in Figure P4.26 using an n-channel MOSFET with $\lambda = 0$. The quiescent values are to be $I_{DQ} = 6$ mA, $V_{GSQ} = 2.8$ V, and $V_{DSQ} = 10$ V. The transconductance is $g_m = 2.2$ mA/V. Let $R_L = 1$ k Ω , $A_v = -1$, and $R_{in} = 100$ k Ω . Find R_1 , R_2 , R_S , R_D , K_n , and V_{TN} .

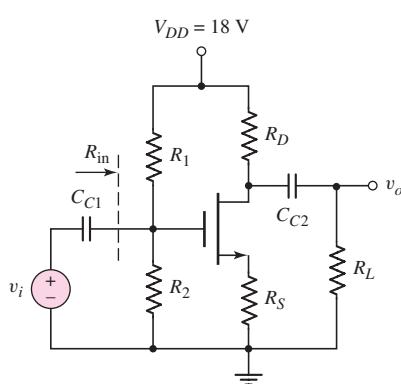


Figure P4.26

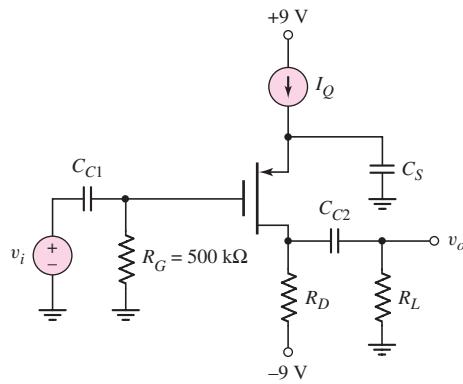


Figure P4.27

- 4.27 For the common-source amplifier shown in Figure P4.27, the transistor parameters are $V_{TP} = -1.2$ V, $K_p = 2$ mA/V 2 , and $\lambda = 0.03$ V $^{-1}$. The

drain resistor is $R_D = 4 \text{ k}\Omega$. (a) Determine I_Q such that $V_{SDQ} = 5 \text{ V}$. (b) Find the small-signal voltage gain for $R_L = \infty$. (c) Repeat part (b) for $R_L = 8 \text{ k}\Omega$.

- D4.28 For the circuit shown in Figure P4.28, the transistor parameters are: $V_{TP} = 0.8 \text{ V}$, $K_p = 0.25 \text{ mA/V}^2$, and $\lambda = 0$. (a) Design the circuit such that $I_{DQ} = 0.5 \text{ mA}$ and $V_{SDQ} = 3 \text{ V}$. (b) Determine the small-signal voltage gain $A_v = v_o/v_i$.

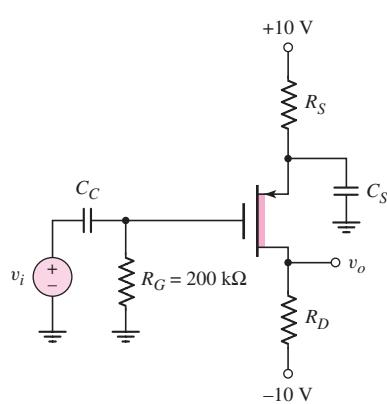


Figure P4.28

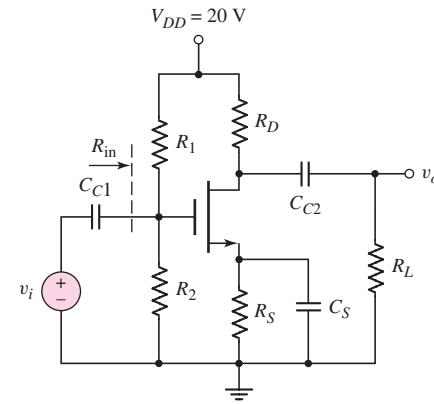


Figure P4.29

- *D4.29 Design a common-source amplifier, such as that in Figure P4.29, to achieve a small-signal voltage gain of at least $A_v = v_o/v_i = -10$ for $R_L = 20 \text{ k}\Omega$ and $R_{in} = 200 \text{ k}\Omega$. Assume the Q -point is chosen at $I_{DQ} = 1 \text{ mA}$ and $V_{DSQ} = 10 \text{ V}$. Let $V_{TN} = 2 \text{ V}$, and $\lambda = 0$.

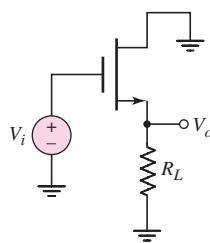


Figure P4.31

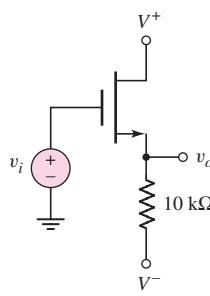


Figure P4.33

Section 4.4 The Source-Follower Amplifier

- 4.30 The small-signal parameters of an enhancement-mode MOSFET source follower are $g_m = 5 \text{ mA/V}$ and $r_o = 100 \text{ k}\Omega$. (a) Determine the no-load small-signal voltage gain and the output resistance. (b) Find the small-signal voltage gain when a load resistance $R_S = 5 \text{ k}\Omega$ is connected.
- 4.31 The open-circuit ($R_L = \infty$) voltage gain of the ac equivalent source-follower circuit shown in Figure P4.31 is $A_v = 0.98$. When R_L is set to $1 \text{ k}\Omega$, the voltage gain is reduced to $A_v = 0.49$. What are the values of g_m and r_o ?
- 4.32 Consider the source-follower circuit in Figure P4.31. The small-signal parameters of the transistor are $g_m = 2 \text{ mA/V}$ and $r_o = 25 \text{ k}\Omega$. (a) Determine the open-circuit ($R_L = \infty$) voltage gain and output resistance. (b) If $R_L = 2 \text{ k}\Omega$ and the small-signal transistor parameters remain constant, determine the voltage gain.
- 4.33 The source follower amplifier in Figure P4.33 is biased at $V^+ = 1.5 \text{ V}$ and $V^- = -1.5 \text{ V}$. The transistor parameters are $V_{TN} = 0.4 \text{ V}$, $k'_n = 100 \mu\text{A/V}^2$, $W/L = 80$, and $\lambda = 0.02 \text{ V}^{-1}$. (a) The dc value of v_O is to be zero volts. What is the current I_{DQ} and the required value of V_{GSQ} ? (b) Determine the small-signal voltage gain. (c) Find the output resistance R_o .

- 4.34 Consider the circuit in Figure P4.34. The transistor parameters are $V_{TN} = 0.6$ V, $k'_n = 100 \mu\text{A/V}^2$, and $\lambda = 0$. The circuit is to be designed such that $V_{DSQ} = 1.25$ V and such that the small-signal voltage gain is $A_v = 0.85$. (a) Find I_{DQ} . (b) Determine the width-to-length ratio of the transistor. (c) What is the required dc value of the input voltage?

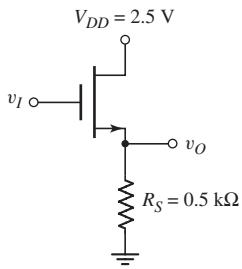


Figure P4.34

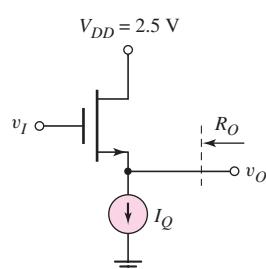


Figure P4.35

- 4.35 The quiescent power dissipation in the circuit in Figure P4.35 is to be limited to 2.5 mW. The parameters of the transistor are $V_{TN} = 0.6$ V, $k'_n = 100 \mu\text{A/V}^2$, and $\lambda = 0.02 \text{ V}^{-1}$. (a) Determine I_Q . (b) Determine W/L such that the output resistance is $R_o = 0.5 \text{ k}\Omega$. (c) Using the results of parts (a) and (b), determine the small-signal voltage gain. (d) Determine the output resistance if the transistor width-to-length ratio is $W/L = 100$.
 4.36 The parameters of the circuit in Figure P4.36 are $R_S = 4 \text{ k}\Omega$, $R_1 = 850 \text{ k}\Omega$, $R_2 = 350 \text{ k}\Omega$, and $R_L = 4 \text{ k}\Omega$. The transistor parameters are $V_{TP} = -1.2$ V, $k'_p = 40 \mu\text{A/V}^2$, $W/L = 80$, and $\lambda = 0.05 \text{ V}^{-1}$. (a) Determine I_{DQ} and V_{SDQ} . (b) Find the small-signal voltage gain $A_v = v_o/v_i$. (c) Determine the small-signal circuit transconductance gain $A_g = i_o/v_i$. (d) Find the small-signal output resistance R_o .

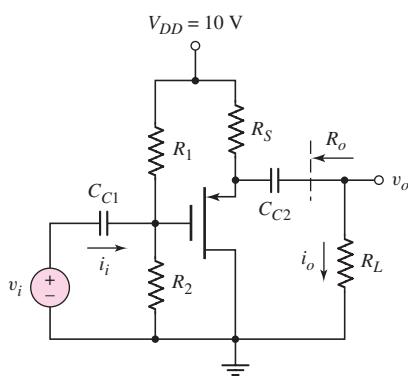


Figure P4.36

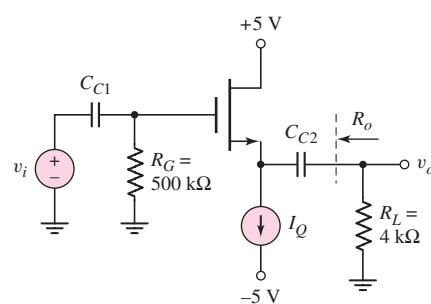


Figure P4.37

- 4.37 Consider the source follower circuit in Figure P4.37 with transistor parameters $V_{TN} = 0.8$ V, $k'_n = 100 \mu\text{A/V}^2$, $W/L = 20$, and $\lambda = 0.02 \text{ V}^{-1}$. (a) Let $I_Q = 5 \text{ mA}$. (i) Determine the small-signal voltage gain. (ii) Find the output resistance R_o . (b) Repeat part (a) for $I_Q = 2 \text{ mA}$.

- 4.38 For the source-follower circuit shown in Figure P4.37, the transistor parameters are: $V_{TN} = 1$ V, $k'_n = 60 \mu\text{A/V}^2$, and $\lambda = 0$. The small-signal voltage gain is to be $A_v = v_o/v_i = 0.95$. (a) Determine the required width-to-length ratio (W/L) for $I_Q = 4$ mA. (b) Determine the required I_Q if $(W/L) = 60$.
- *D4.39 In the source-follower circuit in Figure P4.39 with a depletion NMOS transistor, the device parameters are: $V_{TN} = -2$ V, $K_n = 5 \text{ mA/V}^2$, and $\lambda = 0.01 \text{ V}^{-1}$. Design the circuit such that $I_{DQ} = 5$ mA. Find the small-signal voltage gain $A_v = v_o/v_i$ and the output resistance R_o .

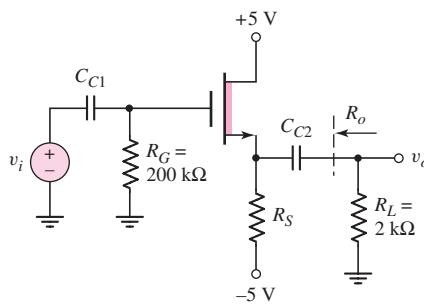


Figure P4.39

- 4.40 For the circuit in Figure P4.39, $R_S = 1 \text{ k}\Omega$ and the quiescent drain current is $I_{DQ} = 5$ mA. The transistor parameters are $V_{TN} = -2$ V, $k'_n = 100 \mu\text{A/V}^2$, and $\lambda = 0.01 \text{ V}^{-1}$. (a) Determine the transistor width-to-length ratio. (b) Using the results of part (a), find the small-signal voltage gain for $R_L = \infty$. (c) Find the small-signal output resistance R_o . (d) Using the results of part (a), find A_v for $R_L = 2 \text{ k}\Omega$.
- D4.41 For the source-follower circuit in Figure P4.39, the transistor parameters are: $V_{TN} = -2$ V, $K_n = 4 \text{ mA/V}^2$, and $\lambda = 0$. Design the circuit such that $R_o \leq 200 \Omega$. Determine the resulting small-signal voltage gain.
- 4.42 The current source in the source-follower circuit in Figure P4.42 is $I_Q = 10$ mA and the transistor parameters are $V_{TP} = -2$ V, $K_p = 5 \text{ mA/V}^2$, and $\lambda = 0.01 \text{ V}^{-1}$. (a) Find the open circuit ($R_L = \infty$) small-signal voltage gain. (b) Determine the small-signal output resistance R_o . (c) What value of R_L will reduce the small-signal voltage gain to $A_v = 0.90$?

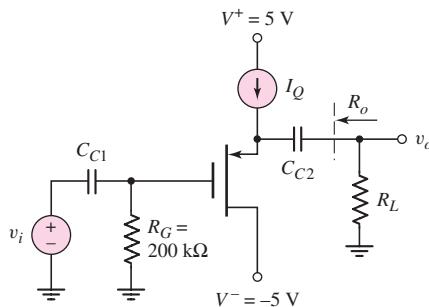


Figure P4.42

- 4.43 Consider the source-follower circuit shown in Figure P4.43. The most negative output signal voltage occurs when the transistor just cuts off. Show that this output voltage $v_o(\min)$ is given by

$$v_o(\min) = \frac{-I_{DQ}R_S}{1 + \frac{R_S}{R_L}}$$

Show that the corresponding input voltage is given by

$$v_i(\min) = -\frac{I_{DQ}}{g_m}(1 + g_m(R_S \parallel R_L))$$

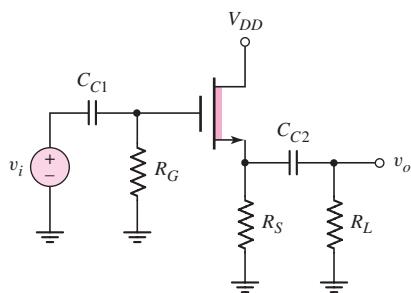


Figure P4.43

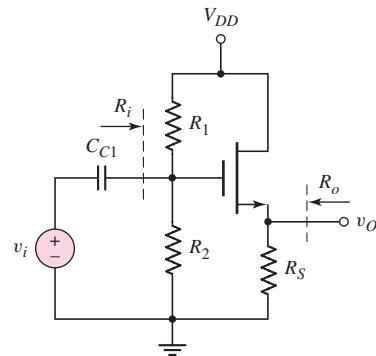


Figure P4.44

- 4.44 The transistor in the circuit in Figure P4.44 has parameters $V_{TN} = 0.4$ V, $K_n = 0.5$ mA/V², and $\lambda = 0$. The circuit parameters are $V_{DD} = 3$ V and $R_i = 300$ k Ω . (a) Design the circuit such that $I_{DQ} = 0.25$ mA and $V_{DSQ} = 1.5$ V. (b) Determine the small-signal voltage gain and the output resistance R_o .

Section 4.5 The Common-Gate Configuration

- 4.45 Figure P4.45 is the ac equivalent circuit of a common-gate amplifier. The transistor parameters are $V_{TN} = 0.4$ V, $k'_n = 100 \mu\text{A}/\text{V}^2$, and $\lambda = 0$. The quiescent drain current is $I_{DQ} = 0.25$ mA. Determine the transistor W/L ratio and the value of R_D such that the small-signal voltage gain is $A_v = V_o/V_i = 20$ and the input resistance is $R_i = 500 \Omega$.
- 4.46 The transistor in the common-gate circuit in Figure P4.46 has the same parameters that are given in Problem 4.45. The output resistance R_o is to

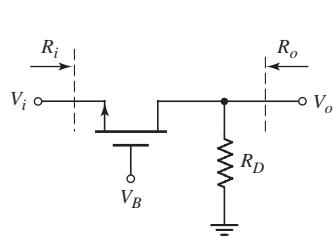


Figure P4.45

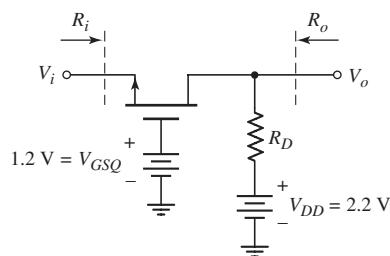


Figure P4.46

be 500Ω and the drain-to-source quiescent voltage is to be $V_{DSQ} = V_{DS}(\text{sat}) + 0.3 \text{ V}$. (a) What is the value of R_D ? (b) What is the quiescent drain current I_{DQ} ? (c) Find the input resistance R_i . (d) Determine the small-signal voltage gain $A_v = V_o/V_i$.

- 4.47 The small-signal parameters of the NMOS transistor in the ac equivalent common-gate circuit shown in Figure P4.47 are $V_{TN} = 0.4 \text{ V}$, $k'_n = 100 \mu\text{A}/\text{V}^2$, $W/L = 80$, and $\lambda = 0$. The quiescent drain current is $I_{DQ} = 0.5 \text{ mA}$. Determine the small-signal voltage gain and the input resistance.

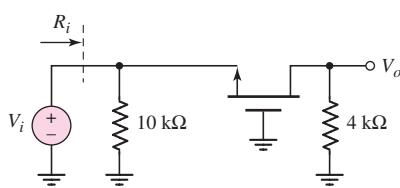


Figure P4.47

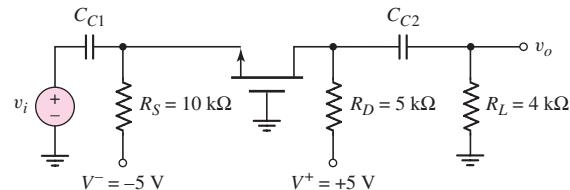


Figure P4.48

- 4.48 For the common-gate circuit in Figure P4.48, the NMOS transistor parameters are: $V_{TN} = 1 \text{ V}$, $K_n = 3 \text{ mA}/\text{V}^2$, and $\lambda = 0$. (a) Determine I_{DQ} and V_{DSQ} . (b) Calculate g_m and r_o . (c) Find the small-signal voltage gain $A_v = v_o/v_i$.
- 4.49 Consider the PMOS common-gate circuit in Figure P4.49. The transistor parameters are: $V_{TP} = -1 \text{ V}$, $K_p = 0.5 \text{ mA}/\text{V}^2$, and $\lambda = 0$. (a) Determine R_S and R_D such that $I_{DQ} = 0.75 \text{ mA}$ and $V_{SDQ} = 6 \text{ V}$. (b) Determine the input impedance R_i and the output impedance R_o . (c) Determine the load current i_o and the output voltage v_o , if $i_i = 5 \sin \omega t \mu\text{A}$.

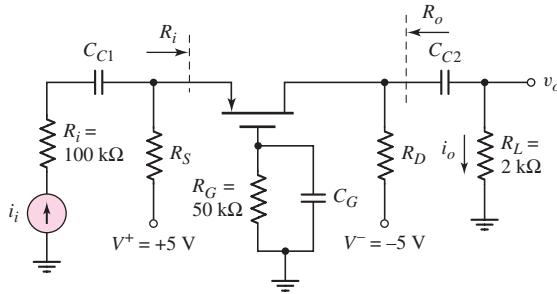


Figure P4.49

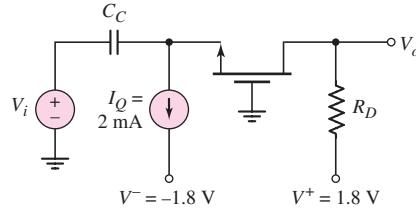


Figure P4.50

- 4.50 The transistor parameters of the NMOS device in the common-gate amplifier in Figure P4.50 are $V_{TN} = 0.4 \text{ V}$, $k'_n = 100 \mu\text{A}/\text{V}^2$, and $\lambda = 0$. (a) Find R_D such that $V_{DSQ} = V_{DS}(\text{sat}) + 0.25 \text{ V}$. (b) Determine the transistor W/L ratio such that the small-signal voltage gain is $A_v = 6$. (c) What is the value of V_{GSQ} ?
- 4.51 The parameters of the circuit shown in Figure 4.32 are $V^+ = 3.3 \text{ V}$, $V^- = -3.3 \text{ V}$, $R_G = 50 \text{ k}\Omega$, $R_L = 4 \text{ k}\Omega$, $R_{Si} = 0$, and $I_Q = 2 \text{ mA}$. The transistor parameters are $V_{TN} = 0.6 \text{ V}$, $K_n = 4 \text{ mA}/\text{V}^2$, and $\lambda = 0$. (a) Find

- R_D such that $V_{DSQ} = 3.5$ V. (b) Determine the small-signal parameters g_m and R_i . (c) Find the small-signal voltage gain A_v .
- 4.52 For the common-gate amplifier in Figure 4.35 in the text, the PMOS transistor parameters are $V_{TP} = -0.8$ V, $K_p = 2.5$ mA/V², and $\lambda = 0$. The circuit parameters are $V^+ = 3.3$ V, $V^- = -3.3$ V, $R_G = 100$ k Ω , and $R_L = 4$ k Ω . (a) Determine R_S and R_D such that $I_{DQ} = 1.2$ mA and $V_{SDQ} = 3$ V. (b) Determine the small-signal voltage gain $A_v = v_o/v_i$.

Section 4.7 Amplifiers with MOSFET Load Devices

- 4.53 Consider the NMOS amplifier with saturated load in Figure 4.39(a). The transistor parameters are $V_{TND} = V_{TNL} = 0.6$ V, $k'_n = 100 \mu\text{A}/\text{V}^2$, $\lambda = 0$, and $(W/L)_L = 1$. Let $V_{DD} = 3.3$ V. (a) Design the circuit such that the small-signal voltage gain is $|A_v| = 5$ and the Q -point is in the center of the saturation region. (b) Determine I_{DQ} and V_{DSDQ} .
- 4.54 For the NMOS amplifier with depletion load in Figure 4.43(a), the transistor parameters are $V_{TND} = 0.6$ V, $V_{TNL} = -0.8$ V, $K_{nD} = 1.2$ mA/V², $K_{nL} = 0.2$ mA/V², and $\lambda_D = \lambda_L = 0.02$ V⁻¹. Let $V_{DD} = 5$ V. (a) Determine the transistor voltages at the transition points A and B. (b) Find V_{GSDQ} and V_{DSDQ} such that the Q -point is in the middle of the saturation region. (c) Determine I_{DQ} . (d) Find the small-signal voltage gain.
- 4.55 Consider a saturated load device in which the gate and drain of an enhancement-mode MOSFET are connected together. The transistor drain current becomes zero when $V_{DS} = 0.6$ V. (a) At $V_{DS} = 1.5$ V, the drain current is 0.5 mA. Determine the small-signal resistance at this operating point. (b) What is the drain current and small-signal resistance at $V_{DS} = 3$ V?
- 4.56 The parameters of the transistors in the circuit in Figure P4.56 are $V_{TND} = -1$ V, $K_{nD} = 0.5$ mA/V² for transistor M_D , and $V_{TNL} = +1$ V, $K_{nL} = 30 \mu\text{A}/\text{V}^2$ for transistor M_L . Assume $\lambda = 0$ for both transistors. (a) Calculate the quiescent drain current I_{DQ} and the dc value of the output voltage. (b) Determine the small-signal voltage gain $A_v = v_o/v_i$ about the Q -point.
- 4.57 A source-follower circuit with a saturated load is shown in Figure P4.57. The transistor parameters are $V_{TND} = 1$ V, $K_{nD} = 1$ mA/V² for M_D , and $V_{TNL} = 1$ V, $K_{nL} = 0.1$ mA/V² for M_L . Assume $\lambda = 0$ for both transistors. Let $V_{DD} = 9$ V. (a) Determine V_{GG} such that the quiescent value of v_{DSL} is 4 V. (b) Show that the small-signal open-circuit ($R_L = \infty$) voltage gain about this Q -point is given by $A_v = 1/[1 + \sqrt{K_{nL}/K_{nD}}]$. (c) Calculate the small-signal voltage gain for $R_L = 4$ k Ω .

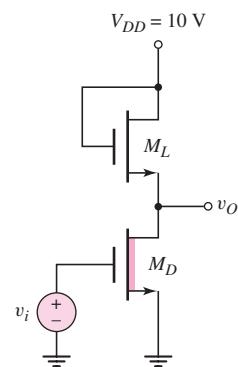


Figure P4.56

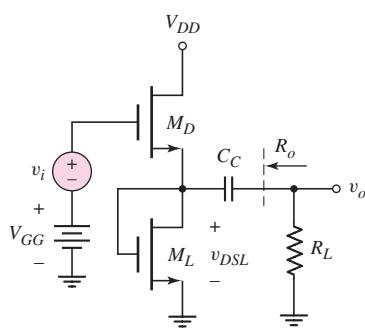


Figure P4.57

- 4.58 For the source-follower circuit with a saturated load as shown in Figure P4.57, assume the same transistor parameters as given in Problem 4.57. (a) Determine the small-signal voltage gain if $R_L = 10 \text{ k}\Omega$. (b) Determine the small-signal output resistance R_o .
- 4.59 The transistor parameters for the common-source circuit in Figure P4.59 are $V_{TND} = 0.4 \text{ V}$, $V_{TPL} = -0.4 \text{ V}$, $(W/L)_L = 50$, $\lambda_D = 0.02 \text{ V}^{-1}$, $\lambda_L = 0.04 \text{ V}^{-1}$, $k'_n = 100 \mu\text{A/V}^2$, and $k'_p = 40 \mu\text{A/V}^2$. At the Q -point, $I_{DQ} = 0.5 \text{ mA}$. (a) Determine $(W/L)_D$ such that the small-signal voltage gain is $A_v = V_o/V_i = -40$. (b) What is the required value of V_B ? (c) What is the value of V_{GSDQ} ?

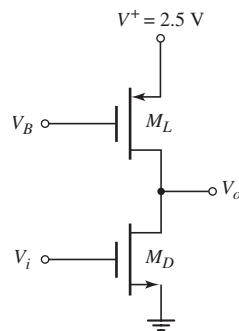


Figure P4.59

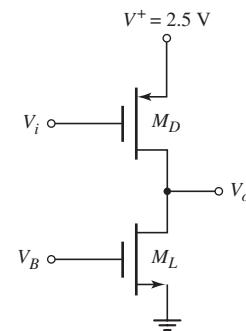


Figure P4.60

- 4.60 Consider the circuit in Figure P4.60. The transistor parameters are $V_{TPD} = -0.6 \text{ V}$, $V_{TNL} = 0.4 \text{ V}$, $k'_n = 100 \mu\text{A/V}^2$, $k'_p = 40 \mu\text{A/V}^2$, $\lambda_L = 0.02 \text{ V}^{-1}$, $\lambda_D = 0.04 \text{ V}^{-1}$, and $(W/L)_L = 10$. (a) At the Q -point, the quiescent drain current is $I_{DQ} = 0.25 \text{ mA}$. (i) Determine $(W/L)_D$ such that the small-signal voltage gain is $A_v = V_o/V_i = -25$. (ii) What is the required value of V_B ? (iii) What is the value of V_{SGDQ} ? (b) Repeat part (a) for $I_{DQ} = 0.1 \text{ mA}$.
- 4.61 The ac equivalent circuit of a CMOS common-source amplifier is shown in Figure P4.61. The transistor parameters for M_1 are $V_{TN} = 0.5 \text{ V}$, $k'_n = 85 \mu\text{A/V}^2$, $(W/L)_1 = 50$, and $\lambda = 0.05 \text{ V}^{-1}$, and for M_2 and M_3 are

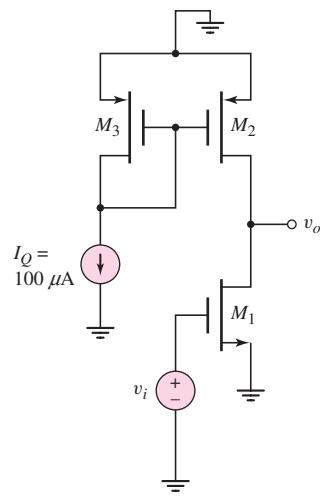


Figure P4.61

- $V_{TP} = -0.5 \text{ V}$, $k'_p = 40 \mu\text{A/V}^2$, $(W/L)_{2,3} = 50$, and $\lambda = 0.075 \text{ V}^{-1}$. Determine the small-signal voltage gain.
- 4.62 Consider the ac equivalent circuit of a CMOS common-source amplifier shown in Figure P4.62. The parameters of the NMOS and PMOS transistors are the same as given in Problem 4.61. Determine the small-signal voltage gain.
- 4.63 The parameters of the transistors in the circuit in Figure P4.63 are $V_{TND} = V_{TNL} = 0.4 \text{ V}$, $K_{nD} = 2 \text{ mA/V}^2$, $K_{nL} = 0.5 \text{ mA/V}^2$, and $\lambda_D = \lambda_L = 0$. (a) Plot V_o versus V_I over the range $0.8 \leq V_I \leq 2.5 \text{ V}$. (b) Plot I_D versus V_I over the same voltage range as part (a). (c) At $I_{DQ} = 0.20 \text{ mA}$, find the small-signal voltage gain $A_v = V_o/V_i = dV_o/dV_I$.

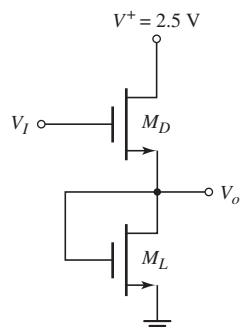


Figure P4.63

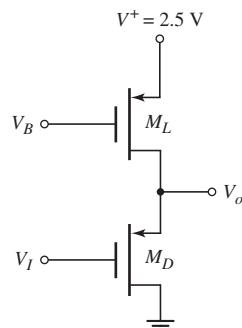


Figure P4.64

- 4.64 Consider the source-follower circuit in Figure P4.64. The transistor parameters are $V_{TP} = -0.4 \text{ V}$, $k'_p = 40 \mu\text{A/V}^2$, $(W/L)_L = 5$, $(W/L)_D = 50$, and $\lambda_D = \lambda_L = 0.025 \text{ V}^{-1}$. Assume $V_B = 1 \text{ V}$. (a) What is the maximum value of V_o such that M_L remains biased in the saturation region? (b) For M_L biased in the saturation region, determine I_D . (c) Using the results of parts (a) and (b), find V_{SGD} . (d) Determine the small-signal voltage gain when the dc value of $V_I = 0.2 \text{ V}$.
- 4.65 Figure P4.65 shows a common-gate amplifier. The transistor parameters are $V_{TN} = 0.6 \text{ V}$, $V_{TP} = -0.6 \text{ V}$, $K_n = 2 \text{ mA/V}^2$, $K_p = 0.5 \text{ mA/V}^2$, and $\lambda_n = \lambda_p = 0$. (a) Find the values of V_{SGLQ} , V_{GSDQ} , and V_{DSDQ} . (b) Derive the expression for the small-signal voltage gain in terms of K_n and K_p . (c) Calculate the value of the small-signal voltage gain $A_v = V_o/V_i$.

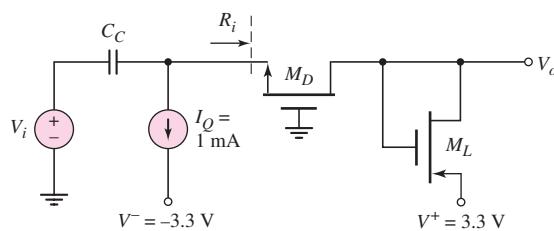


Figure P4.65

- 4.66 The ac equivalent circuit of a CMOS common-gate circuit is shown in Figure P4.66. The parameters of the NMOS and PMOS transistors are the same as given in Problem 4.61. Determine the (a) small-signal parameters of the transistors, (b) small-signal voltage gain $A_v = v_o/v_i$, (c) input resistance R_i , and (d) output resistance R_o .

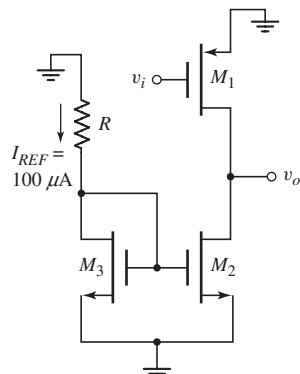


Figure P4.62

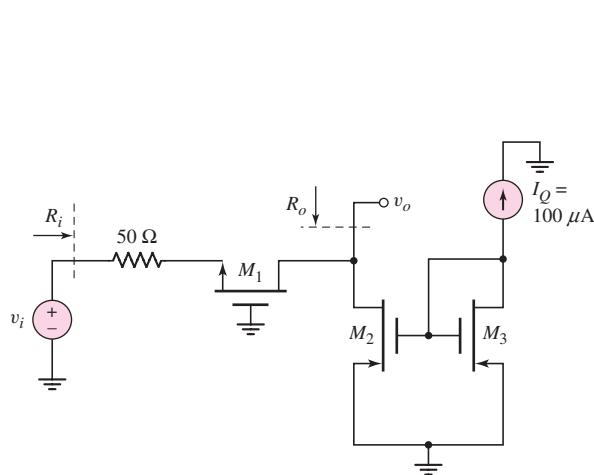


Figure P4.66

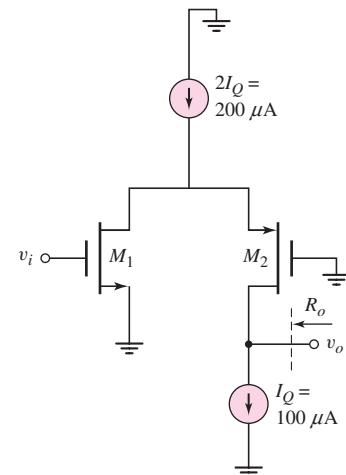


Figure P4.67

- 4.67 The circuit in Figure P4.67 is a simplified ac equivalent circuit of a folded-cascode amplifier. The transistor parameters are $|V_{TN}| = |V_{TP}| = 0.5$ V, $K_n = K_p = 2$ mA/V², and $\lambda_n = \lambda_p = 0.1$ V⁻¹. Assume the current source $2I_Q = 200$ μ A is ideal and the resistance looking into the current source $I_Q = 100$ μ A is 50 k Ω . Determine the (a) small-signal parameters of each transistor, (b) small-signal voltage gain, and (c) output resistance R_o .

Section 4.8 Multistage Amplifiers

- 4.68 The transistor parameters in the circuit in Figure P4.68 are $V_{TN1} = 0.6$ V, $V_{TP2} = -0.6$ V, $K_{n1} = 0.2$ mA/V², $K_{p2} = 1.0$ mA/V², and $\lambda_1 = \lambda_2 = 0$. The circuit parameters are $V_{DD} = 5$ V and $R_{in} = 400$ k Ω . (a) Design the circuit such that $I_{DQ1} = 0.2$ mA, $I_{DQ2} = 0.5$ mA, $V_{DSQ1} = 2$ V, and $V_{SDQ2} = 3$ V. The voltage across R_{S1} is to be 0.6 V. (b) Determine the small-signal voltage gain $A_v = v_o/v_i$.

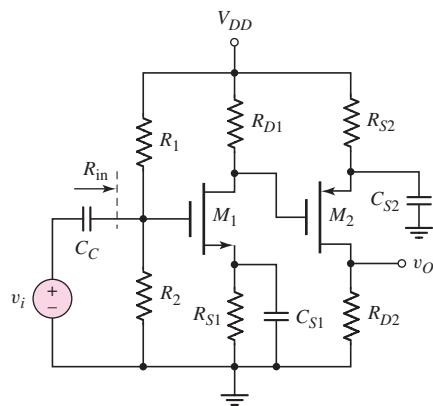


Figure P4.68

- 4.69 The transistor parameters in the circuit in Figure P4.68 are the same as those given in Problem 4.68. The circuit parameters are $V_{DD} = 3.3$ V, $R_{S1} = 1\text{ k}\Omega$, and $R_{in} = 250\text{ k}\Omega$. (a) Design the circuit such that $I_{DQ1} = 0.1$ mA, $I_{DQ2} = 0.25$ mA, $V_{DSQ1} = 1.2$ V, and $V_{SDQ2} = 1.8$ V. (b) Determine the small-signal voltage gain $A_v = v_o/v_i$.
- 4.70 Consider the circuit shown in Figure P4.70. The transistor parameters are $V_{TP1} = -0.4$ V, $V_{TN2} = 0.4$ V, $(W/L)_1 = 20$, $(W/L)_2 = 80$, $k'_p = 40\text{ }\mu\text{A/V}^2$, $k'_n = 100\text{ }\mu\text{A/V}^2$, and $\lambda_1 = \lambda_2 = 0$. Let $R_{in} = 200\text{ k}\Omega$. (a) Design the circuit such that $I_{DQ1} = 0.1$ mA, $I_{DQ2} = 0.3$ mA, $V_{SDQ1} = 1.0$ V, and $V_{DSQ2} = 2.0$ V. The voltage across R_{S1} is to be 0.6 V. (b) Determine the small-signal voltage gain $A_v = v_o/v_i$. (c) Find the small-signal output resistance R_o .

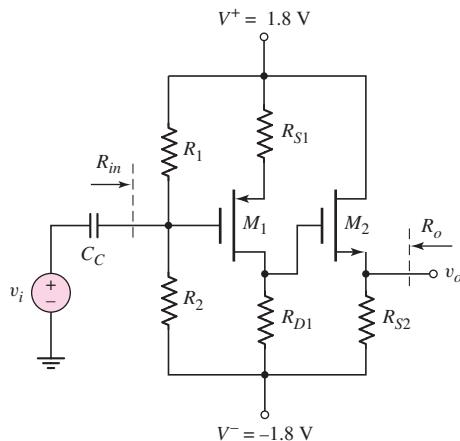


Figure P4.70

- 4.71 For the circuit in Figure P4.71, the transistor parameters are: $K_{n1} = K_{n2} = 4\text{ mA/V}^2$, $V_{TN1} = V_{TN2} = 2$ V, and $\lambda_1 = \lambda_2 = 0$. (a) Determine I_{DQ1} , I_{DQ2} , V_{DSQ1} , and V_{DSQ2} . (b) Determine g_{m1} and g_{m2} . (c) Determine the overall small-signal voltage gain $A_v = v_o/v_i$.

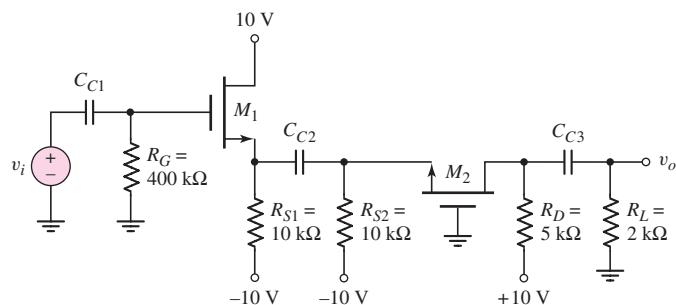


Figure P4.71

- D4.72 For the cascode circuit in Figure 4.51 in the text, the transistor parameters are: $V_{TN1} = V_{TN2} = 1$ V, $K_{n1} = K_{n2} = 2$ mA/V², and $\lambda_1 = \lambda_2 = 0$. (a) Let $R_S = 1.2$ kΩ and $R_1 + R_2 + R_3 = 500$ kΩ. Design the circuit such that $I_{DQ} = 3$ mA and $V_{DSQ1} = V_{DSQ2} = 2.5$ V. (b) Determine the small-signal voltage gain $A_v = v_o/v_i$.
- D4.73 The supply voltages to the cascode circuit in Figure 4.51 in the text are changed to $V^+ = 10$ V and $V^- = -10$ V. The transistor parameters are: $K_{n1} = K_{n2} = 4$ mA/V², $V_{TN1} = V_{TN2} = 1.5$ V, and $\lambda_1 = \lambda_2 = 0$. (a) Let $R_S = 2$ kΩ, and assume the current in the bias resistors is 0.1 mA. Design the circuit such that $I_{DQ} = 5$ mA and $V_{DSQ1} = V_{DSQ2} = 3.5$ V. (b) Determine the resulting small-signal voltage gain.

Section 4.9 Basic JFET Amplifiers

- 4.74 Consider the JFET amplifier in Figure 4.53 with transistor parameters $I_{DSS} = 6$ mA, $V_P = -3$ V, and $\lambda = 0.01$ V⁻¹. Let $V_{DD} = 10$ V. (a) Determine R_D and V_{GS} such that $I_{DQ} = 4$ mA and $V_{DSQ} = 6$ V. (b) Determine g_m and r_o at the Q -point. (c) Determine the small-signal voltage gain $A_v = v_o/v_i$ where v_o is the time-varying portion of the output voltage v_O .
- 4.75 For the JFET amplifier in Figure P4.75, the transistor parameters are: $I_{DSS} = 2$ mA, $V_P = -2$ V, and $\lambda = 0$. Determine g_m , $A_v = v_o/v_i$, and $A_i = i_o/i_i$.

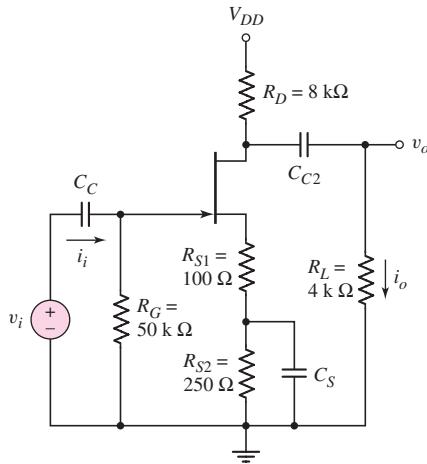


Figure P4.75

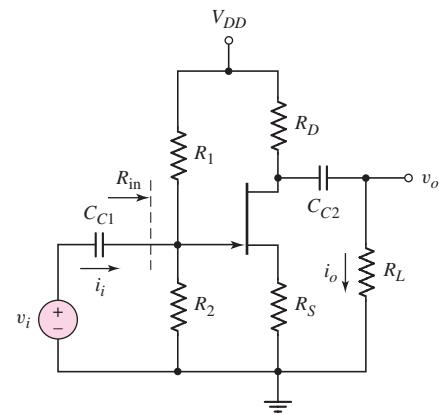


Figure P4.76

- D4.76 The parameters of the transistor in the JFET common-source amplifier shown in Figure P4.76 are: $I_{DSS} = 8$ mA, $V_P = -4.2$ V, and $\lambda = 0$. Let $V_{DD} = 20$ V and $R_L = 16$ kΩ. Design the circuit such that $V_S = 2$ V, $R_1 + R_2 = 100$ kΩ, and the Q -point is at $I_{DQ} = I_{DSS}/2$ and $V_{DSQ} = V_{DD}/2$.

- *D4.77 Consider the source-follower JFET amplifier in Figure P4.77 with transistor parameters $I_{DSS} = 10$ mA, $V_P = -5$ V, and $\lambda = 0.01$ V⁻¹. Let $V_{DD} = 12$ V and $R_L = 0.5$ kΩ. (a) Design the circuit such that $R_{in} = 100$ kΩ, and the Q -point is at $I_{DQ} = I_{DSS}/2$ and $V_{DSQ} = V_{DD}/2$. (b) Determine the resulting small-signal voltage gain $A_v = v_o/v_i$ and the output resistance R_o .

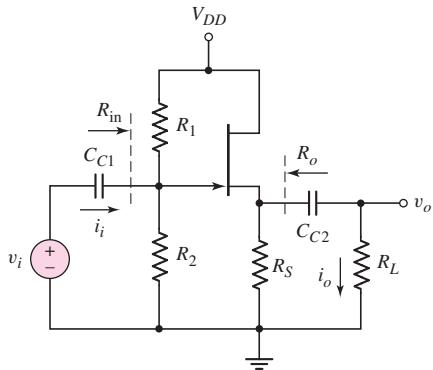


Figure P4.77

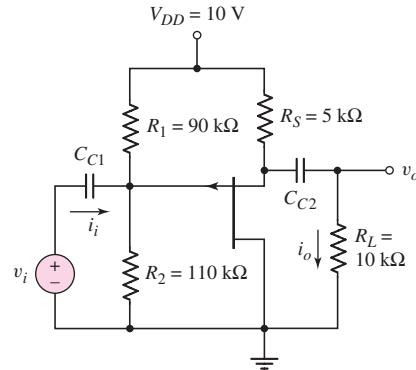


Figure P4.78

- 4.78 For the p-channel JFET source-follower circuit in Figure P4.78, the transistor parameters are: $I_{DSS} = 2 \text{ mA}$, $V_P = +1.75 \text{ V}$, and $\lambda = 0$. (a) Determine I_{DQ} and V_{SDQ} . (b) Determine the small-signal gains $A_v = v_o/v_i$ and $A_i = i_o/i_i$. (c) Determine the maximum symmetrical swing in the output voltage.
- D4.79 The p-channel JFET common-source amplifier in Figure P4.79 has transistor parameters $I_{DSS} = 8 \text{ mA}$, $V_P = 4 \text{ V}$, and $\lambda = 0$. Design the circuit such that $I_{DQ} = 4 \text{ mA}$, $V_{SDQ} = 7.5 \text{ V}$, $A_v = v_o/v_i = -3$, and $R_1 + R_2 = 400 \text{ k}\Omega$.

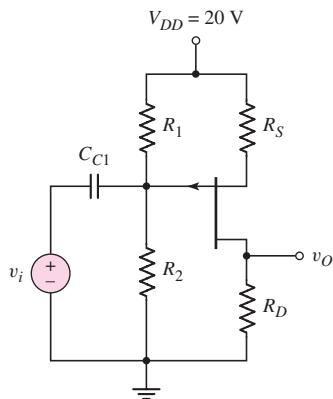


Figure P4.79

COMPUTER SIMULATION PROBLEMS

- 4.80 Consider the common-source circuit described in Example 4.5. (a) Using a computer simulation, verify the results obtained in Example 4.5. (b) Determine the change in the results when the body effect is taken into account.
- 4.81 Using a computer simulation, verify the results of Example 4.7 for the source-follower amplifier.
- 4.82 Using a computer simulation, verify the results of Example 4.10 for the common-gate amplifier.
- 4.83 Using a computer simulation, verify the results of Example 4.17 for the cascode amplifier.

 DESIGN PROBLEMS

[Note: Each design should be correlated with a computer analysis.]

- *D4.84 A discrete common-source circuit with the configuration shown in Figure 4.17 is to be designed to provide a voltage gain of 18 and a symmetrical output voltage swing. The bias voltage is $V_{DD} = 3.3$ V, the output resistance of the signal source is $500\ \Omega$, and the transistor parameters are: $V_{TN} = 0.4$ V, $k'_n = 100\ \mu\text{A}/\text{V}^2$, and $\lambda = 0.01\ \text{V}^{-1}$. Assume a quiescent drain current of $I_{DQ} = 100\ \mu\text{A}$.
- *D4.85 Consider the common-gate amplifier shown in Figure 4.35. The power supply voltages are ± 5 V, the output resistance of the signal source is $500\ \Omega$, and the input resistance of the amplifier is to be $200\ \Omega$. The transistor parameters are $k'_p = 40\ \mu\text{A}/\text{V}^2$, $V_{TP} = -0.6$ V, and $\lambda = 0$. The output load resistance is $R_L = 10\ \text{k}\Omega$. Design the circuit such that the output voltage has a peak-to-peak symmetrical swing of at least 4 V.
- *D4.86 A source-follower amplifier with the configuration shown in Figure 4.31 is to be designed. The power supplies are to be ± 12 V. The transistor parameters are $V_{TN} = 1.2$ V, $k'_n = 100\ \mu\text{A}/\text{V}^2$, and $\lambda = 0$. The load resistance is $R_L = 200\ \Omega$. Design the circuit, as well as a constant-current source, to deliver 250 mW of signal power to the load.
- *D4.87 Consider the multitransistor circuit in Figure 4.49. Assume transistor parameters of $V_{TN} = 0.6$ V, $k'_n = 100\ \mu\text{A}/\text{V}^2$, and $\lambda = 0$. Design the transistors such that the small-signal voltage gain of the first stage is $A_{v1} = -10$ and the small-signal voltage gain of the second stage is $A_{v2} = 0.9$.

Analog Electronics

Part 1 dealt with basic electronic devices and fundamental circuit configurations. Part 2 now deals with more complex analog circuits, including more sophisticated amplifiers.

Chapter 9 introduces the ideal op-amp and related circuits. The op-amp is one of the most common analog integrated circuits and can be used in a plethora of electronic applications. IC biasing techniques, which primarily use constant-current sources, are described in Chapter 10. One of the most widely used amplifier configurations is the differential amplifier, which is analyzed in Chapter 11. Chapter 12 covers the fundamentals of feedback. Feedback is used extensively in analog circuits to set or control gain values more precisely, and to alter, in a favorable way, input and output impedance values.

More complex analog integrated circuits, including circuits that form operational amplifiers, are discussed in Chapter 13. These circuits are composed of the fundamental configurations, such as the differential amplifier, constant-current biasing, active load, and output stage, that have been previously analyzed. Then Chapter 14 considers nonideal effects in operational amplifier circuits, and discusses the effects of these nonideal characteristics on op-amp circuit performance. Additional integrated circuit applications and designs are considered in Chapter 15. Such applications include active filters, oscillators, and integrated circuit power amplifiers.

Ideal Operational Amplifiers and Op-Amp Circuits¹

An operational amplifier (op-amp) is an integrated circuit that amplifies the difference between two input voltages and produces a single output. The op-amp is prevalent in analog electronics, and can be thought of as another electronic device, in much the same way as the bipolar or field-effect transistor.

The term operational amplifier comes from the original applications of the device in the early 1960s. Op-amps, in conjunction with resistors and capacitors, were used in analog computers to perform mathematical operations to solve differential and integral equations. The applications of op-amps have expanded significantly since those early days.

The main reason for postponing the discussion of op-amp circuits until now is that we can use a relatively simple transistor circuit to develop the ideal characteristics of the op-amps, instead of simply stating the ideal parameters as postulates. Once the ideal properties have been developed, the reader can then be more comfortable applying these ideal characteristics in the design of op-amp circuits. Just as we developed equivalent circuits of transistors that include dependent sources representing gain factors, we will develop a basic op-amp equivalent circuit with a dependent source that represents the device gain that can be used to determine some of the nonideal properties of op-amp circuits.

For the most part, this chapter deals with ideal op-amps. Nonideal op-amp effects are considered in Chapter 14.

PREVIEW

In this chapter, we will:

- Discuss and develop the parameters and characteristics of the ideal operational amplifier, and determine the analysis method of ideal op-amp circuits.
- Analyze and understand the characteristics of the inverting operational amplifier.
- Analyze and understand the characteristics of the summing operational amplifier.
- Analyze and understand the characteristics of the noninverting operational amplifier, including the voltage follower or buffer.
- Analyze several ideal op-amp circuits including the difference amplifier and the instrumentation amplifier.
- Discuss the operational transconductance amplifier.
- Design several ideal op-amp circuits with given design specifications.
- As an application, design an electronic thermometer in conjunction with an instrumentation amplifier that will provide the necessary amplification.

¹This chapter, through Section 9.5.5, can easily be studied as a first chapter in electronics for those who wish to cover op-amp circuits first.


9.1

THE OPERATIONAL AMPLIFIER

Objective: • Discuss and develop the parameters and characteristics of the ideal operational amplifier, and determine the analysis method of ideal op-amp circuits.

The integrated circuit operational amplifier evolved soon after development of the first bipolar integrated circuit. The μA -709 was introduced by Fairchild Semiconductor in 1965 and was one of the first widely used general-purpose op-amps. The now classic μA -741, also by Fairchild, was introduced in the late 1960s. Since then, a vast array of op-amps with improved characteristics, using both bipolar and MOS technologies, have been designed. Most op-amps are very inexpensive (less than a dollar) and are available from a wide range of suppliers.

From a signal point of view, the op-amp has two input terminals and one output terminal, as shown in the small-signal circuit symbol in Figure 9.1(a). The op-amp also requires dc power, as do all transistor circuits, so that the transistors are biased in the active region. Also, most op-amps are biased with both a positive and a negative voltage supply, as indicated in Figure 9.1(b). As before, the positive voltage is indicated by V^+ and the negative voltage by V^- .

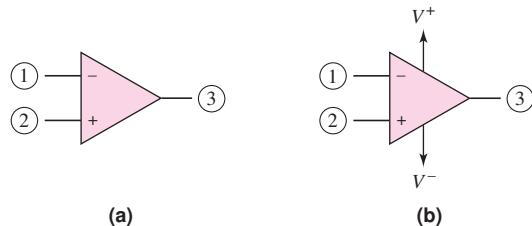


Figure 9.1 (a) Small-signal circuit symbol of the op-amp; (b) op-amp with positive and negative supply voltages

There are normally 20 to 30 transistors that make up an op-amp circuit. The typical IC op-amp has parameters that approach the ideal characteristics. For this reason, then, we can treat the op-amp as a “simple” electronic device, which means that it is quite easy to design a wide range of circuits using the IC op-amp.

In this chapter, we develop the ideal set of op-amp parameters and then consider the analysis and design of a wide variety of op-amp circuits, which will aid in our understanding of the design process of electronic circuits. We generally assume, in this chapter, that the op-amp is ideal. In the following chapters, we consider the differential amplifier, current-source biasing, and feedback, which leads to the development of the actual operational amplifier circuit in Chapter 13. Once the actual op-amp circuit is studied, then the source of nonideal characteristics can be understood. The effect of nonideal op-amp parameters is then considered in Chapter 14. Additional op-amp applications are given in Chapter 15.

9.1.1 Ideal Parameters

The ideal op-amp senses the difference between two input signals and amplifies this difference to produce an output signal. The terminal voltage is the voltage at a terminal

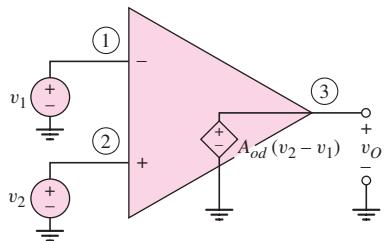


Figure 9.2 Ideal op-amp equivalent circuit

measured with respect to ground. The ideal op-amp equivalent circuit is shown in Figure 9.2.

Ideally, the input resistance R_i between terminals 1 and 2 is infinite, which means that the input current at each terminal is zero. The output terminal of the ideal op-amp acts as the output of an ideal voltage source, meaning that the small-signal output resistance R_o is zero.

The parameter A_{od} shown in the equivalent circuit is the open-loop **differential voltage gain** of the op-amp. The output is out of phase with respect to v_1 and in phase with respect to v_2 . Terminal (1) then is the **inverting input terminal**, designated by the “−” notation, and terminal (2) is the **noninverting input terminal**, designated by the “+” notation. In the ideal op-amp, the open-loop gain A_{od} is very large and approaches infinity.

Since the ideal op-amp responds only to the difference between the two input signals v_1 and v_2 , the ideal op-amp maintains a zero output signal for $v_1 = v_2$. When $v_1 = v_2 \neq 0$, there is what is called a **common-mode input signal**. For the ideal op-amp, the common-mode output signal is zero. This characteristic is referred to as **common-mode rejection**.

Because the device is biased with both positive and negative power supplies, most op-amps are direct-coupled devices (i.e., no coupling capacitors are used on the input). Therefore, the input voltages v_1 and v_2 shown in Figure 9.2 can be dc voltages, which will produce a dc output voltage v_O .

Another characteristic of the op-amp that must be considered in any design is the bandwidth or frequency response. In the ideal op-amp, this parameter is neglected. The frequency response of practical op-amps and other nonideal characteristics are discussed in Chapters 13 and 14. These nonideal parameters are considered after the actual operational amplifier circuits are analyzed in Chapter 13.

The ideal op-amp is being considered in this chapter in order to gain an appreciation of the properties and characteristics of op-amp circuits.

9.1.2 Development of the Ideal Parameters

To develop the ideal op-amp parameters, we start with the basic equivalent circuit shown in Figure 9.2.² We may note that this equivalent circuit is very similar to the MOSFET small-signal equivalent circuit. Figure 9.3(a) shows an n-channel enhancement-mode MOSFET, and Figure 9.3(b) shows the simplified low-frequency small-signal equivalent circuit. In our analysis, the transistor small-signal output resistance r_o is assumed to be infinite.

²For those readers studying this chapter as the first topic in electronics, concentrate on the analysis using the equivalent circuit and ignore any reference to the MOSFET.

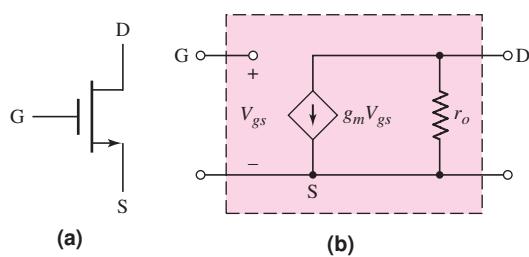


Figure 9.3 (a) n-channel enhancement-mode MOSFET and (b) small-signal equivalent circuit

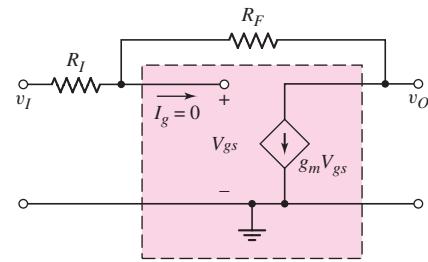


Figure 9.4 Simplified small-signal equivalent circuit of a MOSFET with input and feedback resistors

Figure 9.4 shows the equivalent circuit including two external circuit resistors \$R_I\$ and \$R_F\$. The voltage at the noninverting terminal is set equal to zero, so that the noninverting terminal is at ground potential. An input voltage \$v_I\$ is applied. Resistor \$R_F\$ is a feedback resistor that connects the output back to the input of the amplifier. This circuit is therefore called a feedback circuit. In this example, we use a single device (transistor) as the basic amplifier of the op-amp circuit.

Writing a KCL equation at the gate, or inverting, terminal, we obtain

$$\frac{v_I - V_{gs}}{R_I} = \frac{V_{gs} - v_O}{R_F} \quad (9.1(a))$$

which can be arranged as

$$\frac{v_I}{R_I} + \frac{v_O}{R_F} = V_{gs} \left(\frac{1}{R_I} + \frac{1}{R_F} \right) \quad (9.1(b))$$

Since the input impedance to the transistor is infinite, the current into the device is zero.

A KCL equation at the output node yields

$$\frac{V_{gs} - v_O}{R_F} = g_m V_{gs} \quad (9.2(a))$$

which can be solved for \$V_{gs}\$, as follows:

$$V_{gs} = \frac{v_O}{R_F} \cdot \frac{1}{\left(\frac{1}{R_F} - g_m \right)} \quad (9.2(b))$$

Substituting Equation (9.2(b)) into (9.1(b)) results in the overall voltage gain of the circuit

$$\frac{v_O}{v_I} = -\frac{R_F}{R_I} \cdot \frac{\left(1 - \frac{1}{g_m R_F} \right)}{\left(1 + \frac{1}{g_m R_F} \right)} \quad (9.3)$$

If we let the gain \$g_m\$ of the basic amplifier (i.e., the transistor) go to infinity, then the overall voltage gain becomes

$$\frac{v_O}{v_I} = -\frac{R_F}{R_I} \quad (9.4)$$

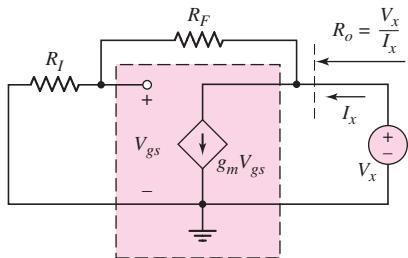


Figure 9.5 Equivalent circuit determining output resistance

Equation (9.4) shows that the overall voltage gain is the ratio of two external circuit resistors, which is one result of using an ideal op-amp. The negative sign indicates a 180 degree phase shift between the input and the output, which means that the input to the transistor corresponds to the inverting terminal of an op-amp. The voltage gain given by Equations (9.3) and (9.4) is called a **closed-loop voltage gain**, since feedback is incorporated into the circuit. Conversely, the voltage gain A_{od} is an **open-loop gain**.

Voltage V_{gs} at the input of the basic amplifier (transistor) is given by Equation (9.2(b)). Again, if we let the gain g_m go to infinity, then $V_{gs} \cong 0$; that is, the voltage at the input terminal to the basic amplifier is almost at ground potential. This terminal is said to be at **virtual ground**, which is another characteristic that we will observe in ideal op-amp circuits. The concept of virtual ground will be discussed in more detail in later sections.

The output resistance of this circuit can be determined from the equivalent circuit shown in Figure 9.5. The input signal source is set at zero. A KCL equation at the output node, written in phasor notation, is

$$I_x = g_m V_{gs} + \frac{V_x}{R_I + R_F} \quad (9.5)$$

Voltage V_{gs} can be written in terms of the test voltage V_x , as

$$V_{gs} = V_x \left(\frac{R_I}{R_I + R_F} \right) \quad (9.6)$$

Substituting Equation (9.6) into (9.5), we find that

$$\frac{I_x}{V_x} = \frac{1}{R_o} = \frac{1 + g_m R_I}{R_I + R_F} \quad (9.7(a))$$

or

$$R_o = \frac{R_I + R_F}{1 + g_m R_I} \quad (9.7(b))$$

If the gain g_m goes to infinity, then $R_o \rightarrow 0$. The output resistance of the circuit with negative feedback included goes to zero. This is also a property of an ideal op-amp circuit.

A simplified MOSFET model with a large gain has thus provided the properties of an ideal op-amp.

9.1.3 Analysis Method

Usually, an op-amp is not used in the open-loop configuration shown in Figure 9.2(a). Instead, feedback is added to close the loop between the output and the input. In this

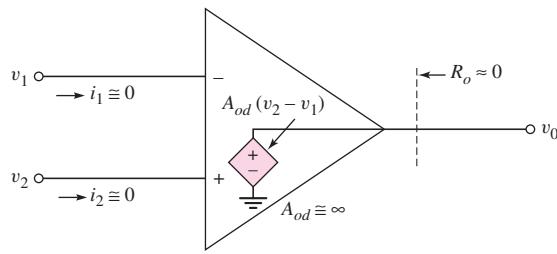


Figure 9.6 Parameters of the ideal op-amp

chapter, we will limit our discussion to **negative feedback**, in which the connection from the output goes to the inverting terminal, or terminal (1). As we will see later, this configuration produces stable circuits; positive feedback, in which the output is connected to the noninverting terminal, can be used to produce oscillators.

The ideal op-amp characteristics resulting from our negative feedback analysis are shown in Figure 9.6 and summarized below.

1. The internal differential gain A_{od} is considered to be infinite.
2. The differential input voltage ($v_2 - v_1$) is assumed to be zero. If A_{od} is very large and if the output voltage v_O is finite, then the two input voltages must be nearly equal.
3. The effective input resistance to the op-amp is assumed to be infinite, so the two input currents, i_1 and i_2 , are essentially zero.
4. The output resistance R_o is assumed to be zero in the ideal case, so the output voltage is connected directly to the dependent voltage source, and the output voltage is independent of any load connected to the output.

We use these ideal characteristics in the analysis and design of op-amp circuits.

9.1.4 Practical Specifications

In the previous discussion, we have considered the properties of an ideal op-amp. Practical op-amps are not ideal, although their characteristics approach those of an ideal op-amp. Figure 9.7(a) is a more accurate equivalent circuit of an op-amp. Also included is a load resistance connected to the output terminal. This load resistance may actually represent another op-amp circuit connected to the output terminal.

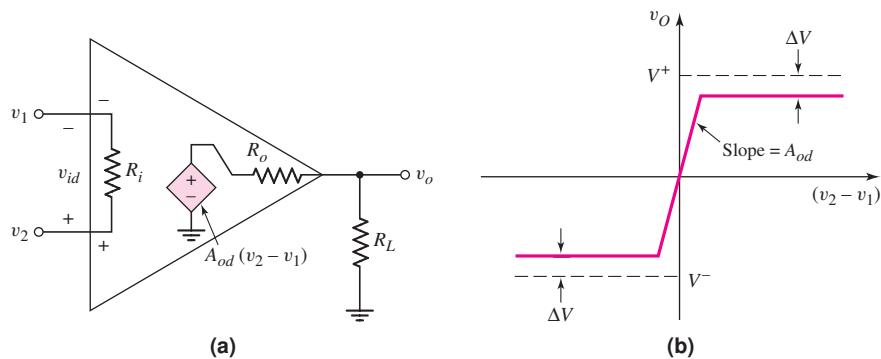


Figure 9.7 (a) Equivalent circuit of the op-amp and (b) simplified voltage transfer characteristic

Output Voltage Swing

Since the op-amp is composed of transistors biased in the active region by the dc input voltages V^+ and V^- , the output voltage is limited. When v_O approaches V^+ , it will saturate, or be limited to a value nearly equal to V^+ , since it cannot go above the positive bias voltage. Similarly, when the output voltage approaches V^- , it will saturate at a value nearly equal to V^- . The output voltage is limited to $V^- + \Delta V < v_O < V^+ - \Delta V$, as shown in Figure 9.7(b). Figure 9.7(b) is a simplified voltage transfer characteristic for the op-amp, showing the saturation effect. In older op-amp designs, such as the 741, the value of ΔV is between 1 and 2 V. We will see this property in Chapter 13. However, in newer CMOS op-amp designs, the value of ΔV may be as low as 10 mV.

Output Currents

As we can see from Figure 9.7(a), if the output voltage v_O becomes either positive or negative, a current is induced in the load resistance. If the output voltage is positive, the load current is supplied by the output of the op-amp. If the output voltage is negative, then the output of the op-amp sinks the load current. A limitation of practical op-amps is the maximum current that an op-amp can supply or sink. A typical value of the maximum current is on the order of ± 20 mA for a general-purpose op-amp.

9.1.5 PSpice Modeling

Three general purpose op-amps are included in the PSpice library. The PSpice circuit simulation uses a macromodel, which is a simplified version of the op-amp, to model the op-amp characteristics. For example, the μ A-741 op-amp has parameters $R_i = 2 \text{ M}\Omega$, $R_o = 75 \Omega$, $A_{od} = 2 \times 10^5$, and a unity-gain bandwidth of $f_{BW} = 1 \text{ MHz}$. This device is also capable of producing output voltages of $\pm 14 \text{ V}$ with dc power supply voltages of $\pm 15 \text{ V}$. We will see in several examples as to whether these nonideal parameters affect actual circuit properties.



9.2 INVERTING AMPLIFIER

Objective: • Analyze and understand the characteristics of the inverting operational amplifier.

One of the most widely used op-amp circuits is the **inverting amplifier**. Figure 9.8 shows the closed-loop configuration of this circuit. We must keep in mind that the op-amp is biased with dc voltages, although those connections are seldom explicitly shown.

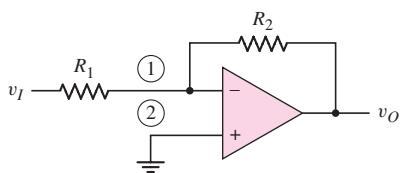


Figure 9.8 Inverting op-amp circuit

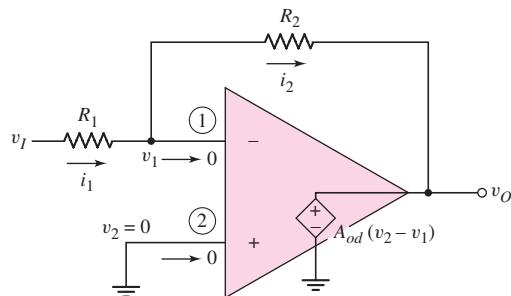


Figure 9.9 Inverting op-amp equivalent circuit

9.2.1 Basic Amplifier

We analyze the circuit in Figure 9.8 by considering the ideal equivalent circuit shown in Figure 9.9. The **closed-loop voltage gain**, or simply the voltage gain, is defined as

$$A_v = \frac{v_O}{v_I} \quad (9.8)$$

We stated that if the open-loop gain A_{od} is very large, then the two inputs v_1 and v_2 must be nearly equal. Since v_2 is at ground potential, voltage v_1 must also be approximately zero volts. We must point out, however, that having v_1 be essentially at ground potential does not imply that terminal (1) is grounded. Rather, terminal (1) is said to be at **virtual ground**; that is, it is essentially zero volts, but it does not provide a current path to ground. The virtual ground concept will be used in the analysis of ideal op-amp circuits. To repeat this important concept, with terminal 1 being at virtual ground means that terminal 1 is essentially at zero volts, but is not connected to ground potential.

From Figure 9.9, we can write

$$i_1 = \frac{v_I - v_1}{R_1} = \frac{v_I}{R_1} \quad (9.9)$$

Since the current into the op-amp is assumed to be zero, current i_1 must flow through resistor R_2 to the output terminal, which means that $i_1 = i_2$.

The output voltage is given by

$$v_O = v_1 - i_2 R_2 = 0 - \left(\frac{v_I}{R_1} \right) R_2 \quad (9.10)$$

Therefore, the closed-loop voltage gain is

$$A_v = \frac{v_O}{v_I} = -\frac{R_2}{R_1} \quad (9.11)$$

For the ideal op-amp, the closed-loop voltage gain is a function of the ratio of two resistors; it is not a function of the transistor parameters within the op-amp circuit. Again, the minus sign implies a phase reversal. If the input voltage v_I is positive, then, because v_1 is essentially at ground potential, the output voltage v_O must be negative, or below ground potential. Also note that if the output terminal is open-circuited, current i_2 must flow back into the op-amp. However, since the output impedance for the ideal case is zero, the output voltage is not a function of this current that flows back into the op-amp and is not dependent on the load.

We can also determine the input resistance seen by the voltage source v_I . Because of the virtual ground, we have, from Equation (9.9)

$$i_1 = v_I / R_1$$

The **input resistance** is then **defined** as

$$R_i = \frac{v_I}{i_1} = R_1 \quad (9.12)$$

This shows that the input resistance seen by the source is a function of R_1 only, and is a result of the “virtual ground” concept. Figure 9.10 summarizes our analysis of the inverting amplifier circuit.

Since there are no coupling capacitors in the op-amp circuit, the input and output voltages, as well as the currents in the resistors, can be dc signals. The inverting op-amp can then amplify dc voltages.

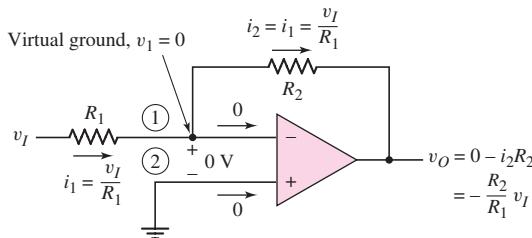


Figure 9.10 Currents and voltages in the inverting op-amp

DESIGN EXAMPLE 9.1

Objective: Design an inverting amplifier with a specified voltage gain.

Specifications: The circuit configuration to be designed is shown in Figure 9.10. Design the circuit such that the voltage gain is $A_v = -5$. Assume the op-amp is driven by an ideal sinusoidal source, $v_s = 0.1 \sin \omega t$ (V), that can supply a maximum current of $5 \mu\text{A}$. Assume that frequency ω is low so that any frequency effects can be neglected.

Design Pointer: If the sinusoidal input signal source has a nonzero output resistance, the op-amp must be redesigned to provide the specified voltage gain.

Initial Solution: The input current is given by

$$i_1 = \frac{v_I}{R_1} = \frac{v_s}{R_1}$$

If $i_1(\max) = 5 \mu\text{A}$, then we can write

$$R_1 = \frac{v_s(\max)}{i_1(\max)} = \frac{0.1}{5 \times 10^{-6}} \Rightarrow 20 \text{ k}\Omega$$

The closed-loop gain is given by

$$A_v = \frac{-R_2}{R_1} = -5$$

We then have

$$R_2 = 5R_1 = 5(20) = 100 \text{ k}\Omega$$

Trade-offs: If the signal source has a finite output resistance and the desired output voltage is $v_o = -0.5 \sin \omega t$, the circuit must be redesigned. Assume the output resistance of the source is $R_S = 1 \text{ k}\Omega$.

Redesign Solution: The output resistance of the signal source is now part of the input resistance to the op-amp. We now write

$$R_1 + R_S = \frac{v_s(\max)}{i_1(\max)} = \frac{0.1}{5 \times 10^{-6}} \Rightarrow 20 \text{ k}\Omega$$

Since $R_S = 1 \text{ k}\Omega$, we then have $R_1 = 19 \text{ k}\Omega$. The feedback resistor is then $R_2 = 5(R_1 + R_S) = 5(19 + 1) = 100 \text{ k}\Omega$.

Comment: The output resistance of the signal source must be included in the design of the op-amp to provide a specified voltage gain.

Computer Verification: Figure 9.11(a) shows the PSpice circuit schematic with the source resistance of $1 \text{ k}\Omega$ and an input resistance of $19 \text{ k}\Omega$. Figure 9.11(b) shows the

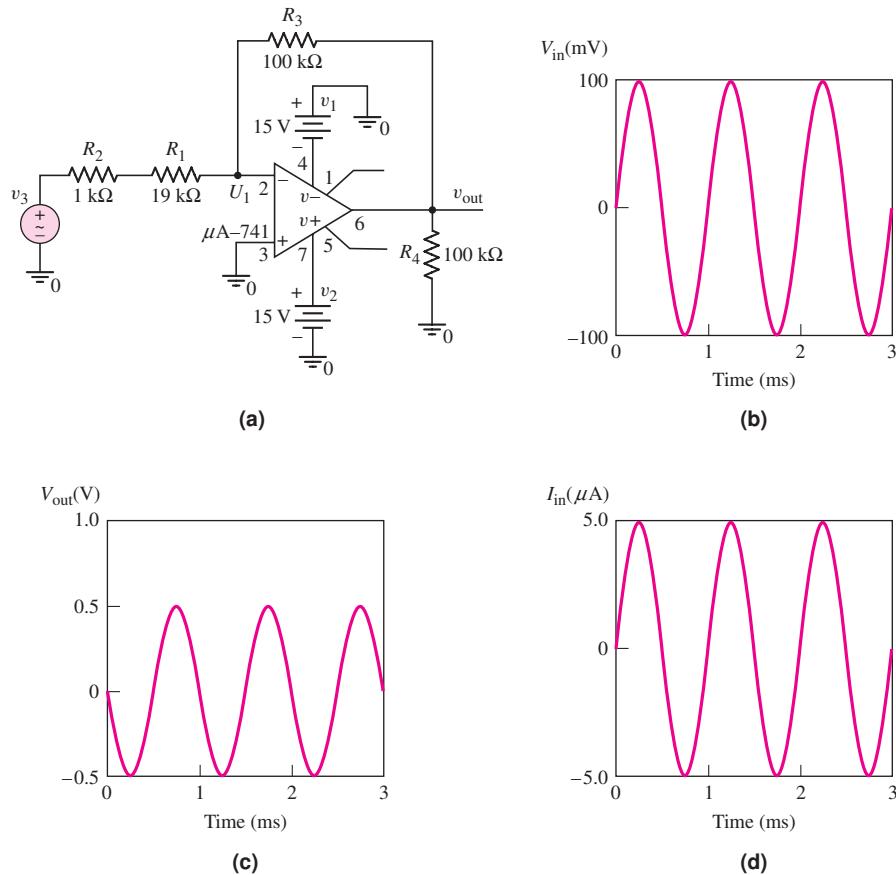


Figure 9.11 (a) PSpice circuit schematic, (b) input signal, (c) output signal, and (d) input current signal for Example 9.1

100 mV sinusoidal input signal. Figure 9.11(c) is the output signal which shows that a gain of 5 (magnitude) has been achieved and also shows that the output signal is 180 degrees out of phase with respect to the input signal. Finally, the input current is shown in Figure 9.11(d) with a maximum value of 5 μ A. The actual circuit characteristics are not influenced to any great extent by the nonideal parameters of the μ A-741 op-amp used in the circuit simulation.

EXERCISE PROBLEM

Ex 9.1: Design an ideal inverting op-amp circuit such that the voltage gain is $A_v = -25$. The maximum current in any resistor is to be limited to 10μ A with the input voltage in the range $-25 \leq v_I \leq +25$ mV. (a) What are the values of R_1 and R_2 ? (b) What is the range of output voltage v_O ? (Ans. (a) $R_1 = 2.5$ k Ω , $R_2 = 62.5$ k Ω ; (b) $-0.625 \leq v_O \leq 0.625$ V)

Problem-Solving Technique: Ideal Op-Amp Circuits

1. If the noninverting terminal of the op-amp is at ground potential, then the inverting terminal is at virtual ground. Sum currents at this node, assuming zero current enters the op-amp itself.
2. If the noninverting terminal of the op-amp is not at ground potential, then the inverting terminal voltage is equal to that at the noninverting terminal. Sum currents at the inverting terminal node, assuming zero current enters the op-amp itself.
3. For the ideal op-amp circuit, the output voltage is determined from either step 1 or step 2 above and is independent of any load connected to the output terminal.

9.2.2 Amplifier with a T-Network

Assume that an inverting amplifier is to be designed having a closed-loop voltage gain of $A_v = -100$ and an input resistance of $R_i = R_1 = 50$ k Ω . The feedback resistor R_2 would then have to be 5 M Ω . However this resistance value is too large for most practical circuits.

Consider the op-amp circuit shown in Figure 9.12 with a T-network in the feedback loop. The analysis of this circuit is similar to that of the inverting op-amp circuit of Figure 9.10. At the input, we have

$$i_1 = \frac{v_I}{R_1} = i_2 \quad (9.13)$$

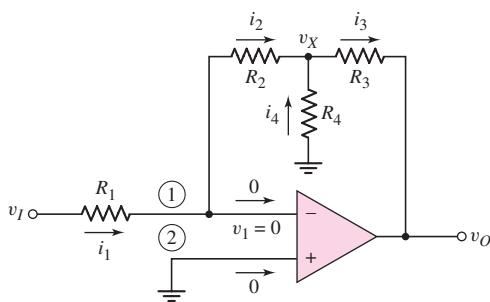


Figure 9.12 Inverting op-amp with T-network

We can also write that

$$v_X = 0 - i_2 R_2 = -v_I \left(\frac{R_2}{R_1} \right) \quad (9.14)$$

If we sum the currents at the node v_X , we have

$$i_2 + i_4 = i_3$$

which can be written

$$-\frac{v_X}{R_2} - \frac{v_X}{R_4} = \frac{v_X - v_O}{R_3} \quad (9.15)$$

or

$$v_X \left(\frac{1}{R_2} + \frac{1}{R_4} + \frac{1}{R_3} \right) = \frac{v_O}{R_3} \quad (9.16)$$

Substituting the expression for v_X from Equation (9.14), we obtain

$$-v_I \left(\frac{R_2}{R_1} \right) \left(\frac{1}{R_2} + \frac{1}{R_4} + \frac{1}{R_3} \right) = \frac{v_O}{R_3} \quad (9.17)$$

The closed-loop voltage gain is therefore

$$A_v = \frac{v_O}{v_I} = -\frac{R_2}{R_1} \left(1 + \frac{R_3}{R_4} + \frac{R_3}{R_2} \right) \quad (9.18)$$

The advantage of using a T-network is demonstrated in the following example.

DESIGN EXAMPLE 9.2

Objective: An op-amp with a T-network is to be designed as a microphone preamplifier.

Specifications: The circuit configuration to be designed is shown in Figure 9.12. The maximum microphone output voltage is 12 mV (rms) and the microphone has an output resistance of 1 kΩ. The op-amp circuit is to be designed such that the maximum output voltage is 1.2 V (rms). The input amplifier resistance should be fairly large, but all resistance values should be less than 500 kΩ.

Choices: The final design should use standard resistor values. In addition, standard resistors with tolerances of ±2 percent are to be considered.

Solution: We need a voltage gain of

$$|A_v| = \frac{1.2}{0.012} = 100$$

Equation (9.18) can be written in the form

$$A_v = -\frac{R_2}{R_1} \left(1 + \frac{R_3}{R_4} \right) - \frac{R_3}{R_1}$$

If, for example, we arbitrarily choose $\frac{R_2}{R_1} = \frac{R_3}{R_1} = 8$, then

$$-100 = -8 \left(1 + \frac{R_3}{R_4} \right) - 8$$

which yields

$$\frac{R_3}{R_4} = 10.5$$

The effective R_1 must include the R_S resistance of the microphone. If we set $R_1 = 49 \text{ k}\Omega$ so that $R_{1,\text{eff}} = 50 \text{ k}\Omega$, then

$$R_2 = R_3 = 400 \text{ k}\Omega$$

and

$$R_4 = 38.1 \text{ k}\Omega$$

Design Pointer: If we need to use standard resistance values in our design, then, using Appendix C, we can choose $R_1 = 51 \text{ k}\Omega$ so that $R_{1,\text{eff}} = 52 \text{ k}\Omega$, and we can choose $R_2 = R_3 = 390 \text{ k}\Omega$. Then, using Equation (9.18), we have

$$A_v = -100 = \frac{-R_2}{R_{1,\text{eff}}} \left(1 + \frac{R_3}{R_4} \right) - \frac{R_3}{R_{1,\text{eff}}} = \frac{-390}{52} \left(1 + \frac{390}{R_4} \right) - \frac{390}{52}$$

which yields $R_4 = 34.4 \text{ k}\Omega$. We may use a standard resistor of $R_4 = 33 \text{ k}\Omega$. This resistance value then produces a voltage gain of $A_v = -103.6$.

Trade-offs: If we consider ± 2 percent tolerances in the standard resistor values, the voltage gain can be written as

$$A_v = \frac{-R_2(1 \pm 0.02)}{1 \text{ k}\Omega + R_1(1 \pm 0.02)} \left[1 + \frac{R_3(1 \pm 0.02)}{R_4(1 \pm 0.02)} \right] - \frac{R_3(1 \pm 0.02)}{1 \text{ k}\Omega + R_1(1 \pm 0.02)}$$

or

$$A_v = \frac{-390(1 \pm 0.02)}{1 + 51(1 \pm 0.02)} \left[1 + \frac{390(1 \pm 0.02)}{33(1 \pm 0.02)} \right] - \frac{390(1 \pm 0.02)}{1 + 51(1 \pm 0.02)}$$

Analyzing this equation, we find the maximum magnitude as $|A_v|_{\max} = 111.6$ or $+7.72$ percent, and the minimum magnitude as $|A_v|_{\min} = 96.3$ or -7.05 percent.

Comment: As required, all resistor values are less than $500 \text{ k}\Omega$. Also the resistance ratios in the voltage gain expression are approximately equal. As with most design problems, there is no unique solution. We must keep in mind that, because of resistor value tolerances, the actual gain of the amplifier will have a range of values.

EXERCISE PROBLEM

Ex 9.2: Design an ideal inverting op-amp circuit with a T-network that has a closed-loop voltage gain of $A_v = -75$ and an input resistance of $R = 20 \text{ k}\Omega$. The maximum resistor value is to be limited to $200 \text{ k}\Omega$. (Ans. Let $R_1 = 20 \text{ k}\Omega$ and $R_2 = R_3 = 160 \text{ k}\Omega$. Then $R_4 = 21.7 \text{ k}\Omega$.)

The amplifier with a T-network allows us to obtain a large gain using reasonably sized resistors.

9.2.3 Effect of Finite Gain

A finite open-loop gain A_{od} , also called the finite differential-mode gain, affects the closed-loop gain of an inverting amplifier. We will consider nonideal effects in

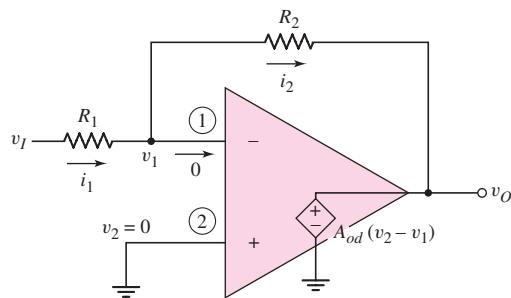


Figure 9.13 Equivalent circuit of the inverting op-amp with a finite differential-mode gain

op-amps in a later chapter; here, we will determine the magnitude of A_{od} required to approach the ideal case.

Consider the inverting op-amps shown in Figure 9.13. As before, we assume an infinite input resistance at terminals (1) and (2), which means the input currents to the op-amp are zero.

The current through R_1 can be written as

$$i_1 = \frac{v_I - v_1}{R_1} \quad (9.19)$$

and the current through R_2 is

$$i_2 = \frac{v_1 - v_O}{R_2} \quad (9.20)$$

The output voltage is now given by

$$v_O = -A_{od}v_1$$

so that the terminal (1) voltage can be written as

$$v_1 = -\frac{v_O}{A_{od}} \quad (9.21)$$

Combining Equations (9.21), (9.19), and (9.20), and setting $i_1 = i_2$, we obtain

$$i_1 = \frac{v_I + \frac{v_O}{A_{od}}}{R_1} = i_2 = \frac{-\frac{v_O}{A_{od}} - v_O}{R_2} \quad (9.22)$$

Solving for the closed-loop voltage gain, we find that

$$A_v = \frac{v_O}{v_I} = -\frac{R_2}{R_1} \left[\frac{1}{1 + \frac{1}{A_{od}} \left(1 + \frac{R_2}{R_1} \right)} \right] \quad (9.23)$$

Equation (9.23) shows that if $A_{od} \rightarrow \infty$, the ideal closed-loop voltage gain reduces to that given by Equation (9.11).

EXAMPLE 9.3

Objective: Determine the deviation from the ideal due to a finite differential gain.

Consider an inverting op-amp with $R_1 = 10 \text{ k}\Omega$ and $R_2 = 100 \text{ k}\Omega$. Determine the closed-loop gain for: $A_{od} = 10^2, 10^3, 10^4, 10^5$, and 10^6 . Calculate the percent deviation from the ideal gain.

Solution: The ideal closed-loop gain is

$$A_v = -\frac{R_2}{R_1} = -\frac{100}{10} = -10$$

If $A_{od} = 10^2$, we have, from Equation (9.23),

$$A_v = -\frac{100}{10} \cdot \frac{1}{1 + \frac{1}{10^2} \left(1 + \frac{100}{10} \right)} = \frac{-10}{(1 + 0.11)} = -9.01$$

which is a 9.9 percent deviation from the ideal. For the other differential gain values we have the following results:

A_{od}	A_v	Deviation (%)
10^2	-9.01	9.9
10^3	-9.89	1.1
10^4	-9.989	0.11
10^5	-9.999	0.01
10^6	-9.9999	0.001

Comment: For this case, the open-loop gain must be on the order of at least 10^3 in order to be within 1 percent of the ideal gain. If the ideal closed-loop gain changes, a new value of open-loop gain must be determined in order to meet the specified requirements. As we will see in Chapter 14, at low frequencies, most op-amp circuits have gains on the order of 10^5 , so achieving the required accuracy is not difficult.

EXERCISE PROBLEM

Ex 9.3: (a) An inverting op-amp circuit is to be designed using an op-amp with a finite differential voltage gain of $A_{od} = 10^4$. The closed-loop voltage gain is to be $A_v = -15.0$ and the input resistance is to be $R = 25 \text{ k}\Omega$. What is the required value of R_2 ? (b) Using the results of part (a), what is the closed-loop voltage gain if (i) $A_{od} = 10^5$ and (ii) $A_{od} = 10^3$? (Ans. (a) $R_1 = 25 \text{ k}\Omega$, $R_2 = 375.6 \text{ k}\Omega$; (b) (i) $A_v = -15.0216$, (ii) $A_v = -14.787$)

Test Your Understanding

TYU 9.1 (a) Design an ideal inverting op-amp circuit such that $A_v = -12$. Let $R_2 = 240 \text{ k}\Omega$. (b) Using the results of part (a), find i_1 when (i) $v_I = -0.15 \text{ V}$ and (ii) $v_I = +0.25 \text{ V}$. (Ans. (a) $R_1 = 20 \text{ k}\Omega$; (b) (i) $i_1 = -7.5 \mu\text{A}$, (ii) $i_1 = 12.5 \mu\text{A}$)

TYU 9.2 Consider Example 9.1. Suppose the source resistance is not a constant, but varies within the range $0.7 \text{ k}\Omega \leq R_S \leq 1.3 \text{ k}\Omega$. Using the results of Example 9.1, what is the range in (a) the voltage gain A_v and (b) the input current i_1 . (c) Is the specified maximum input current still maintained? (Ans. (a) $4.926 \leq A_v \leq 5.076$, (b) $4.926 \leq i_1 \leq 5.076 \mu\text{A}$)

TYU 9.3 Consider an inverting op-amp circuit as shown in Figure 9.13 with $R_1 = 20 \text{ k}\Omega$ and $R_2 = 200 \text{ k}\Omega$. The op-amp is ideal except the open-loop gain is $A_{od} = 10^4$. Determine (a) v_1 and v_O when $v_I = 50 \text{ mV}$, (b) v_1 and v_I when $v_O = 5 \text{ V}$, and (c) v_I and v_O when $v_I = 0.20 \text{ mV}$. (Ans. (a) $v_O = -0.49945 \text{ V}$, $v_1 = +49.945 \mu\text{V}$; (b) $v_I = -0.50055 \text{ V}$, $v_1 = -0.5 \text{ mV}$; (c) $v_O = -2.0 \text{ V}$, $v_I = 0.20022 \text{ V}$)

9.3 SUMMING AMPLIFIER

Objective: • Analyze and understand the characteristics of the summing operational amplifier.

To analyze the op-amp circuit shown in Figure 9.14(a), we will use the superposition theorem and the concept of virtual ground. Using the superposition theorem, we will determine the output voltage due to each input acting alone. We will then algebraically sum these terms to determine the total output.

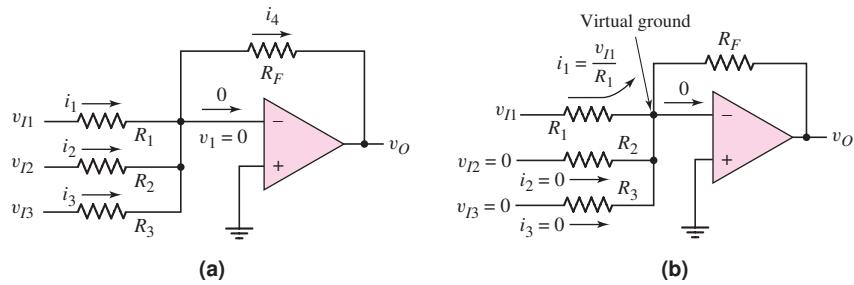


Figure 9.14 (a) Summing op-amp amplifier circuit and (b) currents and voltages in the summing amplifier

If we set \$v_{I2} = v_{I3} = 0\$, the current \$i_1\$ is

$$i_1 = \frac{v_{I1}}{R_1} \quad (9.24)$$

Since \$v_{I2} = v_{I3} = 0\$ and the inverting terminal is at virtual ground, the currents \$i_2\$ and \$i_3\$ must both be zero. Current \$i_1\$ does not flow through either \$R_2\$ or \$R_3\$, but the entire current must flow through the feedback resistor \$R_F\$, as indicated in Figure 9.14(b). The output voltage due to \$v_{I1}\$ acting alone is

$$v_O(v_{I1}) = -i_1 R_F = -\left(\frac{R_F}{R_1}\right) v_{I1} \quad (9.25)$$

Similarly, the output voltages due to \$v_{I2}\$ and \$v_{I3}\$ acting individually are

$$v_O(v_{I2}) = -i_2 R_F = -\left(\frac{R_F}{R_2}\right) v_{I2} \quad (9.26)$$

and

$$v_O(v_{I3}) = -i_3 R_F = -\left(\frac{R_F}{R_3}\right) v_{I3} \quad (9.27)$$

The total output voltage is the algebraic sum of the individual output voltages, or

$$v_O = v_O(v_{I1}) + v_O(v_{I2}) + v_O(v_{I3}) \quad (9.28)$$

which becomes

$$v_O = -\left(\frac{R_F}{R_1} v_{I1} + \frac{R_F}{R_2} v_{I2} + \frac{R_F}{R_3} v_{I3}\right) \quad (9.29)$$

The output voltage is the sum of the three input voltages, with different weighting factors. This circuit is therefore called the **inverting summing amplifier**. The number of input terminals and input resistors can be changed to add more or fewer voltages.

A special case occurs when the three input resistances are equal. When $R_1 = R_2 = R_3 \equiv R$, then

$$v_O = -\frac{R_F}{R_1}(v_{I1} + v_{I2} + v_{I3}) \quad (9.30)$$

This means that the output voltage is the sum of the input voltages, with a single amplification factor.

Discussion: Up to this point, we have seen that op-amps can be used to multiply a signal by a constant and sum a number of signals with prescribed weights. These are mathematical operations. Later in the chapter, we will see that op-amps can also be used to integrate and differentiate. These circuits are the building blocks needed to perform analog computations—hence the original name of operational amplifier. Op-amps, however, are versatile and can do much more than just perform mathematical operations, as we will continue to observe through the remainder of the chapter.

DESIGN EXAMPLE 9.4

Objective: Design a summing amplifier to produce a specified output signal.

Specifications: The output signal generated from an ideal amplifier circuit is $v_{O1} = 1.2 - 0.5 \sin \omega t$ (V). Design a summing amplifier to be connected to the amplifier circuit such that the output signal is $v_O = 2 \sin \omega t$ (V).

Choices: Standard precision resistors with tolerances of ± 1 percent are to be used in the final design. Assume an ideal op-amp is available.

Solution: In this case, we need only two inputs to the summing amplifier, as shown in Figure 9.14. One input to the summing amplifier is the output of the ideal amplifier circuit and the second input should be a dc voltage to cancel the +1.2 V signal from the amplifier circuit. If the voltage gains of each input to the summing amplifier are equal, then an input of -1.2 V at the second input will cancel the +1.2 V from the amplifier circuit.

For a -0.5 V sinusoidal input signal and a desired 2 V sinusoidal output signal, the summing amplifier gain must be

$$A_v = \frac{-R_F}{R_1} = \frac{2}{-0.5} = -4$$

If we choose the input resistances to be $R_1 = R_2 = 30 \text{ k}\Omega$, then the feedback resistance must be $R_F = 120 \text{ k}\Omega$.

Trade-offs: From Appendix C, we can choose precision resistor values of $R_F = 124 \text{ k}\Omega$ and $R_1 = R_2 = 30.9 \text{ k}\Omega$. The ratio of the ideal resistors is 4.013. Considering the ± 1 percent tolerance values, the output of the summing amplifier is

$$v_O = \frac{-R_F(1 \pm 0.01)}{R_1(1 \pm 0.01)} \cdot (1.2 - 0.5 \sin \omega t) - \frac{R_F(1 \pm 0.01)}{R_2(1 \pm 0.01)} \cdot (-1.2)$$

The dc output voltage is in the range $-0.1926 \leq v_O(\text{dc}) \leq 0.1926$ V and the peak ac output voltage is in the range $1.967 \leq v_O(\text{ac}) \leq 2.047$ V.

Comment: In this example, we have used a summing amplifier to amplify a time-varying signal and eliminate a dc voltage (ideally).

EXERCISE PROBLEM

Ex 9.4: (a) Design an inverting summing amplifier that will produce an output voltage of $v_O = -3(v_{I1} + 2v_{I2} + 0.3v_{I3} + 4v_{I4})$. The maximum resistance is to be limited to $400 \text{ k}\Omega$. (b) Using the results of part (a), determine v_O for (i) $v_{I1} = 0.1 \text{ V}$, $v_{I2} = -0.2 \text{ V}$, $v_{I3} = -1 \text{ V}$, $v_{I4} = 0.05 \text{ V}$; and for (ii) $v_{I1} = -0.2 \text{ V}$, $v_{I2} = 0.3 \text{ V}$, $v_{I3} = 1.5 \text{ V}$, $v_{I4} = -0.1 \text{ V}$. (Ans. (a) Let $R_3 = 400 \text{ k}\Omega$, $R_F = 360 \text{ k}\Omega$, $R_1 = 120 \text{ k}\Omega$, $R_2 = 60 \text{ k}\Omega$, $R_4 = 30 \text{ k}\Omega$; (b) (i) $v_O = +1.2 \text{ V}$, (ii) $v_O = -1.35 \text{ V}$)

Test Your Understanding

TYU 9.4 Consider an ideal summing amplifier as shown in Figure 9.14(a) with $R_1 = 20 \text{ k}\Omega$, $R_2 = 40 \text{ k}\Omega$, $R_3 = 50 \text{ k}\Omega$, and $R_F = 200 \text{ k}\Omega$. Determine the output voltage v_O for (a) $v_{I1} = -0.25 \text{ mV}$, $v_{I2} = +0.30 \text{ mV}$, $v_{I3} = -0.50 \text{ mV}$; and (b) $v_{I1} = +10 \text{ mV}$, $v_{I2} = -40 \text{ mV}$, $v_{I3} = +25 \text{ mV}$. (Ans. (a) $v_O = 3 \text{ mV}$, (b) $v_O = 0$)

TYU 9.5 Design the summing amplifier in Figure 9.14 to produce the average (magnitude) of three input voltages, i.e., $v_O = (v_{I1} + v_{I2} + v_{I3})/3$. The amplifier is to be designed such that each input signal sees the maximum possible input resistance under the condition that the maximum allowed resistance in the circuit is $1 \text{ M}\Omega$. (Ans. $R_1 = R_2 = R_3 = 1 \text{ M}\Omega$, $R_F = 333 \text{ k}\Omega$)

9.4 NONINVERTING AMPLIFIER

Objective: • Analyze and understand the characteristics of the noninverting operational amplifier, including the voltage follower or buffer.

In our previous discussions, the feedback element was connected between the output and the inverting terminal. However, a signal can be applied to the noninverting terminal while still maintaining negative feedback.

9.4.1 Basic Amplifier

Figure 9.15 shows the basic **noninverting amplifier**. The input signal v_I is applied directly to the noninverting terminal, while one side of resistor R_1 is connected to the inverting terminal and the other side is at ground.

Previously, when v_2 was at ground potential, we argued that v_1 was also essentially at ground potential, and we stated that terminal (1) was at virtual ground. The same principle applies to the circuit in Figure 9.15, with slightly different terminology. The negative feedback connection forces the terminal voltages v_1 and v_2 to be essentially equal. Such a condition is referred to as a **virtual short**. This condition exists since a change in v_2 will cause the output voltage v_O to change in such a way that v_1 is forced to track v_2 . The virtual short means that the voltage difference

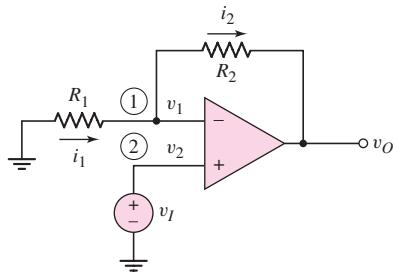


Figure 9.15 Noninverting op-amp circuit

between v_1 and v_2 is, for all practical purposes, zero. However, unlike a true short circuit, there is no current flow directly from one terminal to the other. We use the virtual short concept, i.e. $v_1 = v_2$, as an ideal op-amp characteristic and use this property in our circuit analysis.

The analysis of the noninverting amplifier is essentially the same as for the inverting amplifier. We assume that no current enters the input terminals. Since $v_1 = v_2$, then $v_1 = v_I$, and current i_1 is given by

$$i_1 = -\frac{v_1}{R_1} = -\frac{v_I}{R_1} \quad (9.31)$$

Current i_2 is given by

$$i_2 = \frac{v_1 - v_O}{R_2} = \frac{v_I - v_O}{R_2} \quad (9.32)$$

As before, $i_1 = i_2$, so that

$$-\frac{v_I}{R_1} = \frac{v_I - v_O}{R_2} \quad (9.33)$$

Solving for the closed-loop voltage gain, we find

$$A_v = \frac{v_O}{v_I} = 1 + \frac{R_2}{R_1} \quad (9.34)$$

From this equation, we see that the output is in phase with the input, as expected. Also note that the gain is always greater than unity.

The input signal v_I is connected directly to the noninverting terminal; therefore, since the input current is essentially zero, the input impedance seen by the source is very large, ideally infinite. The ideal equivalent circuit of the noninverting op-amp is shown in Figure 9.16.

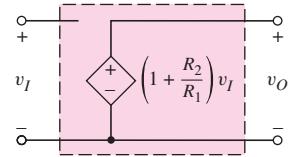


Figure 9.16 Equivalent circuit of ideal noninverting op-amp

9.4.2 Voltage Follower

An interesting property of the noninverting op-amp occurs when $R_1 = \infty$, an open circuit. The closed-loop gain then becomes

$$A_v = \frac{v_O}{v_I} = 1 \quad (9.35)$$

Since the output voltage follows the input, this op-amp circuit is called a **voltage follower**. The closed-loop gain is independent of resistor R_2 (except when $R_2 = \infty$), so we can set $R_2 = 0$ to create a short circuit.

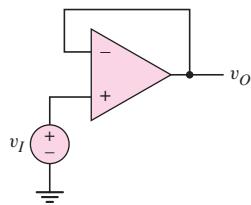


Figure 9.17 Voltage-follower op-amp

The voltage-follower op-amp circuit is shown in Figure 9.17. At first glance, it might seem that this circuit, with unity voltage gain, would be of little value. However, other terms used for the voltage follower are **impedance transformer** or **buffer**. The input impedance is essentially infinite, and the output impedance is essentially zero. If, for example, the output impedance of a signal source is large, a voltage follower inserted between the source and a load will prevent loading effects, that is, it will act as a buffer between the source and the load.

Consider the case of a voltage source with a $100\text{ k}\Omega$ output impedance driving a $1\text{ k}\Omega$ load impedance, as shown in Figure 9.18(a). This situation may occur if the source is a transducer. (We will see an example of this later in the chapter when we consider a temperature-sensitive resistor, or thermistor, in a bridge circuit.) The ratio of output voltage to input voltage is

$$\frac{v_O}{v_I} = \frac{R_L}{R_L + R_S} = \frac{1}{1 + 100} \cong 0.01$$

This equation indicates that, for this case, there is a severe loading effect, or **attenuation**, in the signal voltage.

Figure 9.18(b) shows a voltage follower inserted between the source and the load. Since the input impedance to the noninverting terminal is usually much greater than $100\text{ k}\Omega$, then $v_O \cong v_I$ and the loading effect is eliminated.

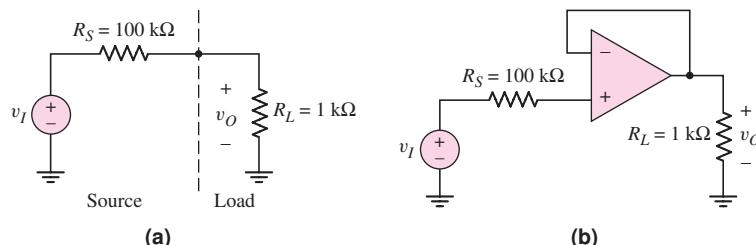


Figure 9.18 (a) Source with a $100\text{ k}\Omega$ output resistance driving a $1\text{ k}\Omega$ load and (b) source with a $100\text{ k}\Omega$ output resistance, voltage follower, and $1\text{ k}\Omega$ load

Test Your Understanding

TYU 9.6 (a) Design a noninverting amplifier such that the closed-loop gain is $A_v = 10$. The maximum resistance is to be $180\text{ k}\Omega$ and the output voltage is to be in the range $-9 \leq v_O \leq +9\text{ V}$. (b) Repeat part (a) for a closed-loop gain of $A_v = 5$. The maximum current in any resistor is to be limited to $100\text{ }\mu\text{A}$ when the output voltage is in the range $-5 \leq v_O \leq +5\text{ V}$. (Ans. (a) $R_2 = 180\text{ k}\Omega$, $R_1 = 20\text{ k}\Omega$; (b) $R_2 = 40\text{ k}\Omega$, $R_1 = 10\text{ k}\Omega$)

TYU 9.7 The noninverting op-amp in Figure 9.15 has a finite differential gain of A_{od} . Show that the closed-loop gain is

$$A_v = \frac{v_O}{v_I} = \frac{\left(1 + \frac{R_2}{R_1}\right)}{\left[1 + \frac{1}{A_{od}} \left(1 + \frac{R_2}{R_1}\right)\right]}$$

TYU 9.8 Use superposition to determine the output voltage v_O in the ideal op-amp circuit in Figure 9.19. (Ans. $v_O = 10v_{I1} + 5v_{I2}$)

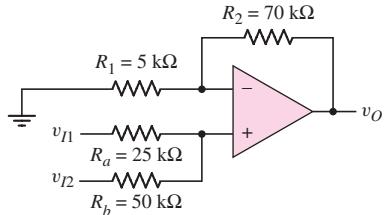


Figure 9.19 Figure for Exercise TYU 9.8



9.5 OP-AMP APPLICATIONS

Objective: • Analyze several ideal op-amp circuits including the difference amplifier and the instrumentation amplifier.

The summing amplifier is one example of special functional capabilities that can be provided by the op-amp. In this section, we will look at other examples of op-amp versatility.

9.5.1 Current-to-Voltage Converter

In some situations, the output of a device or circuit is a current. An example is the output of a photodiode or photodetector. We may need to convert this output current to an output voltage.

Consider the circuit in Figure 9.20. The input resistance R_i at the virtual ground node is

$$R_i = \frac{v_1}{i_1} \cong 0 \quad (9.36)$$

In most cases, we can assume that $R_S \gg R_i$; therefore, current i_1 is essentially equal to the signal current i_S . Then,

$$i_2 = i_1 = i_S \quad (9.37)$$

and

$$v_O = -i_2 R_F = -i_S R_F \quad (9.38)$$

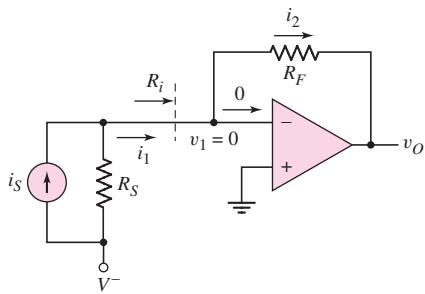


Figure 9.20 Current-to-voltage converter

The output voltage is directly proportional to the signal current, and the feedback resistance R_F is the magnitude of the ratio of the output voltage to the signal current.

9.5.2 Voltage-to-Current Converter

The complement of the current-to-voltage converter is the voltage-to-current converter. For example, we may want to drive a coil in a magnetic circuit with a given current, using a voltage source. We could use the inverting op-amp shown in Figure 9.21. For this circuit,

$$i_2 = i_1 = \frac{v_I}{R_1} \quad (9.39)$$

which means that current i_2 is directly proportional to input voltage v_I and is independent of the load impedance or resistance R_2 . However, one side of the load device might need to be at ground potential, so the circuit in Figure 9.21 would not be practical for such applications.

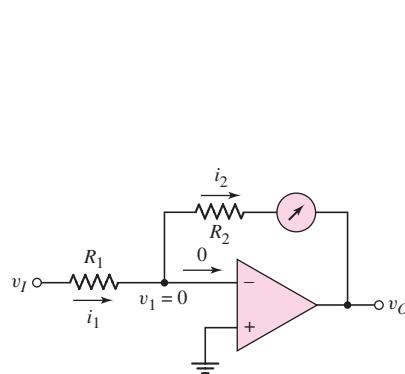


Figure 9.21 Simple voltage-to-current converter

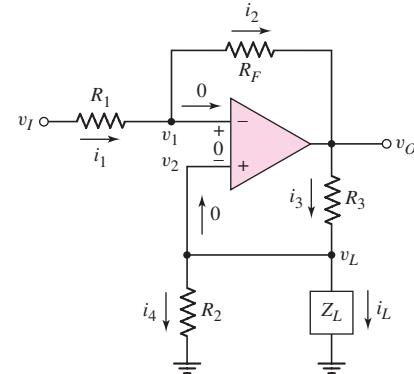


Figure 9.22 Voltage-to-current converter

Consider the circuit in Figure 9.22. In this case, one terminal of the load device, which has an impedance of Z_L , is at ground potential. The inverting terminal (1) is not at virtual ground. From the virtual short concept, $v_1 = v_2$. We also note that $v_1 = v_2 = v_L = i_L Z_L$. Equating the currents i_1 and i_2 , we have

$$\frac{v_I - i_L Z_L}{R_1} = \frac{i_L Z_L - v_O}{R_F} \quad (9.40)$$

Summing the currents at the noninverting terminal gives

$$\frac{v_O - i_L Z_L}{R_3} = i_L + \frac{i_L Z_L}{R_2} \quad (9.41)$$

Solving for $(v_O - i_L Z_L)$ from Equation (9.40) and substituting into Equation (9.41) produces

$$\frac{R_F}{R_1} \cdot \frac{(i_L Z_L - v_I)}{R_3} = i_L + \frac{i_L Z_L}{R_2} \quad (9.42)$$

Combining terms in i_L , we obtain

$$i_L \left(\frac{R_F Z_L}{R_1 R_3} - 1 - \frac{Z_L}{R_2} \right) = v_I \left(\frac{R_F}{R_1 R_3} \right) \quad (9.43)$$

In order to make i_L independent of Z_L , we can design the circuit such that the coefficient of Z_L is zero, or

$$\frac{R_F}{R_1 R_3} = \frac{1}{R_2} \quad (9.44)$$

Equation (9.43) then reduces to

$$i_L = -v_I \left(\frac{R_F}{R_1 R_3} \right) = \frac{-v_I}{R_2} \quad (9.45)$$

which means that the load current is proportional to the input voltage and is independent of the load impedance Z_L , as long as the output voltage remains between allowed limits.

We may note that the input resistance seen by the source v_I is finite, and is actually a function of the load impedance Z_L . For a constant i_L , a change in Z_L produces a change in $v_L = v_2 = v_1$, which causes a change in i_1 . A voltage follower may be inserted between the voltage source v_I and the resistor R_1 to eliminate any loading effects due to a variable input resistance.

EXAMPLE 9.5

Objective: Determine a load current in a voltage-to-current converter.

Consider the circuit in Figure 9.22. Let $Z_L = 100 \Omega$, $R_1 = 10 \text{ k}\Omega$, $R_2 = 1 \text{ k}\Omega$, $R_3 = 1 \text{ k}\Omega$, and $R_F = 10 \text{ k}\Omega$. If $v_I = -5 \text{ V}$, determine the load current i_L and the output voltage v_O .

Solution: We note first that the condition expressed by Equation (9.44) is satisfied; that is,

$$\frac{1}{R_2} = \frac{R_F}{R_1 R_3} = \frac{10}{(10)(1)} \rightarrow \frac{1}{1}$$

The load current is

$$i_L = \frac{-v_I}{R_2} = \frac{-(-5)}{1 \text{ k}\Omega} = 5 \text{ mA}$$

and the voltage across the load is

$$v_L = i_L Z_L = (5 \times 10^{-3})(100) = 0.5 \text{ V}$$

Currents i_4 and i_3 are

$$i_4 = \frac{v_L}{R_2} = \frac{0.5}{1} = 0.5 \text{ mA}$$

and

$$i_3 = i_4 + i_L = 0.5 + 5 = 5.5 \text{ mA}$$

The output voltage is then

$$v_O = i_3 R_3 + v_L = (5.5 \times 10^{-3})(10^3) + 0.5 = 6 \text{ V}$$

We could also calculate i_1 and i_2 as

$$i_1 = i_2 = -0.55 \text{ mA}$$

Comment: In this example, we implicitly assume that the op-amp is not in saturation, which means that the applied dc bias voltage must be greater than 6 V. In addition, since currents i_2 (which is negative) and i_3 must be supplied by the op-amp, we are assuming that the op-amp is capable of supplying 6.05 mA.

Computer Verification: The PSpice circuit schematic of the voltage-to-current converter is shown in Figure 9.23(a). The input voltage was varied between 0 and -10 V . Figure 9.23(b) shows the current through the 100Ω load and Figure 9.23(c)

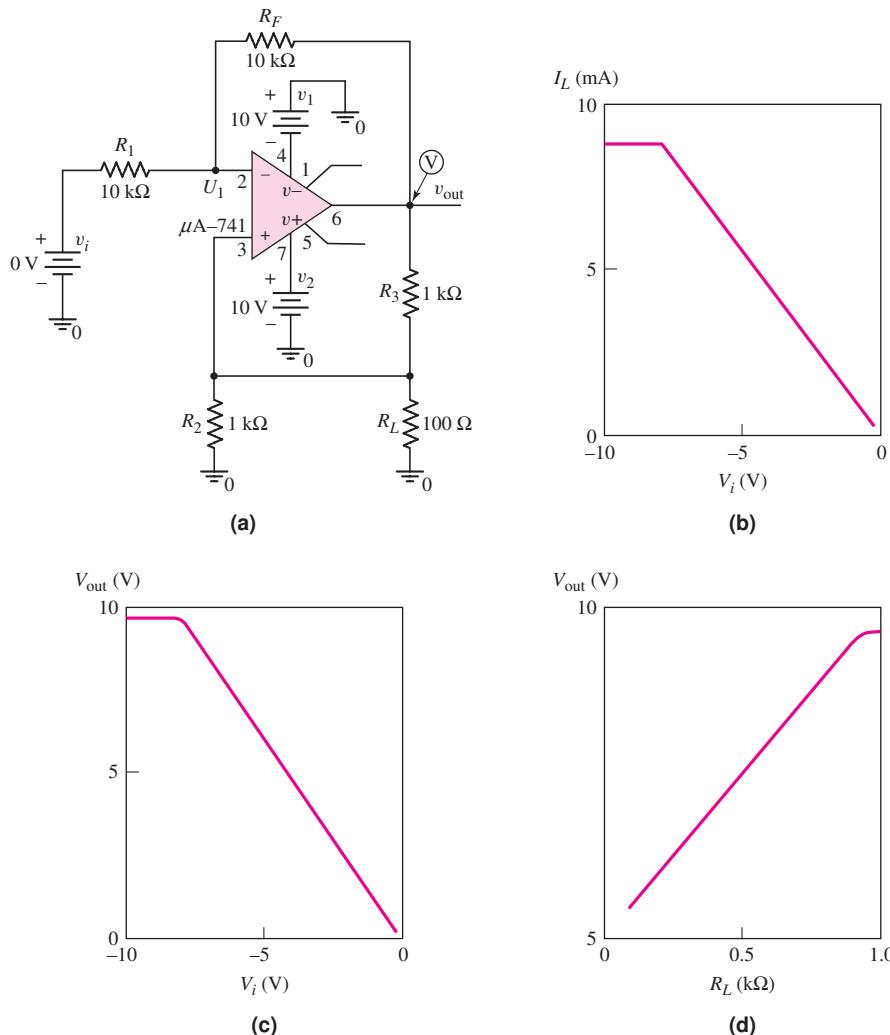


Figure 9.23 (a) PSpice circuit schematic; (b) load current and (c) op-amp output voltage versus input voltage; (d) op-amp output voltage versus load resistance for $v_1 = -5 \text{ V}$

shows the op-amp output voltage as a function of the input voltage. At approximately $v_I = -7.5$ V, the op-amp saturates, so the load current and output voltage no longer increase with input voltage. This result demonstrates that the ideal voltage-to-current conversion is valid only if the op-amp is operating in its linear region. Figure 9.23(d) shows the output voltage as a function of load resistance for an input voltage of $v_I = -5$ V. At a load resistance greater than approximately 900Ω , the op-amp saturates. The range over which the op-amp remains linear could be increased by increasing the bias to ± 15 V, for example.

EXERCISE PROBLEM

Ex 9.5: Consider the voltage-to-current converter shown in Figure 9.22. The load impedance is $Z_L = 200 \Omega$ and the input voltage is $v_I = -3$ V. Determine i_L and v_O if $R_1 = 10 \text{ k}\Omega$, $R_2 = 1.5 \text{ k}\Omega$, $R_3 = 3 \text{ k}\Omega$, and $R_F = 20 \text{ k}\Omega$. (Ans. $i_L = 2 \text{ mA}$, $v_O = 7.2 \text{ V}$)

9.5.3 Difference Amplifier

An ideal difference amplifier amplifies only the difference between two signals; it rejects any common signals to the two input terminals. For example, a microphone system amplifies an audio signal applied to one terminal of a difference amplifier, and rejects any 60 Hz noise signal or “hum” existing on both terminals. The basic op-amp also amplifies the difference between two input signals. However, we would like to make a difference amplifier, in which the output is a function of the ratio of resistors, as we had for the inverting and noninverting amplifiers.

Consider the circuit shown in Figure 9.24(a), with inputs v_{I1} and v_{I2} . To analyze the circuit, we will use superposition and the virtual short concept. Figure 9.24(b)

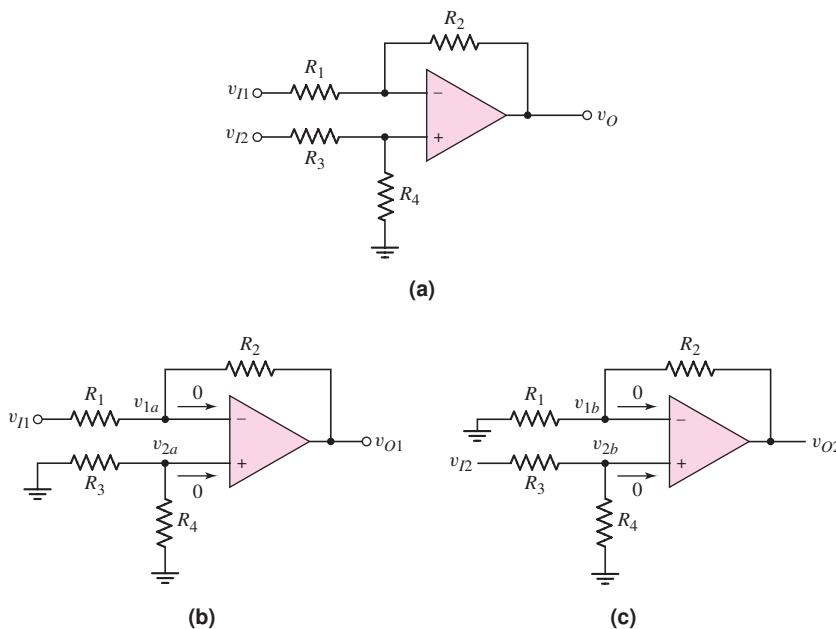


Figure 9.24 (a) Op-amp difference amplifier, (b) difference amplifier with $v_{I2} = 0$ and (c) difference amplifier with $v_{I1} = 0$

shows the circuit with input $v_{I2} = 0$. There are no currents in R_3 and R_4 ; therefore, $v_{2a} = 0$. The resulting circuit is the inverting amplifier previously considered, for which

$$v_{O1} = -\frac{R_2}{R_1} v_{I1} \quad (9.46)$$

Figure 9.24(c) shows the circuit with $v_{I1} = 0$. Since the current into the op-amp is zero, R_3 and R_4 form a voltage divider. Therefore,

$$v_{2b} = \frac{R_4}{R_3 + R_4} v_{I2} \quad (9.47)$$

From the virtual short concept, $v_{1b} = v_{2b}$ and the circuit becomes a noninverting amplifier, for which

$$v_{O2} = \left(1 + \frac{R_2}{R_1}\right) v_{1b} = \left(1 + \frac{R_2}{R_1}\right) v_{2b} \quad (9.48)$$

Substituting Equation (9.47) into (9.48), we obtain

$$v_{O2} = \left(1 + \frac{R_2}{R_1}\right) \left(\frac{R_4}{R_3 + R_4}\right) v_{I2} \quad (9.49(a))$$

which can be rearranged as follows:

$$v_{O2} = (1 + R_2/R_1) \left(\frac{R_4/R_3}{1 + R_4/R_3}\right) v_{I2} \quad (9.49(b))$$

Since the net output voltage is the sum of the individual terms, we have

$$v_O = v_{O1} + v_{O2} \quad (9.50(a))$$

or

$$v_O = \left(1 + \frac{R_2}{R_1}\right) \left(\frac{\frac{R_4}{R_3}}{1 + \frac{R_4}{R_3}}\right) v_{I2} - \left(\frac{R_2}{R_1}\right) v_{I1} \quad (9.50(b))$$

A property of the ideal difference amplifier is that the output voltage is zero when $v_{I1} = v_{I2}$. An inspection of Equation (9.50(b)) shows that this condition is met if

$$\frac{R_4}{R_3} = \frac{R_2}{R_1} \quad (9.51)$$

The output voltage is then

$$v_O = \frac{R_2}{R_1} (v_{I2} - v_{I1}) \quad (9.52)$$

which indicates that this amplifier has a differential gain of $A_d \equiv R_2/R_1$. This factor is a closed-loop differential gain, rather than the open-loop differential gain A_{od} of the op-amp itself.

As previously stated, another important characteristic of electronic circuits is the input resistance. The **differential input resistance** of the differential amplifier can be determined by using the circuit shown in Figure 9.25. In the figure, we have

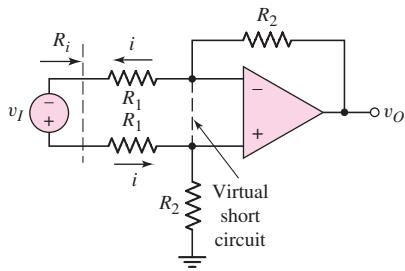


Figure 9.25 Circuit for measuring differential input resistance of op-amp difference amplifier

imposed the condition given in Equation (9.51) and have set $R_1 = R_3$ and $R_2 = R_4$. The input resistance is then defined as

$$R_i = \frac{v_I}{i} \quad (9.53)$$

Taking into account the virtual short concept, we can write a loop equation, as follows:

$$v_I = iR_1 + iR_1 = i(2R_1) \quad (9.54)$$

Therefore, the input resistance is

$$R_i = 2R_1 \quad (9.55)$$

DESIGN EXAMPLE 9.6

Objective: Design a difference amplifier with a specified gain.

Specifications: Design the difference amplifier with the configuration shown in Figure 9.24 such that the differential gain is 30. Standard valued resistors are to be used and the maximum resistor value is to be 500 kΩ.

Choices: An ideal op-amp is available.

Solution: The differential gain is given by

$$\frac{R_2}{R_1} = \frac{R_4}{R_3} = 30$$

From Appendix C, we can use standard resistors of

$$R_2 = R_4 = 390 \text{ k}\Omega \quad \text{and} \quad R_1 = R_3 = 13 \text{ k}\Omega$$

These resistor values are obviously less than 500 kΩ and will give an input resistance of $R_i = 2R_1 = 2(13) = 26 \text{ k}\Omega$.

Trade-offs: Resistor tolerances must be considered as we have in other designs. This effect is considered in end-of-chapter Problem 9.62. Resistor tolerances also affect the common-mode rejection ratio, as analyzed in the following example.

Comment: This example illustrated one disadvantage of this differential amplifier design. It cannot achieve both high gain and high input impedance without using extremely large resistor values.

EXERCISE PROBLEM

Ex 9.6: Consider the difference amplifier in Figure 9.24(a). (a) Design the circuit with $R_2 = R_4$, $R_1 = R_3$, and such that the differential voltage gain is $A_d = 50$. For input voltages in the range of -50 mV to $+50 \text{ mV}$, the maximum current in R_2 is to be limited to $50 \mu\text{A}$. (b) Using the results of part (a), what is the maximum current in R_3 ? (Ans. (a) Set $R_2 = R_4 = 100 \text{ k}\Omega$, then $R_1 = R_2 = 2 \text{ k}\Omega$; (b) $0.49 \mu\text{A}$)

In the ideal difference amplifier, the output v_O is zero when $v_{I1} = v_{I2}$. However, an inspection of Equation (9.50(b)) shows that this condition is not satisfied if $R_4/R_3 \neq R_2/R_1$. When $v_{I1} = v_{I2}$, the input is called a **common-mode input signal**. The common-mode input voltage is defined as

$$v_{cm} = (v_{I1} + v_{I2})/2 \quad (9.56)$$

The common-mode gain is then defined as

$$A_{cm} = \frac{v_O}{v_{cm}} \quad (9.57)$$

Ideally, when a common-mode signal is applied, $v_O = 0$ and $A_{cm} = 0$.

A nonzero common-mode gain may be generated in actual op-amp circuits. This is discussed in Chapter 14.

A figure of merit for a difference amplifier is the **common-mode rejection ratio (CMRR)**, which is defined as the magnitude of the ratio of differential gain to common-mode gain, or

$$\text{CMRR} = \left| \frac{A_d}{A_{cm}} \right| \quad (9.58)$$

Usually, the CMRR is expressed in decibels, as follows:

$$\text{CMRR(dB)} = 20 \log_{10} \left| \frac{A_d}{A_{cm}} \right| \quad (9.59)$$

Ideally, the common-mode rejection ratio is infinite. In an actual differential amplifier, we would like the common-mode rejection ratio to be as large as possible.

EXAMPLE 9.7

Objective: Calculate the common-mode rejection ratio of a difference amplifier.

Consider the difference amplifier shown in Figure 9.24(a). Let $R_2/R_1 = 10$ and $R_4/R_3 = 11$. Determine CMRR(dB).

Solution: From Equation (9.50(b)), we have

$$v_O = (1 + 10) \left(\frac{11}{1 + 11} \right) v_{I2} - (10)v_{I1}$$

or

$$v_O = 10.0833v_{I2} - 10v_{I1} \quad (9.60)$$

The differential-mode input voltage is defined as

$$v_d = v_{I2} - v_{I1}$$

and the common-mode input voltage is defined as

$$v_{cm} = (v_{I1} + v_{I2})/2$$

Combining these two equations produces

$$v_{I1} = v_{cm} - \frac{v_d}{2} \quad (9.61(a))$$

and

$$v_{I2} = v_{cm} + \frac{v_d}{2} \quad (9.61(b))$$

If we substitute Equations (9.61(a)) and (9.61(b)) in Equation (9.60), we obtain

$$v_o = (10.0833) \left(v_{cm} + \frac{v_d}{2} \right) - (10) \left(v_{cm} - \frac{v_d}{2} \right)$$

or

$$v_o = 10.042v_d + 0.0833v_{cm} \quad (9.62)$$

The output voltage is also

$$v_o = A_d v_d + A_{cm} v_{cm} \quad (9.63)$$

If we compare Equations (9.62) and (9.63), we see that

$$A_d = 10.042 \quad \text{and} \quad A_{cm} = 0.0833$$

Therefore, from Equation (9.59), the common-mode rejection ratio, is

$$\text{CMRR(dB)} = 20 \log_{10} \left(\frac{10.042}{0.0833} \right) = 41.6 \text{ dB}$$

Comment: For good differential amplifiers, typical CMRR values are in the range of 80–100 dB. This example shows how close the ratios R_2/R_1 and R_4/R_3 must be in order to achieve a CMRR value in that range.

Computer Verification: A PSpice analysis was performed on the differential amplifier in this example with a μ A-741 op-amp. For input voltages of $v_{I1} = -50$ mV and $v_{I2} = +50$ mV, the output voltage is $v_o = 1.0043$ V, which gives a differential voltage gain of 10.043. For input voltages of $v_{I1} = v_{I2} = 5$ V, the output voltage is $v_o = 0.4153$ V, which gives a common-mode voltage gain of $A_{cm} = 0.4153/5 = 0.0831$. The common-mode rejection ratio is then $\text{CMRR} = 10.043/0.0831 = 120.9 \Rightarrow 41.6$ dB, which agrees with the hand analysis. This result demonstrates that at this point, the nonideal characteristics of the μ A-741 op-amp do not affect these results.

EXERCISE PROBLEM

***Ex 9.7:** In the difference amplifier shown in Figure 9.24(a), $R_1 = R_3 = 10$ k Ω , $R_2 = 20$ k Ω , and $R_4 = 21$ k Ω . Determine v_o when: (a) $v_{I1} = +1$ V, $v_{I2} = -1$ V; and (b) $v_{I1} = v_{I2} = +1$ V. (c) Determine the common-mode gain. (d) Determine the CMRR(dB). (Ans. (a) $v_o = -4.032$ V, (b) $v_o = 0.0323$ V, (c) $A_{cm} = 0.0323$, (d) CMRR(db) = 35.9 dB)

9.5.4 Instrumentation Amplifier

We saw in the last section that it is difficult to obtain a high input impedance and a high gain in a difference amplifier with reasonable resistor values. One solution is to insert a voltage follower between each source and the corresponding input. However, a disadvantage of this design is that the gain of the amplifier cannot easily be changed. We would need to change two resistance values and still maintain equal ratios between R_2/R_1 and R_4/R_3 . Optimally, we would like to be able to change the gain by changing only a single resistance value. The circuit in Figure 9.26, called an instrumentation amplifier, allows this flexibility. Note that two noninverting amplifiers, A_1 and A_2 , are used as the input stage, and a difference amplifier, A_3 is the second, or amplifying, stage.

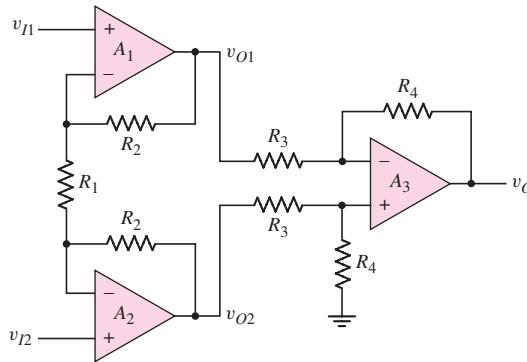


Figure 9.26 Instrumentation amplifier

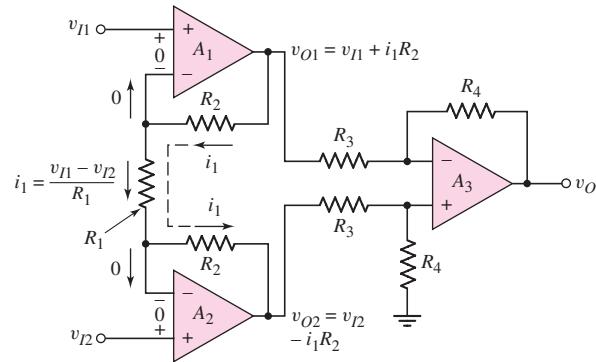


Figure 9.27 Voltages and currents in instrumentation amplifier

We begin the analysis using the virtual short concept. The voltages at the inverting terminals of the voltage followers are equal to the input voltages. The currents and voltages in the amplifier are shown in Figure 9.27. The current in resistor R_1 is then

$$i_1 = \frac{v_{I1} - v_{I2}}{R_1} \quad (9.64)$$

The current in resistors R_2 is also i_1 , as shown in the figure, and the output voltages of op-amps A_1 and A_2 are, respectively,

$$v_{O1} = v_{I1} + i_1 R_2 = \left(1 + \frac{R_2}{R_1}\right) v_{I1} - \frac{R_2}{R_1} v_{I2} \quad (9.65(a))$$

and

$$v_{O2} = v_{I2} - i_1 R_2 = \left(1 + \frac{R_2}{R_1}\right) v_{I2} - \frac{R_2}{R_1} v_{I1} \quad (9.65(b))$$

From previous results, the output of the difference amplifier is given as

$$v_O = \frac{R_4}{R_3} (v_{O2} - v_{O1}) \quad (9.66)$$

Substituting Equations (9.65(a)) and (9.65(b)) into Equation (9.66), we find the output voltage, as follows:

$$v_O = \frac{R_4}{R_3} \left(1 + \frac{2R_2}{R_1}\right) (v_{I2} - v_{I1}) \quad (9.67)$$

Since the input signal voltages are applied directly to the noninverting terminals of A_1 and A_2 , the input impedance is very large, ideally infinite, which is one desirable characteristic of the instrumentation amplifier. Also, the differential gain is a function of resistor R_1 , which can easily be varied by using a potentiometer, thus providing a variable amplifier gain with the adjustment of only one resistance.

EXAMPLE 9.8

Objective: Determine the range required for resistor R_1 , to realize a differential gain adjustable from 5 to 500.

The instrumentation amplifier circuit is shown in Figure 9.26. Assume that $R_4 = 2R_3$, so that the difference amplifier gain is 2.

Solution: Assume that resistance R_1 is a combination of a fixed resistance R_{1f} and a variable resistance R_{1v} , as shown in Figure 9.28. The fixed resistance ensures that the gain is limited to a maximum value, even if the variable resistance is set equal to zero. Assume the variable resistance is a $100\text{ k}\Omega$ potentiometer.

From Equation (9.67), the maximum differential gain is

$$500 = 2 \left(1 + \frac{2R_2}{R_{1f}} \right)$$

and the minimum differential gain is

$$5 = 2 \left(1 + \frac{2R_2}{R_{1f} + 100} \right)$$

From the maximum gain expression, we find that

$$2R_2 = 249R_{1f}$$

Substituting this R_2 value into the minimum gain expression, we have

$$1.5 = \frac{2R_2}{R_{1f} + 100} = \frac{249R_{1f}}{R_{1f} + 100}$$

The resulting value of R_{1f} is $R_{1f} = 0.606\text{ k}\Omega$, which yields $R_2 = 75.5\text{ k}\Omega$.

Comment: We can select standard resistance values that are close to the values calculated, and the range of the gain will be approximately in the desired range.

Design Pointer: An amplifier with a wide range of gain and designed with a potentiometer would normally not be used with standard integrated circuits in electronic systems. However, such a circuit might be very useful in specialized test equipment.

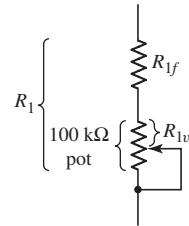


Figure 9.28 Equivalent resistance R_1 in instrumentation amplifier

EXERCISE PROBLEM

Ex 9.8: For the instrumentation amplifier in Figure 9.26, the parameters are $R_4 = 90\text{ k}\Omega$, $R_3 = 30\text{ k}\Omega$, and $R_2 = 50\text{ k}\Omega$. Resistance R_1 is a series combination of a fixed $2\text{ k}\Omega$ resistor and a $100\text{ k}\Omega$ potentiometer. (a) Determine the range of the differential voltage gain. (b) Determine the maximum current in R_1 for input voltages in the range -25 mV to $+25\text{ mV}$. (Ans. (a) $5.94 \leq A_d \leq 153$, (b) $25\text{ }\mu\text{A}$)

9.5.5 Integrator and Differentiator

In the op-amp circuits previously considered, the elements exterior to the op-amp have been resistors. Other elements can be used, with differing results. Figure 9.29 shows a generalized inverting amplifier for which the voltage transfer function has the same general form as before, that is,

$$\frac{v_O}{v_I} = -\frac{Z_2}{Z_1} \quad (9.68)$$

where Z_1 and Z_2 are generalized impedances. Two special circuits can be developed from this generalized inverting amplifier.

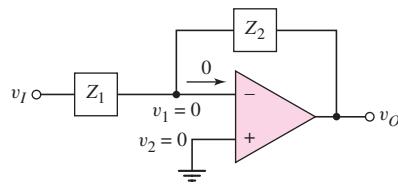


Figure 9.29 Generalized inverting amplifier

In the first, Z_1 corresponds to a resistor and Z_2 to a capacitor. The impedances are then $Z_1 = R_1$ and $Z_2 = 1/sC_2$, where s again is the complex frequency. The output voltage is

$$v_O = -\frac{Z_2}{Z_1}v_I = \frac{-1}{sR_1C_2}v_I \quad (9.69)$$

Equation (9.69) represents integration in the time domain. If V_C is the voltage across the capacitor at $t = 0$, the output voltage is

$$v_O = V_C - \frac{1}{R_1C_2} \int_0^t v_I(t') dt' \quad (9.70)$$

where t' is the variable of integration. Figure 9.30 summarizes these results.

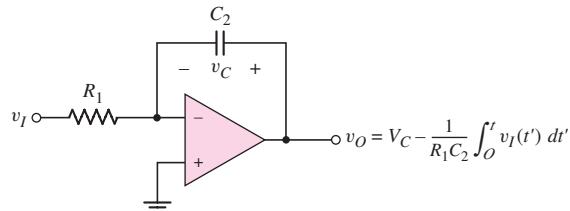


Figure 9.30 Op-amp integrator

Equation (9.70) is the output response of the integrator circuit, shown in Figure 9.30, for any input voltage v_I . Note that if $v_I(t)$ is a finite step function, output v_O will be a linear function of time. The output v_O will be a ramp function and will eventually saturate at a voltage near either the positive or negative supply voltage. We will use the integrator in filter circuits, which are covered in Chapter 15.

We will show in Chapter 14 that nonzero bias currents into the op-amp greatly influence the characteristics of this circuit. A dc current through the capacitor will cause the output voltage to linearly change with time until the positive or negative

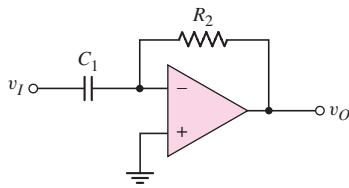


Figure 9.31 Op-amp differentiator

supply voltage is reached. In many applications, a transistor switch needs to be added in parallel with the capacitor to periodically set the capacitor voltage to zero.

The second generalized inverting op-amp uses a capacitor for Z_1 and a resistor for Z_2 , as shown in Figure 9.31. The impedances are $Z_1 = 1/sC_1$ and $Z_2 = R_2$, and the voltage transfer function is

$$\frac{v_O}{v_I} = -\frac{Z_2}{Z_1} = -sR_2C_1 \quad (9.71(a))$$

The output voltage is

$$v_O = -sR_2C_1v_I \quad (9.71(b))$$

Equation (9.71(b)) represents differentiation in the time domain, as follows:

$$v_O(t) = -R_2C_1 \frac{dv_I(t)}{dt} \quad (9.72)$$

The circuit in Figure 9.31 is therefore a differentiator.

Differentiator circuits are more susceptible to noise than are the integrator circuits. Input noise fluctuations of small amplitudes may have large derivatives. When differentiated, these noise fluctuations may generate large noise signals at the output, creating a poor output signal to noise ratio. This problem may be alleviated by placing a resistor in series with the input capacitor. This modified circuit then differentiates low-frequency signals but has a constant high-frequency gain.

EXAMPLE 9.9

Objective: Determine the time constant required in an integrator.

Consider the integrator shown in Figure 9.30. Assume that voltage V_C across the capacitor is zero at $t = 0$. A step input voltage of $v_I = -1$ V is applied at $t = 0$. Determine the time constant required such that the output reaches +10 V at $t = 1$ ms.

Solution: From Equation (9.70), we have

$$v_o = \frac{-1}{R_1C_2} \int_0^t (-1) dt' = \frac{1}{R_1C_2} t' \Big|_0^t = \frac{t}{R_1C_2}$$

At $t = 1$ ms, we want $v_O = 10$ V. Therefore,

$$10 = \frac{10^{-3}}{R_1C_2}$$

which means the time constant is $R_1C_2 = 0.1$ ms.

Comment: As an example, for a time constant of 0.1 ms, we could have $R_1 = 10$ k Ω and $C_2 = 0.01$ μ F, which are reasonable values of resistance and capacitance.

EXERCISE PROBLEM

Ex 9.9: An integrator with input and output voltages that are zero at $t = 0$ is driven by the input signal shown in Figure 9.32. (a) For circuit parameters $R_1 = 10 \text{ k}\Omega$ and $C_2 = 0.1 \mu\text{F}$, determine the output voltage at $t =$ (i) 1 ms, (ii) 2 ms, (iii) 3 ms, and (iv) 4 ms. (b) Repeat part (a) for circuit parameters $R_1 = 10 \text{ k}\Omega$ and $C_2 = 1 \mu\text{F}$. (Ans. (a) (i) -1 V, (ii) 0, (iii) -1 V, (iv) 0; (b) (i) -0.1 V, (ii) 0, (iii) -0.1 V, (iv) 0)

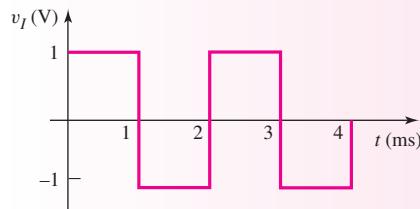


Figure 9.32 Figure for Exercise Ex 9.9

9.5.6 Nonlinear Circuit Applications

Up to this point in the chapter, we have used linear passive elements in conjunction with the op-amp. Many useful circuits can be fabricated if nonlinear elements, such as diodes or transistors, are used in the op-amp circuits. We will consider three simple examples to illustrate the types of nonlinear characteristics that can be generated and to illustrate the general analysis technique.

Precision Half-Wave Rectifier

An op-amp and diode are combined as shown in Figure 9.33 to form a precision half-wave rectifier. For $v_I > 0$, the circuit behaves as a voltage follower. The output voltage is $v_O = v_I$, the load current i_L is positive, and a positive diode current is induced such that $i_D = i_L$. The feedback loop is closed through the forward-biased diode. The output voltage of the op-amp, v_{O1} , adjusts itself to exactly absorb the forward voltage drop of the diode.

For $v_I < 0$, the output voltage tends to go negative, which tends to produce negative load and diode currents. However, a negative diode current cannot exist, so the diode cuts off, the feedback loop is broken, and $v_O = 0$.

The voltage transfer characteristics are shown in Figure 9.34. The rectification is precise in that, even at small positive input voltages, $v_O = v_I$ and we do not observe a diode cut-in voltage.

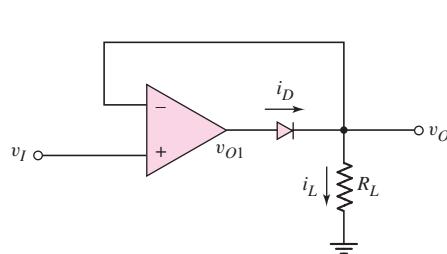


Figure 9.33 Precision half-wave rectifier circuit

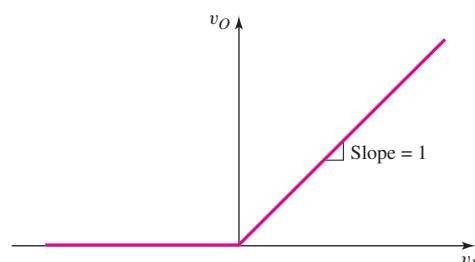


Figure 9.34 Voltage transfer characteristics of precision half-wave rectifier

A potential problem in this circuit exists for $v_I < 0$. The feedback loop is broken so that the op-amp output voltage v_{O1} will saturate near the negative supply voltage. When v_I switches positive, it will take time for the internal circuit to recover, so the response time of the output voltage may be relatively slow. In addition, for $v_I < 0$ and $v_O = 0$, there is now a voltage difference applied across the input terminals of the op-amp. Most op-amps provide input voltage protection so the op-amp will not be damaged in this case. However, if the op-amp does not have input protection, the op-amp may be damaged if the input voltage is larger than 5 or 6 V.

Log Amplifier

Consider the circuit in Figure 9.35. The diode is to be forward biased, so the input signal voltage is limited to positive values. The diode current is

$$i_D = I_S(e^{v_D/V_T} - 1) \quad (9.73(a))$$

If the diode is sufficiently forward biased, the (-1) term is negligible, and

$$i_D \cong I_S e^{v_D/V_T} \quad (9.73(b))$$

The input current can be written

$$i_1 = \frac{v_I}{R_1} \quad (9.74)$$

and the output voltage, since v_1 is at virtual ground, is given by

$$v_O = -v_D \quad (9.75)$$

Noting that $i_1 = i_D$, we can write

$$i_1 = \frac{v_I}{R_1} = i_D = I_S e^{-v_O/V_T} \quad (9.76)$$

If we take the natural log of both sides of this equation, we obtain

$$\ln\left(\frac{v_I}{I_S R_1}\right) = -\frac{v_O}{V_T} \quad (9.77(a))$$

or

$$v_O = -V_T \ln\left(\frac{v_I}{I_S R_1}\right) \quad (9.77(b))$$

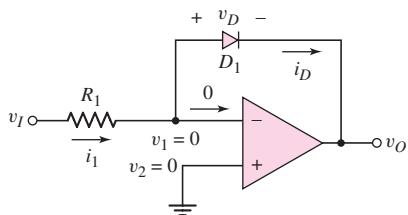


Figure 9.35 Simple log amplifier

Equation (9.77(b)) indicates that, for this circuit, the output voltage is proportional to the log of the input voltage. One disadvantage of this circuit is that the reverse-saturation current I_S is a strong function of temperature, and it varies substantially from one diode to another. A more sophisticated circuit uses bipolar transistors to eliminate the I_S parameter in the log term. This circuit will not be considered here.

Antilog or Exponential Amplifier

The complement, or inverse function, of the log amplifier is the antilog, or exponential, amplifier. A simple example using a diode is shown in Figure 9.36. Since v_1 is at virtual ground, we can write for $v_I > 0$

$$i_D \cong I_S e^{v_I/V_T} \quad (9.78)$$

and

$$v_O = -i_2 R = -i_D R \quad (9.79(a))$$

or

$$v_O = -I_S R \cdot e^{v_I/V_T} \quad (9.79(b))$$

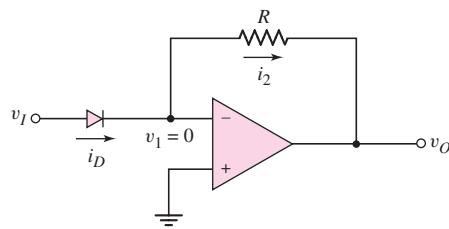


Figure 9.36 A simple antilog, or exponential, amplifier

The output voltage is an exponential function of the input voltage. Again, there are more sophisticated circuits that perform this function, but they will not be considered here.

Test Your Understanding

TYU 9.9 A current source has an output impedance of $R_S = 100 \text{ k}\Omega$. Design a current-to-voltage converter with an output voltage of $v_O = -10 \text{ V}$ when the signal current is $i_S = 100 \mu\text{A}$. (Ans. Figure 9.20 with $R_F = 100 \text{ k}\Omega$)

TYU 9.10 Design the voltage-to-current converter shown in Figure 9.22 such that the load current in a 500Ω load can be varied between 0 and 1 mA with an input voltage between 0 and -5 V . Assume the op-amp is biased at $\pm 10 \text{ V}$. (Ans. $R_2 = 5 \text{ k}\Omega$; for example, let $R_3 = 7 \text{ k}\Omega$, $R_1 = 10 \text{ k}\Omega$, $R_F = 14 \text{ k}\Omega$)

TYU 9.11 All parameters associated with the instrumentation amplifier in Figure 9.26 are as given in Exercise Ex 9.8, except that resistor R_2 associated with the A_1 op-amp is $R_2 = 50 \text{ k}\Omega \pm 5\%$. (a) Determine the maximum and minimum possible values of the common-mode gain. (b) Determine the maximum and minimum possible values of the differential-mode gain. (c) Determine the minimum CMRR(dB). (Ans. (a) $A_{cm} = 0$; (b) $A_d(\min) = 5.87$, $A_d(\max) = 156.75$; (c) CMRR = ∞)

TYU 9.12 Design the instrumentation amplifier in Figure 9.26 such that the variable differential voltage gain is in the range of 5 to 500. The range of the input voltages is between -2 mV and $+2 \text{ mV}$, and the maximum current in R_1 is to be limited to $2 \mu\text{A}$. Set the gain of the difference amplifier to 2.5. (Ans. $R_1(\text{fixed}) = 2 \text{ k}\Omega$, $R_2 = 199 \text{ k}\Omega$, $R_1(\text{var}) = 396 \text{ k}\Omega$)

TYU 9.13 An integrator is driven by the series of pulses shown in Figure 9.37. At the end of the tenth pulse, the output voltage is to be $v_O = -5 \text{ V}$. Assume $V_C = 0$ at $t = 0$.

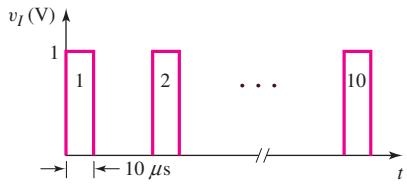


Figure 9.37 Figure for Exercise TYU 9.13

Determine the time constant and values of R_1 and C_2 that will meet these specifications. (Ans. $\tau = 20 \mu s$; for example, let $C_2 = 0.01 \mu F$, $R_1 = 2 k\Omega$)



9.6 OPERATIONAL TRANSCONDUCTANCE AMPLIFIERS

Objective: • Discuss the operational transconductance amplifier.

The operational amplifiers considered up to this point have been voltage amplifiers. The input signal is a voltage and the output signal is a voltage.

Another type of op-amp is an operational transconductance amplifier (OTA). This op-amp is a voltage-input, current-output amplifier. Its circuit symbol is shown in Figure 9.38(a) and the equivalent circuit model is given in Figure 9.38(b). For the ideal OTA, both the input and output impedance is infinite. (The output impedance of an ideal current source is infinite.) The output current for the ideal circuit can be written as

$$i_o = g_m v_d \quad (9.80)$$

where g_m is called the *unloaded transconductance*, with units of amperes per volt. The transconductance can be varied by changing the control current in the op-amp circuit. The OTA can then be used to electronically program functions.

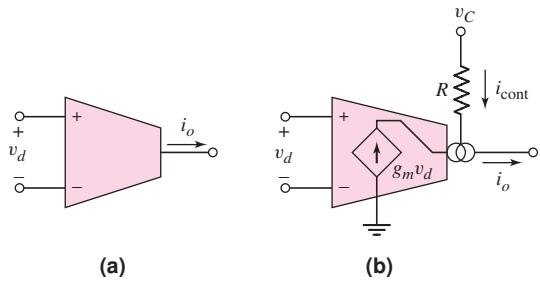


Figure 9.38 (a) Circuit symbol of the OTA. (b) Equivalent circuit model of the OTA.

We will see examples of actual OTA circuits in Chapter 13.

One example of an OTA application is shown in Figure 9.39. This circuit is a simple voltage-controlled amplifier. The output op-amp is configured as a current-to-voltage converter. We see that

$$v_o = -i_o R_F = -i_o (25 k\Omega) \quad (9.81)$$

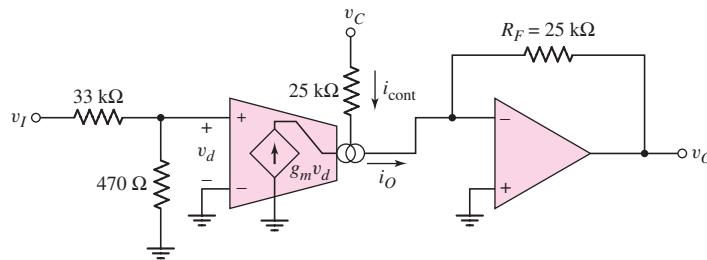


Figure 9.39 Example of a voltage-controlled voltage amplifier using an OTA

and

$$i_O = g_m v_d \quad (9.82)$$

where

$$v_d = \frac{470}{470 + 33,000} \cdot v_I = 0.014v_I \quad (9.83)$$

From the OTA circuit, we have

$$g_m = \frac{i_{\text{cont}}}{2V_T} \quad (9.84)$$

where $V_T = 0.026$ V at room temperature. The control current is given by

$$i_{\text{cont}} = \frac{v_C}{25 \text{ k}\Omega} \quad (9.85)$$

where v_C may be in the range $2 \leq v_C \leq 10$ V. The transconductance of the transconductance operational amplifier is controlled by the control voltage v_C .

Combining equations, we can write the voltage gain as

$$A_v = \frac{v_O}{v_I} = 0.269v_C \quad (9.86)$$

The amplifier shown in Figure 9.39 is then a voltage-controlled voltage amplifier. The amplification factor is a function of the control voltage v_C . This circuit can be used as an amplitude modulator. The v_I input may be the carrier signal and the v_C input may be the audio signal.

OTAs can also be used to design voltage-controlled filters and voltage-controlled oscillators.



9.7

OP-AMP CIRCUIT DESIGN

Objective: • Design several ideal op-amp circuits with given design specifications.

Up to this point, we have mainly been concerned with analyzing ideal op-amp circuits and designing a few basic op-amp circuits. In this section, we will design three specific op-amp circuits. We will assume that these circuits will be fabricated as integrated circuits so that we are not limited to standard resistor values.

9.7.1 Summing Op-Amp Circuit Design

In an inverting summing op-amp, each input is connected to the inverting terminal through a resistor. The summing op-amp can be designed such that the output is

$$v_O = -a_1 v_{I1} - a_2 v_{I2} + a_3 v_{I3} + a_4 v_{I4} \quad (9.87)$$

where the coefficients a_i are all positive. In one design, we could apply voltages v_{I3} and v_{I4} to inverter amplifiers and use the summing op-amp considered previously. This design would require three such op-amps. Alternatively, we could use the results of Exercise TYU 9.8 to design a summing circuit that uses only one op-amp and is more versatile.

Consider the circuit shown in Figure 9.40. Resistor R_C provides more versatility in the design. When we consider nonideal effects, such as bias currents, in op-circuits, in Chapter 14, we will impose a design constraint on the relationship between the resistors connected to the inverting and noninverting terminals. In this section, we will continue to use the ideal op-amp.

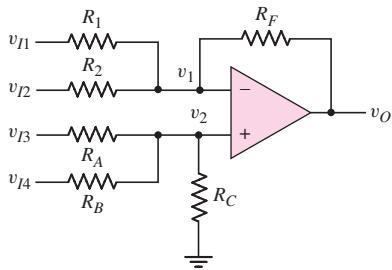


Figure 9.40 Generalized op-amp summing amplifier

To determine the output voltage of our circuit, we use superposition. The inputs v_{I1} and v_{I2} produce the usual outputs, as follows:

$$v_O(v_{I1}) = -\frac{R_F}{R_1} v_{I1} \quad (9.88(a))$$

and

$$v_O(v_{I2}) = -\frac{R_F}{R_2} v_{I2} \quad (9.88(b))$$

We then determine the output due to v_{I3} , with all other inputs set equal to zero. We can write

$$v_2(v_{I3}) = \frac{R_B \| R_C}{R_A + R_B \| R_C} v_{I3} = v_1(v_{I3}) \quad (9.89)$$

Since $v_{I1} = v_{I2} = 0$, the voltage $v_2(v_{I3})$ is the input to a noninverting op-amp with R_1 and R_2 in parallel.

Then,

$$v_O(v_{I3}) = \left(1 + \frac{R_F}{R_1 \| R_2}\right) v_1(v_{I3}) = \left(1 + \frac{R_F}{R_1 \| R_2}\right) \left(\frac{R_B \| R_C}{R_A + R_B \| R_C}\right) v_{I3} \quad (9.90)$$

which can be rearranged as follows:

$$v_O(v_{I3}) = \left(1 + \frac{R_F}{R_N}\right) \left(\frac{R_P}{R_A}\right) v_{I3} \quad (9.91)$$

Here, we define

$$R_N = R_1 \parallel R_2 \quad (9.92(a))$$

and

$$R_P = R_A \parallel R_B \parallel R_C \quad (9.92(b))$$

The output voltage due to v_{I4} is similarly determined and is

$$v_O(v_{I4}) = \left(1 + \frac{R_F}{R_N}\right) \left(\frac{R_P}{R_B}\right) v_{I4} \quad (9.93)$$

The total output voltage is then the sum of the individual terms, or

$$v_O = -\frac{R_F}{R_1} v_{I1} - \frac{R_F}{R_2} v_{I2} + \left(1 + \frac{R_F}{R_N}\right) \left[\frac{R_P}{R_A} v_{I3} + \frac{R_P}{R_B} v_{I4} \right] \quad (9.94)$$

This form of the output voltage is the same as the desired output given by Equation (9.87).

DESIGN EXAMPLE 9.10

Objective: Design a summing op-amp to produce the output

$$v_O = -10v_{I1} - 4v_{I2} + 5v_{I3} + 2v_{I4}$$

The smallest resistor value allowable is $20\text{k}\Omega$. Consider the circuit in Figure 9.40.

Solution: First we determine the values of resistors R_1 , R_2 , and R_F , and then we can determine the noninverting terms. We know that

$$\frac{R_F}{R_1} = 10 \quad \text{and} \quad \frac{R_F}{R_2} = 4$$

Resistor R_1 will be the smallest value, so we can set $R_1 = 20\text{k}\Omega$. Then,

$$R_F = 200\text{k}\Omega \quad \text{and} \quad R_2 = 50\text{k}\Omega$$

The multiplying factor in the noninverting terms becomes

$$\left(1 + \frac{R_F}{R_1 \parallel R_2}\right) = \left(1 + \frac{200}{20 \parallel 50}\right) = 15$$

We then need

$$(15) \left(\frac{R_P}{R_A}\right) = 5 \quad \text{and} \quad (15) \left(\frac{R_P}{R_B}\right) = 2$$

If we take the ratio of these two expressions, we have

$$\frac{R_B}{R_A} = \frac{5}{2}$$

If we choose $R_A = 80 \text{ k}\Omega$, then $R_B = 200 \text{ k}\Omega$, $R_P = 26.67 \text{ k}\Omega$, and R_C becomes $R_C = 50 \text{ k}\Omega$.

Comment: We could change the number of inputs to either the inverting or noninverting terminal, depending on the desired output versus input voltage response.

EXERCISE PROBLEM

Ex 9.10: Consider the summing op-amp in Figure 9.40. Let $R_F = 80 \text{ k}\Omega$, $R_1 = 40 \text{ k}\Omega$, $R_2 = 20 \text{ k}\Omega$, $R_A = R_B = 50 \text{ k}\Omega$, and $R_C = 100 \text{ k}\Omega$. (a) Determine the output voltage in terms of the input voltages. (b) Determine v_O for (i) $v_{I1} = 0.1 \text{ V}$, $v_{I2} = 0.15 \text{ V}$, $v_{I3} = 0.2 \text{ V}$, $v_{I4} = 0.3 \text{ V}$; and for (ii) $v_{I1} = -0.2 \text{ V}$, $v_{I2} = 0.25 \text{ V}$, $v_{I3} = -0.1 \text{ V}$, $v_{I4} = 0.2 \text{ V}$. (Ans. (a) $v_O = -2v_{I1} - 4v_{I2} + 2.8v_{I3} + 2.8v_{I4}$; (b) (i) $v_O = 0.6 \text{ V}$, (ii) $v_O = -0.32 \text{ V}$)

9.7.2 Reference Voltage Source Design

In Chapter 2, we discussed the use of Zener diodes to provide a constant or reference voltage source. A limitation, however, was that the reference voltage could never be greater than the Zener voltage. Now, we can combine a Zener diode with an op-amp to provide more flexibility in the design of reference voltage sources.

Consider the circuit shown in Figure 9.41. Voltage source V_S and resistor R_S bias the Zener diode in the breakdown region. The op-amp is then used as a noninverting amplifier. The output voltage is

$$V_O = \left(1 + \frac{R_2}{R_1}\right)V_Z \quad (9.95)$$

The output current to the load circuit is supplied by the op-amp. A change in the load current will not produce a change in the Zener diode current; consequently, voltage regulation is much improved compared to the simple Zener diode voltage source previously considered.

Since the incremental Zener resistance is not zero, the Zener diode voltage is a slight function of the diode current. The circuit shown in Figure 9.42 is less affected by variations in V_S , since V_S is used only to start up the circuit. The Zener diode begins to conduct when

$$\frac{R_4}{R_3 + R_4} V_S > V_Z + V_D \cong V_Z + 0.7 \quad (9.96)$$

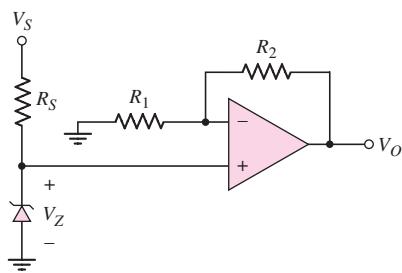


Figure 9.41 Simple op-amp voltage reference circuit

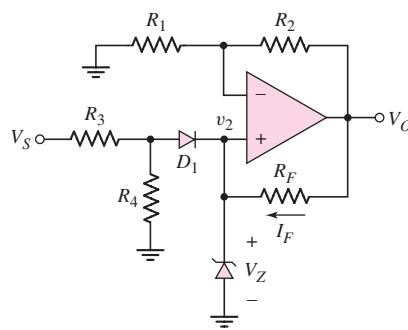


Figure 9.42 Op-amp voltage reference circuit

At this specific voltage, we have

$$V_O = \left(1 + \frac{R_2}{R_1}\right) V_Z \quad (9.97)$$

and

$$I_F = \frac{V_O - V_Z}{R_F} = \frac{R_2 V_Z}{R_1 R_F} \quad (9.98)$$

If V_S decreases and diode D_1 becomes reverse biased, the Zener diode continues to conduct; the Zener diode current is then constant. However, if diode D_1 is conducting, the circuit can be designed such that variations in Zener diode current will be small.

DESIGN EXAMPLE 9.11

Objective: Design a voltage reference source with an output of 10.0 V. Use a Zener diode with a breakdown voltage of 5.6 V. Assume the voltage regulation will be within specifications if the Zener diode is biased between 1–1.2 mA.

Solution: Consider the circuit shown in Figure 9.42. For this example, we need

$$\frac{V_O}{V_Z} = \left(1 + \frac{R_2}{R_1}\right) = \frac{10.0}{5.6}$$

Therefore,

$$\frac{R_2}{R_1} = 0.786$$

We know that

$$I_F = \frac{V_O - V_Z}{R_F}$$

If we set I_F equal to the minimum bias current, we have

$$1 \text{ mA} = \frac{10 - 5.6}{R_F}$$

which means that $R_F = 4.4 \text{ k}\Omega$. If we choose $R_2 = 30 \text{ k}\Omega$, then $R_1 = 38.17 \text{ k}\Omega$.

Resistors R_3 and R_4 can be determined from Figure 9.43. The maximum Zener current supplied by V_S , R_3 , and R_4 should be no more than 0.2 mA. We set the current through D_1 equal to 0.2 mA, for $V_S = 10 \text{ V}$. We then have

$$V'_2 = V_Z + 0.7 = 5.6 + 0.7 = 6.3 \text{ V}$$

Also,

$$I_4 = \frac{V'_2}{R_4} = \frac{6.3}{R_4}$$

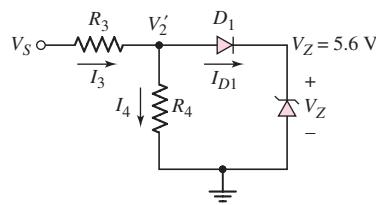


Figure 9.43 Input circuit of the op-amp voltage reference circuit

and

$$I_3 = \frac{V_S - V'_2}{R_3} = \frac{10 - 6.3}{R_3} = \frac{3.7}{R_3}$$

If we set $I_4 = 0.2 \text{ mA}$, then

$$I_3 = 0.4 \text{ mA} \quad R_3 = 9.25 \text{ k}\Omega \quad R_4 = 31.5 \text{ k}\Omega$$

Comment: Voltage V_S is used as a start-up source. Once the Zener diode is biased in breakdown, the output will be maintained at 10.0 V, even if V_S is reduced to zero.

EXERCISE PROBLEM

Ex 9.11: Consider the op-amp voltage reference circuit in Figure 9.42 with parameters given in Example 9.11. Initially set $V_S = 10 \text{ V}$ and then plot, using PSpice, v_O and I_F versus V_S as V_S decreases from 10 to 0 V. Bias the op-amp at $\pm 15 \text{ V}$.

9.7.3 Difference Amplifier and Bridge Circuit Design

A transducer is a device that transforms one form of energy into another form. One type of transducer uses nonelectrical inputs to produce electrical outputs. For example, a microphone converts acoustical energy into electrical energy. A pressure transducer is a device in which, for example, a resistance is a function of pressure, so that pressure can be converted to an electrical signal. Often, the output characteristics of these transducers are measured with a bridge circuit.

Figure 9.44 shows a bridge circuit. Resistance R_3 represents the transducer, and parameter δ is the deviation of R_3 from R_2 due to the input response of the transducer. The output voltage v_{O1} is a measure of δ . If v_{O1} is an open-circuit voltage, then

$$v_{O1} = \left[\frac{R_2(1 + \delta)}{R_2(1 + \delta) + R_1} - \frac{R_2}{R_1 + R_2} \right] V^+ \quad (9.99)$$

which reduces to

$$v_{O1} = \delta \left(\frac{R_1 \| R_2}{R_1 + R_2} \right) V^+ \quad (9.100)$$

Since neither side of voltage v_{O1} is at ground potential, we must connect v_{O1} to an instrumentation amplifier. In addition, v_{O1} is directly proportional to supply voltage V^+ ; therefore, this bias should be a well-defined voltage reference.

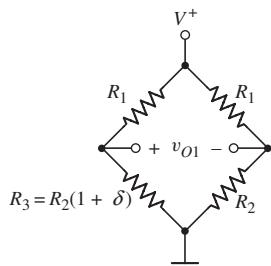


Figure 9.44 Bridge circuit

DESIGN EXAMPLE 9.12

Objective: Design an amplifier system that will produce an output voltage of $\pm 5 \text{ V}$ when the resistance R_3 deviates by $\pm 1\%$ from the value of R_2 . This would occur, for example, in a system where R_3 is a thermistor whose resistance is given by

$$R_3 = 200 \left[1 + \frac{(0.040)(T - 300)}{300} \right] \text{k}\Omega$$

where T is the absolute temperature. For R_3 to vary by $\pm 1\%$ means the temperature is in the range $225 \leq T \leq 375$ K.

Consider biasing the bridge circuit at $V^+ = 7.5$ V using a 5.6 V Zener diode. Assume ± 10 V is available for biasing the op-amp and reference voltage source, and that $R_1 = R_2 = 200$ k Ω .

Solution: With $R_1 = R_2$, from Equation (9.100), we have

$$v_{O1} = \left(\frac{\delta}{4} \right) V^+$$

For $V^+ = 7.5$ V and $\delta = 0.01$, the maximum output of the bridge circuit is $v_{O1} = 0.01875$ V. If the output of the amplifier system is to be +5 V, the gain of the instrumentation amplifier must be $5/0.01875 = 266.7$. Consider the instrumentation amplifier shown in Figure 9.26. The output voltage is given by Equation (9.67), which can be written

$$\frac{v_O}{v_{O1}} = \frac{R'_4}{R'_3} \left(1 + \frac{2R'_2}{R'_1} \right) = 266.7$$

We would like the ratios R'_4/R'_3 and R'_2/R'_1 to be the same order of magnitude. If we let $R'_3 = 15.0$ k Ω and $R'_4 = 187.0$ k Ω , then $R'_4/R'_3 = 12.467$ and $R'_2/R'_1 = 10.195$. If we set $R'_2 = 200.0$ k Ω , then $R'_1 = 19.62$ k Ω .

Resistance R'_1 can be a combination of a fixed resistance in series with a potentiometer, to permit adjustment of the gain.

Comment: The complete design of this instrumentation amplifier is shown in Figure 9.45. Correlation of the reference voltage source design is left as an exercise.

Design Pointer: The design of fairly sophisticated op-amp circuits is quite straightforward when the ideal op-amp parameters are used.

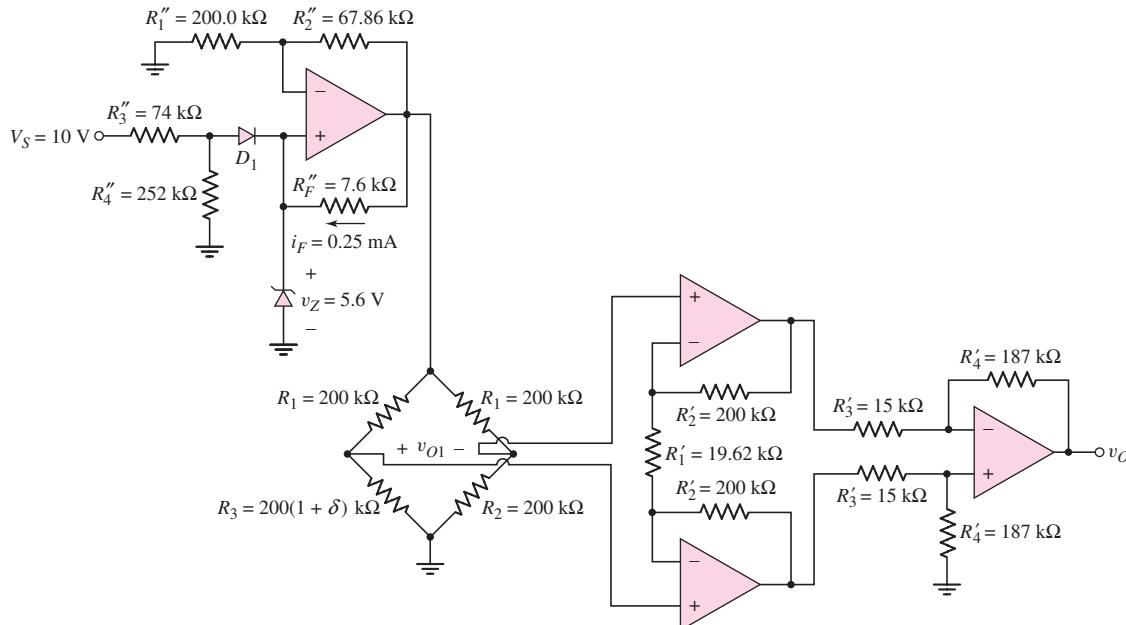


Figure 9.45 Complete amplifier system

Test Your Understanding

TYU 9.14 Consider the bridge circuit in Figure 9.46. The resistance is $R = 20 \text{ k}\Omega$ and the variable resistance ΔR ranges between -100Ω and $+100 \Omega$. The circuit is biased at $V^+ = 5 \text{ V}$. (a) Find v_{O1} as a function of ΔR . (b) Design an amplifier system such that the output is -3 V when $\Delta R = -100 \Omega$. (Ans. (a) $|v_{O1}| = 2.5 \times 10^{-4}(\Delta R)$. (b) For an instrumentation amplifier, let $R_4/R_3 = 10$ and $R_2/R_1 = 5.5$)

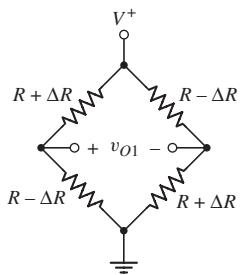


Figure 9.46 Figure for Exercise TYU 9.14

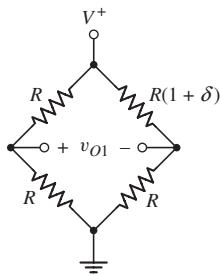


Figure 9.47 Figure for Exercise TYU 9.15

TYU 9.15 The resistance R in the bridge circuit in Figure 9.47 is $50 \text{ k}\Omega$. The circuit is biased at $V^+ = 3 \text{ V}$. (a) Find v_{O1} as a function of δ . (b) Design an amplifier system such that the output varies between $+3 \text{ V}$ and -3 V as the parameter δ varies between $+0.025$ and -0.025 . (Ans. (a) $v_{O1} \cong 0.75\delta$. (b) For an instrumentation amplifier, let $R_4/R_3 = 10$ and $R_2/R_1 = 7.5$)



9.8

DESIGN APPLICATION: ELECTRONIC THERMOMETER WITH AN INSTRUMENTATION AMPLIFIER

Objective: • Design an electronic thermometer with an instrumentation amplifier to provide the necessary amplification.

Specifications: The temperature range to be measured is 0 to 100°F . The output voltage is to be in the range of 0 to 5 V with 0 V corresponding to 0°F and 5 V corresponding to 100°F .

Design Approach: In Chapter 1, we began a design of an electronic thermometer using the temperature characteristics of a pn junction diode. Here, we expand on that design.

Figure 9.48(a) shows a circuit with two diodes, each biased with a constant current source. Figure 9.48(b) shows the same circuit, but with the constant current sources implemented with transistor circuits. The current source circuits were briefly

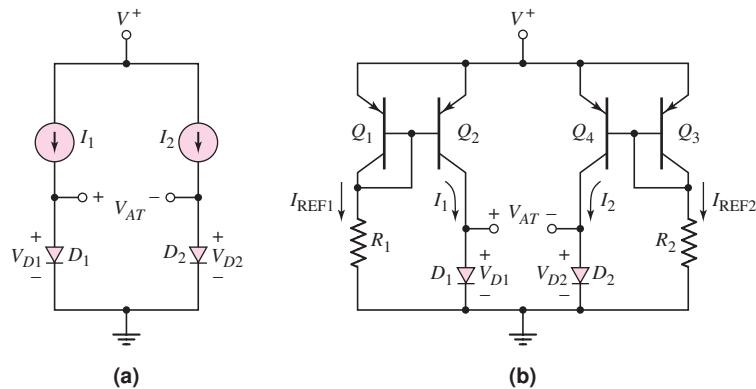


Figure 9.48 (a) Two diodes biased with constant current sources. (b) The same circuit with the constant current sources implemented with transistor circuits.

described and analyzed in Chapter 5. The two diodes, D_1 and D_2 , are assumed to be matched or identical devices. We also assume that all transistors are matched. Neglecting base currents, we have $I_1 = I_{\text{REF1}}$ and $I_2 = I_{\text{REF2}}$.

Choices: Ideal matched silicon diodes and bipolar transistors are available. In addition, ideal op-amps are available.

Solution (Diodes): From Chapter 1, we can write the voltage drops across each diode as

$$V_{D1} = V_T \ln\left(\frac{I_1}{I_S}\right) \quad (9.101(\text{a}))$$

and

$$V_{D2} = V_T \ln\left(\frac{I_2}{I_S}\right) \quad (9.101(\text{b}))$$

We may note that, since the two diodes are matched, the reverse-saturation current, I_S , is the same in the two expressions.

The output voltage is defined as the difference between the voltages across the two diodes, or

$$V_{AT} = V_{D1} - V_{D2} = V_T \left[\ln\left(\frac{I_1}{I_S}\right) - \ln\left(\frac{I_2}{I_S}\right) \right] \quad (9.102(\text{a}))$$

or

$$V_{AT} = V_T \ln\left(\frac{I_1}{I_2}\right) = \frac{kT}{e} \ln\left(\frac{I_{\text{REF1}}}{I_{\text{REF2}}}\right) \quad (9.102(\text{b}))$$

The output voltage, V_{AT} , is now directly proportional to absolute temperature T , hence the subscript AT .

If we let $I_{\text{REF1}}/I_{\text{REF2}} = 5$, then Equation (9.102(b)) can be written as

$$V_{AT} = (0.0259) \left(\frac{T}{300} \right) \ln(5) = (1.3895 \times 10^{-4}) T \quad (9.103)$$

Letting $I_{\text{REF}1}/I_{\text{REF}2} > 0$ provides a small amount of gain. Converting absolute temperature to degrees Celsius and then to degrees Fahrenheit, we find

$$T = T_C + 273.15 \quad (9.104)$$

and

$$T_F = 32 + \frac{9}{5}T_C \Rightarrow T_C = (T_F - 32)\left(\frac{5}{9}\right) \quad (9.105)$$

where T_C and T_F are temperatures in degrees Celsius and degrees Fahrenheit, respectively.

Combining Equations (9.104) and (9.105), we obtain

$$T = (T_F - 32)\left(\frac{5}{9}\right) + 273.15 = \frac{5}{9}T_F + 255.37 \quad (9.106)$$

The output voltage from Equation (9.103) can now be written as

$$\begin{aligned} V_{AT} &= (1.3895 \times 10^{-4}) \left(\frac{5}{9}T_F + 255.37 \right) \\ &= (7.719 \times 10^{-5})T_F + 3.5484 \times 10^{-2} \end{aligned} \quad (9.107)$$

Solution (Instrumentation Amplifier): Since neither terminal of the output voltage is at ground potential, we can apply this voltage to an instrumentation amplifier to obtain a voltage gain. The output of the instrumentation amplifier will be applied to a summing amplifier in addition to an offset voltage. The objective of the design is to obtain an output voltage of zero volts at $T_F = 0$ and an output voltage of 5 V at $T_F = 100^{\circ}\text{F}$.

If the gain of the instrumentation amplifier is $A = -129.55$, then the output of the instrumentation amplifier is as follows:

T_F	V_{AT}	V_{O1}
0	0.035484	-4.5970
100	0.043203	-5.5970

Solution (Output Stage): The offset voltage can be generated by using the noninverting op-amp circuit with a Zener diode, as shown in Figure 9.49. If we use a Zener diode with a breakdown voltage of 3.60 V and if we set $R_3/R_4 = 0.277$, then the output voltage is $V_{O2} = +4.597$ V. Applying the output voltage of the instrumentation amplifier, V_{O1} , and the offset voltage, V_{O2} , to a summing amplifier with a gain of -5 as shown in Figure 9.49, we achieve the desired specifications. That is $V_O = 0$ at $T_F = 0$ and $V_O = 5$ V at $T_F = 100^{\circ}\text{F}$.

Comment: The primary advantage of this system is that the output voltage is a linear function of temperature.

In Chapter 16, we can apply the analog output voltage V_O to an A/D converter and use a seven-segment display so that the output signal is actually displayed in terms of degrees Fahrenheit.

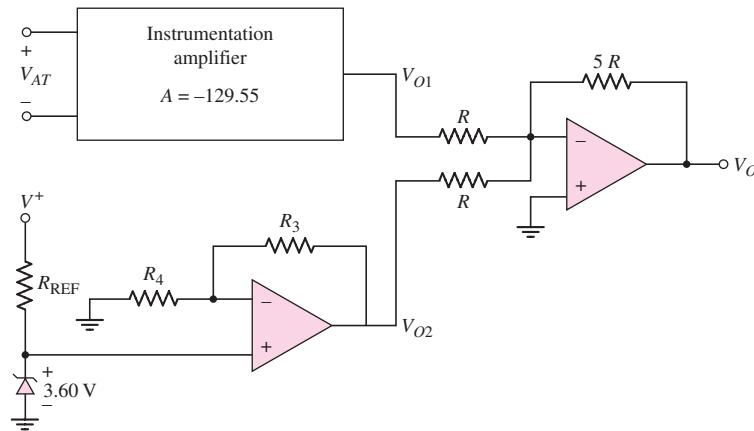


Figure 9.49 The output voltage V_{AT} applied to an instrumentation amplifier, an offset voltage generated by a Zener diode and a noninverting amplifier, and the final output voltage obtained from a summing amplifier

9.9 SUMMARY

- In this chapter, we considered the ideal operational amplifier (op-amp) and various op-amp applications. The op-amp is a three-terminal device (three signal terminals) that ideally amplifies only the difference between two input signals. The op-amp, then, is a high-gain differential amplifier.
- The ideal op-amp model has infinite input impedance (zero input bias currents), infinite open-loop differential voltage gain (zero voltage between the two input terminals), and zero output impedance.
- Two basic op-amp circuits are the inverting amplifier and the noninverting amplifier. For an ideal op-amp, the voltage gain of these circuits is just a function of the ratio of resistors.
- Other amplifier configurations considered were the summing amplifier, voltage follower, current-to-voltage converter, and voltage-to-current converter.
- A versatile circuit is the instrumentation amplifier. The input resistance is essentially infinite and the amplifier gain can be varied by changing a single resistor value.
- If a capacitor is included as a feedback element, the output voltage is the integral of the input voltage. If a capacitor is included as an input element, the output voltage is the derivative of the input voltage. Nonlinear feedback elements, such as diodes or transistors, produce nonlinear transfer functions such as a logarithmic function.
- As an application, an electronic thermometer in conjunction with an instrumentation amplifier was designed to yield a given amplification.

CHECKPOINT

After studying this chapter, the reader should have the ability to:

- ✓ Describe the characteristics of an ideal op-amp.
- ✓ Analyze various op-amp circuits using the ideal op-amp model.
- ✓ Analyze various op-amp circuits, taking into account the finite gain of the op-amp.

- ✓ Understand and describe the characteristics and operation of various op-amp circuits, such as the summing amplifier, difference amplifier, and instrumentation amplifier.
- ✓ Design various op-amp circuits to perform specific functions using the ideal op-amp model.
- ✓ Analyze and design op-amp circuits using nonlinear feedback elements.

REVIEW QUESTIONS

1. Describe the ideal op-amp model and describe the implications of this ideal model in terms of input currents and voltages.
2. Describe the op-amp model including the effect of a finite op-amp voltage gain.
3. Describe the operation and characteristics of the ideal inverting amplifier.
4. What is the concept of virtual ground?
5. What is the significance of a zero output resistance?
6. When a finite op-amp gain is taken into account, is the magnitude of the resulting amplifier voltage gain less than or greater than the ideal value?
7. Describe the operation and characteristics of the ideal summing amplifier.
8. Describe the operation and characteristics of the ideal noninverting amplifier.
9. Describe the voltage follower. What are the advantages of using this circuit?
10. What is the input resistance of an ideal current-to-voltage converter?
11. Describe the operation and characteristics of a difference amplifier.
12. Describe the operation and characteristics of an instrumentation amplifier.
13. Describe the operation and characteristics of an op-amp circuit using a capacitor as a feedback element.
14. Describe the operation and characteristics of an op-amp circuit using a diode as a feedback element.

PROBLEMS

Section 9.1 The Operational Amplifier

- 9.1 Assume an op-amp is ideal, except for having a finite open-loop differential gain. Measurements were made with the op-amp in the open-loop mode. Determine the open-loop gain and complete the following table, which shows the results of those measurements.

v_1	v_2	v_o
-1 mV	+1 mV	1 V
+1 mV		1 V
	1 V	5 V
-1 V	-1 V	
-0.5 V		-3 V

- 9.2 The op-amp in the circuit shown in Figure P9.2 is ideal except it has a finite open-loop gain. (a) If $A_{od} = 10^4$ and $v_o = -2$ V, determine v_I . (b) If $v_I = 2$ V and $v_o = 1$ V, determine A_{od} .
- 9.3 An op-amp is in an open-loop configuration as shown in Figure 9.2. (a) If $v_1 = 2.0010$ V, $v_2 = 2.000$ V, and $A_{od} = 5 \times 10^3$, determine v_o .

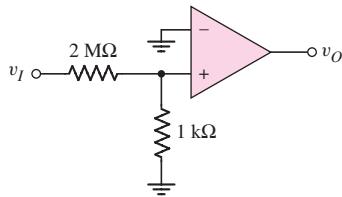


Figure P9.2

- (b) If $v_2 = 3.0025$ V, $v_O = -3.00$ V, and $A_{od} = 2 \times 10^4$, what is v_1 ? (c) If $v_1 = -0.01$ mV, $v_2 = +0.01$ mV, and $v_O = 1.80$ V, determine A_{od} .
- 9.4 Consider the equivalent circuit of the op-amp shown in Figure 9.7(a). Assume terminal v_1 is grounded and the input to terminal v_2 is from a transducer that can be represented by a 0.8 mV voltage source in series with a $25\text{ k}\Omega$ resistance. What is the minimum input resistance R_i such that the minimum differential input voltage is $v_{id} = 0.790$ mV?

Section 9.2 Inverting Amplifier

- 9.5 Consider the ideal inverting op-amp circuit shown in Figure 9.8. Determine the voltage gain $A_v = v_O/v_I$ for (a) $R_2 = 200\text{ k}\Omega$, $R_1 = 20\text{ k}\Omega$; (b) $R_2 = 120\text{ k}\Omega$, $R_1 = 40\text{ k}\Omega$; and (c) $R_2 = 40\text{ k}\Omega$, $R_1 = 40\text{ k}\Omega$.
- 9.6 Assume the op-amps in Figure P9.6 are ideal. Find the voltage gain $A_v = v_O/v_I$ and the input resistance R_i of each circuit.

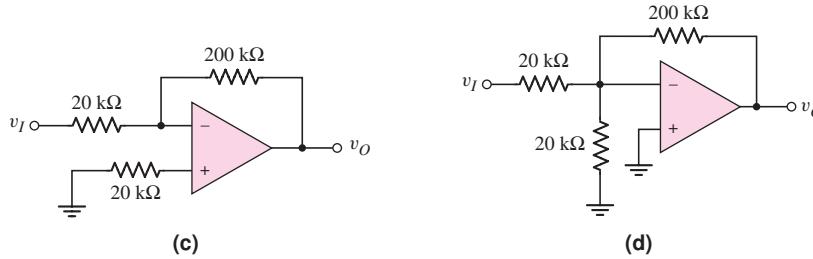
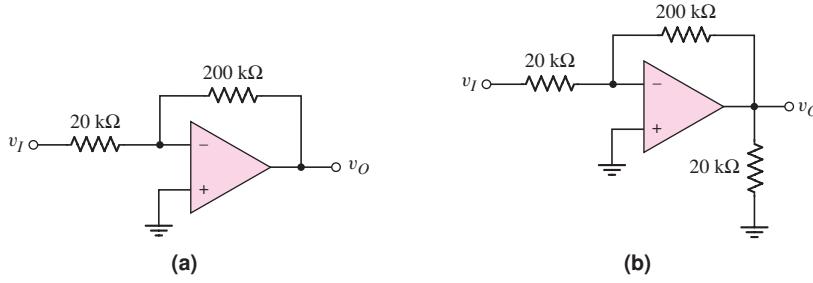


Figure P9.6

- 9.7 Consider an ideal inverting op-amp with $R_2 = 100\text{ k}\Omega$ and $R_1 = 10\text{ k}\Omega$. (a) Determine the ideal voltage gain and input resistance R_i . (b) Repeat part (a) for a second $100\text{ k}\Omega$ resistor connected in parallel with R_2 . (c) Repeat part (a) for a second $10\text{ k}\Omega$ resistance connected in series with R_1 .

- D9.8 (a) Design an inverting op-amp circuit with a closed-loop voltage gain of $A_v = v_O/v_I = -12$. The current in each resistor is to be no larger than $20 \mu\text{A}$ when the output voltage is -4.0 V . (b) Using the results of part (a), determine v_I and the current in each resistor when $v_O = +1.5 \text{ V}$.
- 9.9 Consider an ideal op-amp used in an inverting configuration as shown in Figure 9.8. Determine the closed-loop voltage gain for the following resistor values.
- $R_1 = 20 \text{ k}\Omega, R_2 = 200 \text{ k}\Omega$
 - $R_1 = 20 \text{ k}\Omega, R_2 = 20 \text{ k}\Omega$
 - $R_1 = 20 \text{ k}\Omega, R_2 = 4 \text{ k}\Omega$
 - $R_1 = 50 \text{ k}\Omega, R_2 = 500 \text{ k}\Omega$
 - $R_1 = 50 \text{ k}\Omega, R_2 = 100 \text{ k}\Omega$
 - $R_1 = 50 \text{ k}\Omega, R_2 = 50 \text{ k}\Omega$
- 9.10 Consider the inverting amplifier shown in Figure 9.8. Assume the op-amp is ideal. Determine the resistor values R_1 and R_2 to produce a closed-loop voltage gain of (a) -3.0 , (b) -8.0 , (c) -20 , and (d) -0.50 . In each case the largest resistor is to be limited to $200 \text{ k}\Omega$.
- D9.11 (a) Design an inverting op-amp circuit with a closed-loop voltage gain of $A_v = -6.5$. When in the input voltage is $v_I = -0.25 \text{ V}$, the magnitude of the currents is to be $50 \mu\text{A}$. Determine R_1 and R_2 . (b) Using the results of part (a), find v_I, i_1 , and i_2 when $v_O = -4.0 \text{ V}$.
- D9.12 (a) Design an inverting op-amp circuit such that the closed-loop voltage gain is $A_v = -20$ and the smallest resistor value is $25 \text{ k}\Omega$. (b) Repeat part (a) for the case when the largest resistor value is $1 \text{ M}\Omega$. (c) Determine i_1 in both parts (a) and (b) when the input voltage is $v_I = -0.20 \text{ V}$.
- 9.13 (a) In an inverting op-amp circuit, the nominal resistance values are $R_2 = 300 \text{ k}\Omega$ and $R_1 = 15 \text{ k}\Omega$. The tolerance of each resistor is $\pm 5\%$, which means that each resistance can deviate from its nominal value by $\pm 5\%$. What is the maximum deviation in the voltage gain from its nominal value? (b) Repeat part (a) if the resistor tolerance is reduced to $\pm 1\%$.
- 9.14 (a) The input to the circuit shown in Figure P9.14 is $v_I = -0.20 \text{ V}$. (i) What is v_O ? (ii) Determine i_2, i_O , and i_L . (b) Repeat part (a) for $v_I = +0.05 \text{ V}$. (c) Repeat part (a) for $v_I = 8 \sin \omega t \text{ mV}$.

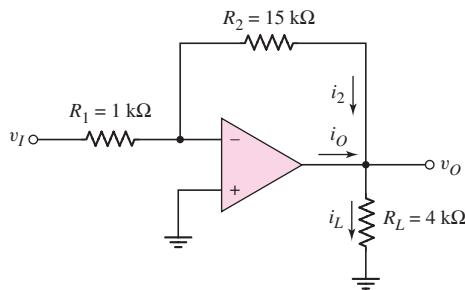


Figure P9.14

- D9.15 Design an inverting amplifier to provide a nominal closed-loop voltage gain of $A_v = -30$. The maximum input voltage signal is 25 mV with a source resistance in the range $1 \text{ k}\Omega \leq R_S \leq 2 \text{ k}\Omega$. The variable source resistance should introduce no more than a 5 percent difference in the gain factor. What is the range in output voltage?

- 9.16 The parameters of the two inverting op-amp circuits connected in cascade in Figure P9.16 are $R_1 = 10 \text{ k}\Omega$, $R_2 = 80 \text{ k}\Omega$, $R_3 = 20 \text{ k}\Omega$, and $R_4 = 100 \text{ k}\Omega$. For $v_I = -0.15 \text{ V}$, determine v_{O1} , v_O , i_1 , i_2 , i_3 , and i_4 . Also determine the current into or out of the output terminal of each op-amp.

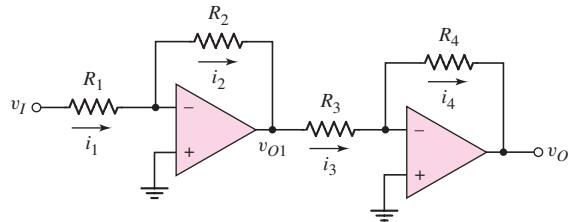


Figure P9.16

- D9.17 Design the cascade inverting op-amp circuit in Figure P9.16 such that the overall closed-loop voltage gain is $A_v = v_O/v_I = 100$ and such that the maximum current in any resistor is limited to $50 \mu\text{A}$ when $v_I = 50 \text{ mV}$, but under the condition that the minimum resistance is $10 \text{ k}\Omega$.
- D9.18 Design an amplifier system with three inverting op-amps circuits in cascade such that the overall closed-loop voltage gain is $A_v = v_O/v_I = -300$. The maximum resistance is limited to $200 \text{ k}\Omega$ and the minimum resistance is limited to $20 \text{ k}\Omega$. In addition, the maximum current in any resistor is to be limited to $60 \mu\text{A}$ when $v_O = 6 \text{ V}$.
- 9.19 Consider the circuit shown in Figure P9.19. (a) Determine the ideal output voltage v_O if $v_I = -0.40 \text{ V}$. (b) Determine the actual output voltage if the open-loop gain of the op-amp is $A_{od} = 5 \times 10^3$. (c) Determine the required value of A_{od} in order that the actual voltage gain be within 0.2 percent of the ideal value.

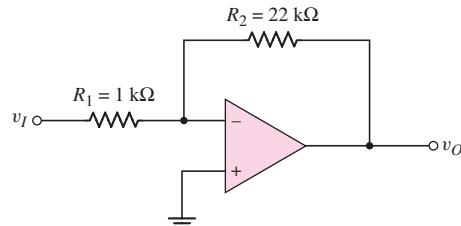


Figure P9.19

- 9.20 The inverting op-amp shown in Figure 9.9 has parameters $R_1 = 25 \text{ k}\Omega$, $R_2 = 100 \text{ k}\Omega$, and $A_{od} = 5 \times 10^3$. The input voltage is from an ideal voltage source whose value is $v_I = 1.0000 \text{ V}$. (a) Calculate the closed-loop voltage gain. (b) Determine the actual output voltage. (c) What is the percentage difference between the actual output voltage and the ideal output voltage. (d) What is the voltage at the inverting terminal of the op-amp?
- 9.21 (a) An op-amp with an open-loop gain of $A_{od} = 7 \times 10^3$ is to be used in an inverting op-amp circuit. Let $R_2 = 100 \text{ k}\Omega$ and $R_1 = 10 \text{ k}\Omega$. If the output voltage is $v_O = 7 \text{ V}$, determine the input voltage and the voltage at the inverting terminal of the op-amp. (b) If the output voltage is $v_O = -5 \text{ V}$ and

- the voltage at the inverting terminal of the op-amp is 0.2 mV, what is the input voltage and the value of A_{od} ?
- 9.22 (a) For the ideal inverting op-amp circuit with T-network, shown in Figure 9.12, the circuit parameters are $R_1 = 10\text{ k}\Omega$, $R_2 = R_3 = 50\text{ k}\Omega$, and $R_4 = 5\text{ k}\Omega$. Determine the closed-loop voltage gain. (b) Determine a new value of R_4 to produce a voltage gain of (i) $A_v = -100$ and (ii) $A_v = -150$.
- D9.23 Consider the ideal inverting op-amp circuit with T-network in Figure 9.12. (a) Design the circuit such that the input resistance is $500\text{ k}\Omega$ and the gain is $A_v = -80$. Do not use resistor values greater than $500\text{ k}\Omega$. (b) For the design in part (a), determine the current in each resistor if $v_I = -0.05\text{ V}$.
- 9.24 An ideal inverting op-amp circuit is to be designed with a closed-loop voltage gain of $A_v = -1000$. The largest resistor value to be used is $500\text{ k}\Omega$. (a) If the simple two-resistor design shown in Figure 9.8 is used, what is the input resistance? (b) If the T-network design shown in Figure 9.12 with $R_3 = 500\text{ k}\Omega$ and $R_2 = R_4 = 250\text{ k}\Omega$ is used, what is the input resistance?
- 9.25 For the op-amp circuit shown in Figure P9.25, determine the gain $A_v = v_O/v_I$. Compare this result to the gain of the circuit shown in Figure 9.12, assuming all resistor values are equal.

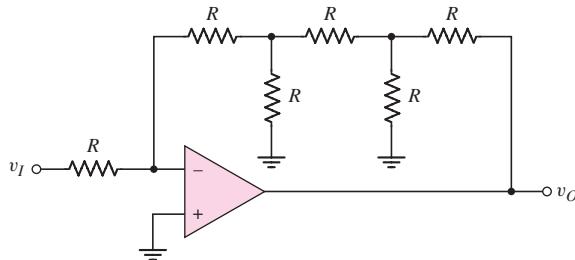


Figure P9.25

- 9.26 The inverting op-amp circuit in Figure 9.9 has parameters $R_1 = 20\text{ k}\Omega$, $R_2 = 200\text{ k}\Omega$, and $A_{od} = 5 \times 10^4$. The output voltage is $v_O = -4.80\text{ V}$. (a) Determine the closed-loop voltage gain. (b) Find the input voltage. (c) Determine the voltage at the inverting terminal of the op-amp. (d) Using v_I from part (b), find the percent error in output voltage compared to the ideal value.
- 9.27 (a) Consider the op-amp circuit in Figure P9.27. The open-loop gain of the op-amp is $A_{od} = 2.5 \times 10^3$. (i) Determine v_O when $v_I = -0.80\text{ V}$. (ii) What is the percent error in output voltage compared to the ideal value? (b) Repeat part (a) for $A_{od} = 200$.

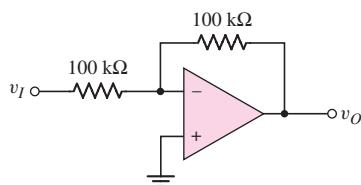


Figure P9.27

- *9.28 The circuit in Figure P9.28 is similar to the inverting amplifier except the resistor R_3 has been added. (a) Derive the expression for v_O in terms of v_I and the resistors. (b) Derive the expression for i_3 in terms of v_I and the resistors.

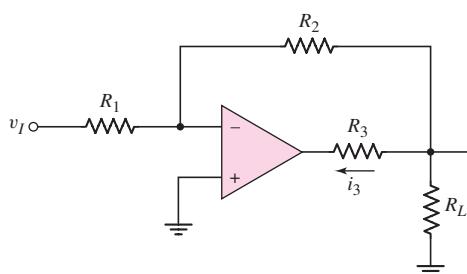


Figure P9.28

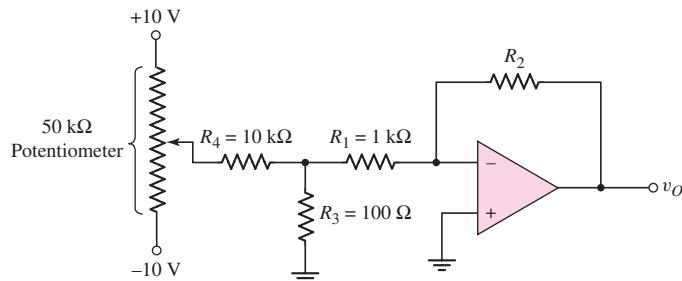


Figure P9.29

- *D9.29 Design the amplifier in Figure P9.29 such that the output voltage varies between ± 10 V as the wiper arm of the potentiometer changes from -10 V to $+10$ V. What is the purpose of including R_3 and R_4 instead of connecting R_1 directly to the wiper arm?

Section 9.3 Summing Amplifier

- 9.30 Consider the ideal inverting summing amplifier in Figure 9.14(a) with parameters $R_1 = 40 \text{ k}\Omega$, $R_2 = 20 \text{ k}\Omega$, $R_3 = 60 \text{ k}\Omega$, and $R_F = 120 \text{ k}\Omega$.
 (a) Determine v_O for $v_{I1} = -0.25 \text{ V}$, $v_{I2} = +0.10 \text{ V}$, and $v_{I3} = +1.5 \text{ V}$.
 (b) Determine v_{I1} for $v_{I2} = +0.25 \text{ V}$, $v_{I3} = -1.2 \text{ V}$, and $v_O = +0.50 \text{ V}$.
- D9.31 (a) Design an ideal inverting summing amplifier to produce an output voltage of $v_O = -2.5(1.2v_{I1} + 2.5v_{I2} + 0.25v_{I3})$. Design the circuit to produce the largest possible input resistance, assuming the largest resistance in the circuit is limited to $400 \text{ k}\Omega$. (b) Using the results of part (a), determine the current in the feedback resistor for $v_{I1} = -1.0 \text{ V}$, $v_{I2} = +0.25 \text{ V}$, and $v_{I3} = +2 \text{ V}$.
- D9.32 Design an ideal inverting summing amplifier to produce an output voltage of $v_O = -2(v_{I1} + 3v_{I2})$. The input voltages are limited to the ranges of $-1 \leq v_{I1} \leq +1 \text{ V}$ and $-0.5 \leq v_{I2} \leq +0.2 \text{ V}$. The current in any resistor is to be limited to a maximum of $80 \mu\text{A}$.
- 9.33 Consider the summing amplifier in Figure 9.14 with $R_F = 10 \text{ k}\Omega$, $R_1 = 1 \text{ k}\Omega$, $R_2 = 5 \text{ k}\Omega$, and $R_3 = 10 \text{ k}\Omega$. If v_{I1} is a 1 kHz sine wave with an rms value of 50 mV , if v_{I2} is a 100 Hz square wave with an amplitude of $\pm 1 \text{ V}$, and if $v_{I3} = 0$, sketch the output voltage v_O .
- 9.34 The parameters for the summing amplifier in Figure 9.14 are $R_F = 100 \text{ k}\Omega$ and $R_3 = \infty$. The two input voltages are $v_{I1} = 4 + 125 \sin \omega t \text{ mV}$ and $v_{I2} = -6 \text{ mV}$. Determine R_1 and R_2 to produce an output voltage of $v_O = -0.5 \sin \omega t \text{ V}$.
- D9.35 (a) Design an ideal summing op-amp circuit to provide an output voltage of $v_O = -2[(v_{I1}/4) + 2v_{I2} + v_{I3}]$. The largest resistor value is to be $250 \text{ k}\Omega$.
 (b) Using the results of part (a), determine the range in output voltage and

the maximum current in R_F if the input voltages are in the ranges $-2 \leq v_{I1} \leq +2$ V, $0 \leq v_{I2} \leq 0.5$ V, and $-1 \leq v_{I3} \leq 0$ V.

- D9.36 An ideal three-input inverting summing amplifier is to be designed. The input voltages are $v_{I1} = 2 + 2 \sin \omega t$ V, $v_{I2} = 0.5 \sin \omega t$ V, and $v_{I3} = -4$ V. The desired output voltage is $v_O = -6 \sin \omega t$ V. The maximum current in any resistor is to be limited to $120 \mu\text{A}$.

- 9.37 A summing amplifier can be used as a digital-to-analog converter (DAC). An example of a 4-bit DAC is shown in Figure P9.37. When switch S_3 is connected to the -5 V supply, the most significant bit is $a_3 = 1$; when S_3 is connected to ground, the most significant bit is $a_3 = 0$. The same condition applies to the other switches S_2 , S_1 , and S_o , corresponding to bits a_2 , a_1 , and a_o , where a_o is the least significant bit. (a) Show that the output voltage is given by

$$v_O = \frac{R_F}{10} \left[\frac{a_3}{2} + \frac{a_2}{4} + \frac{a_1}{8} + \frac{a_o}{16} \right] \quad (5)$$

where R_F is in $\text{k}\Omega$. (b) Find the value of R_F such that $v_O = 2.5$ V when the digital input is $a_3a_2a_1a_o = 1000$. (c) Using the results of part (b), find v_o for: (i) $a_3a_2a_1a_o = 0001$, and (ii) $a_3a_2a_1a_o = 1111$.

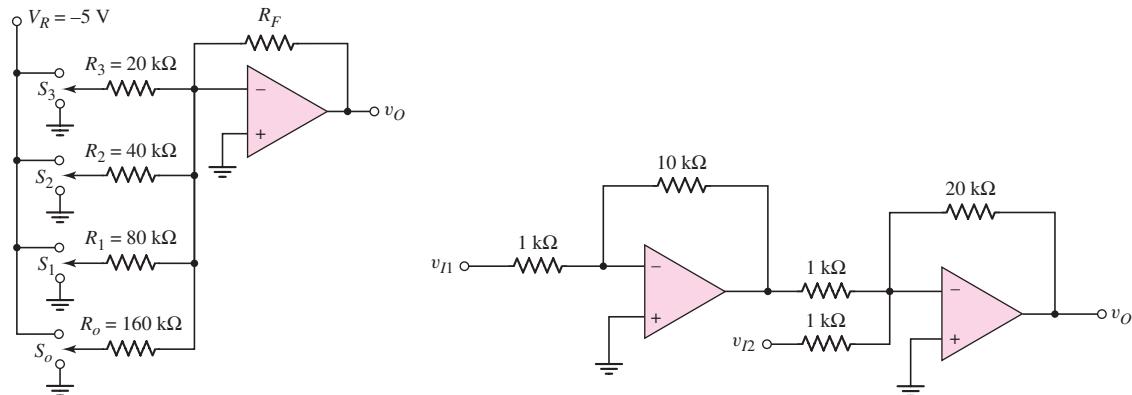


Figure P9.37

Figure P9.38

- 9.38 Consider the circuit in Figure P9.38. (a) Derive the expression for the output voltage v_O in terms of v_{I1} and v_{I2} . (b) Determine v_O for $v_{I1} = +5$ mV and $v_{I2} = -25 - 50 \sin \omega t$ mV. (c) Determine the peak currents in the $10 \text{k}\Omega$ and $20 \text{k}\Omega$ resistors.
- *9.39 Consider the summing amplifier in Figure 9.14 (a). Assume the op-amp has a finite open-loop differential gain A_{od} . Using the principle of superposition, show that the output voltage is given by

$$v_O = \frac{-1}{1 + \frac{(1 + R_F/R_P)}{A_{od}}} \left[\frac{R_F}{R_1} v_{I1} + \frac{R_F}{R_2} v_{I2} + \frac{R_F}{R_3} v_{I3} \right]$$

where $R_P = R_1 \parallel R_2 \parallel R_3$. Demonstrate how the expression will change if more or fewer inputs are included.

Section 9.4 Noninverting Amplifier

- 9.40 Consider the ideal noninverting op-amp circuit in Figure 9.15. Determine the closed-loop gain for the following circuit parameters: (a) $R_1 = 15 \text{ k}\Omega$, $R_2 = 150 \text{ k}\Omega$; (b) $R_1 = 50 \text{ k}\Omega$, $R_2 = 150 \text{ k}\Omega$; (c) $R_1 = 50 \text{ k}\Omega$, $R_2 = 20 \text{ k}\Omega$; and (d) $R_1 = 20 \text{ k}\Omega$, $R_2 = 20 \text{ k}\Omega$.
- D9.41 (a) Design an ideal noninverting op-amp circuit with the configuration shown in Figure 9.15 to have a closed-loop gain of $A_v = 15$. When $v_O = -7.5 \text{ V}$, the current in any resistor is to be limited to a maximum value of $120 \mu\text{A}$. (b) Using the results of part (a), determine the output voltage v_O and the currents in the resistors for $v_I = 0.25 \text{ V}$.
- 9.42 Consider the noninverting amplifier in Figure 9.15. Assume the op-amp is ideal. Determine the resistor values R_1 and R_2 to produce a closed-loop gain of (a) 3, (b) 9, (c) 30, and (d) 1.0. The maximum resistor value is to be limited to $290 \text{ k}\Omega$.
- 9.43 For the circuit in Figure P9.43, the input voltage is $v_I = 5 \text{ V}$. (a) If $v_O = 2.5 \text{ V}$, determine the finite open-loop differential gain of the op-amp. (b) If the open-loop differential gain of the op-amp is 5000, determine v_O .

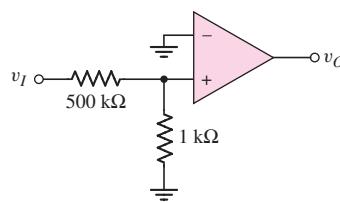


Figure P9.43

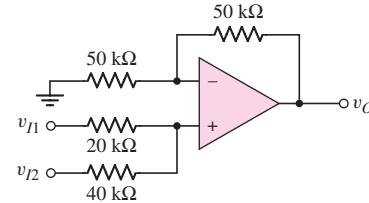


Figure P9.44

- 9.44 Determine v_O as a function of v_{I1} and v_{I2} for the ideal noninverting op-amp circuit in Figure P9.44.
- 9.45 Consider the ideal noninverting op-amp circuit in Figure P9.45. (a) Derive the expression for v_O as a function of v_{I1} and v_{I2} . (b) Find v_O for $v_{I1} = 0.2 \text{ V}$ and $v_{I2} = 0.3 \text{ V}$. (c) Find v_O for $v_{I1} = +0.25 \text{ V}$ and $v_{I2} = -0.40 \text{ V}$.

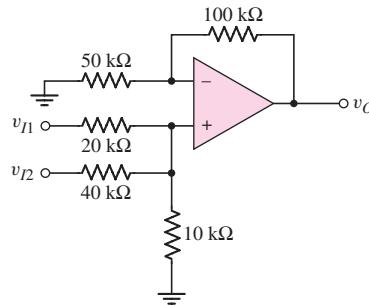


Figure P9.45

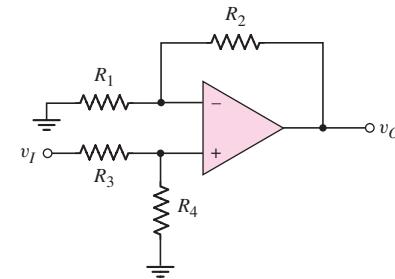


Figure P9.46

- 9.46 (a) Derive the expression for the closed-loop voltage gain $A_v = v_O/v_I$ for the circuit shown in Figure P9.46. Assume an ideal op-amp. (b) Let

$R_4 = 50 \text{ k}\Omega$ and $R_3 = 25 \text{ k}\Omega$. Determine R_1 and R_2 such that $A_v = 6$, assuming the maximum resistor value is limited to $200 \text{ k}\Omega$.

- 9.47 The circuit shown in Figure P9.47 can be used as a variable noninverting amplifier. The circuit uses a $50 \text{ k}\Omega$ potentiometer in conjunction with an ideal op-amp. (a) Derive the expression for the closed-loop voltage gain v_O/v_I in terms of the potentiometer setting x . (b) What is the range of closed-loop voltage gain? (c) Is there a potential problem with this circuit? If so, what is the problem?

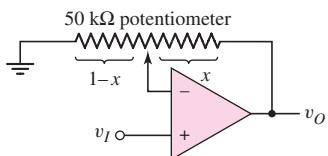


Figure P9.47

- 9.48 (a) Determine the closed-loop voltage gain $A_v = v_O/v_I$ for the ideal op-amp circuit in Figure P9.48. (b) Determine v_O for $v_I = 0.25 \text{ V}$. (c) Let $R = 30 \text{ k}\Omega$. For $v_I = -0.15 \text{ V}$, determine the current in the resistor R in the T-network.

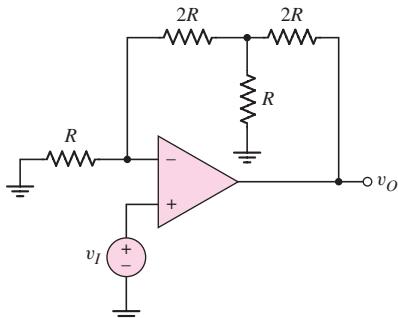


Figure P9.48

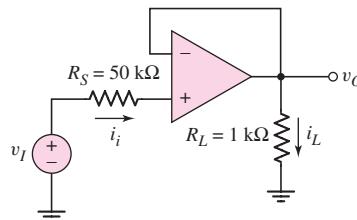


Figure P9.49

- 9.49 For the amplifier in Figure P9.49, determine (a) the ideal closed-loop voltage gain, (b) the actual closed-loop voltage gain if the open-loop gain is $A_{od} = 150,000$, and (c) the open-loop gain such that the actual closed-loop gain is within 1 percent of the ideal.
- 9.50 Consider the voltage-follower circuit in Figure 9.17. Determine the closed-loop voltage gain if the op-amp open-loop voltage gain A_{od} is (a) 20, (b) 200, (c) 2×10^3 , and (d) 2×10^4 .
- 9.51 (a) Consider the ideal op-amp circuit shown in Figure P9.51. Determine the voltage gains $A_{v1} = v_{O1}/v_I$ and $A_{v2} = v_{O2}/v_I$. What is the relationship between v_{O1} and v_{O2} ? (b) For $R_2 = 60 \text{ k}\Omega$, $R_1 = 20 \text{ k}\Omega$, and $R = 50 \text{ k}\Omega$, determine v_{O1} and v_{O2} for $v_I = -0.50 \text{ V}$. (c) Determine $(v_{O1} - v_{O2})$ for $v_I = +0.8 \text{ V}$.

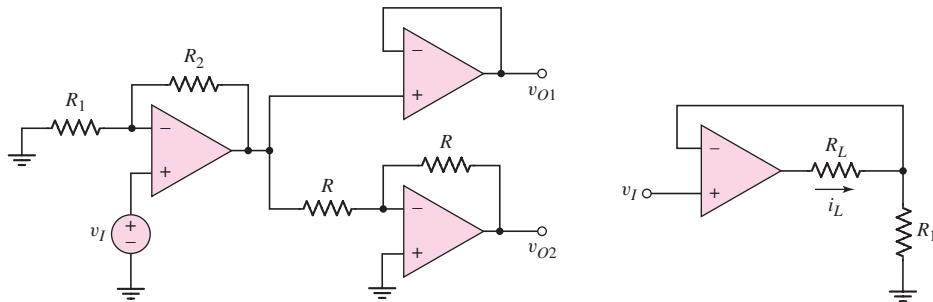


Figure P9.51

Figure P9.52

- 9.52 (a) Assume the op-amp in the circuit in Figure P9.52 is ideal. Determine i_L as a function of v_I . (b) Let $R_1 = 9\text{ k}\Omega$ and $R_L = 1\text{ k}\Omega$. If the op-amp saturates at $\pm 10\text{ V}$, determine the maximum value of v_I and i_L before the op-amp saturates.
- 9.53 Consider the three circuits shown in Figure P9.53. Determine each output voltage for (i) $v_I = 3\text{ V}$ and (ii) $v_I = -5\text{ V}$.

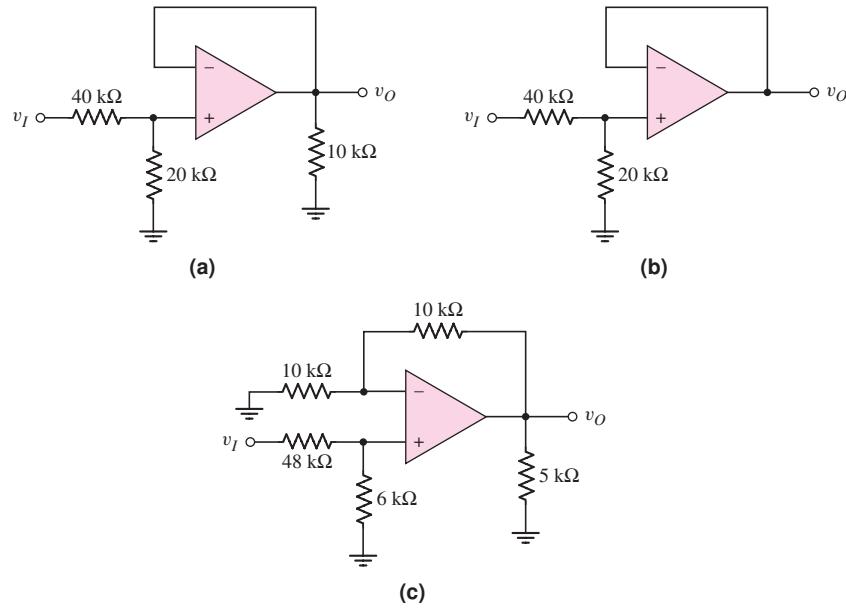


Figure P9.53

Section 9.5 Op-Amp Applications

- *9.54 A current-to-voltage converter is shown in Figure P9.54. The current source has a finite output resistance R_S , and the op-amp has a finite open-loop differential gain A_{od} . (a) Show that the input resistance is given by

$$R_{in} = \frac{R_F}{1 + A_{od}}$$

- (b) If $R_F = 10\text{ k}\Omega$ and $A_{od} = 1000$, determine the range of R_S such that the output voltage deviates from its ideal value by less than 1 percent.

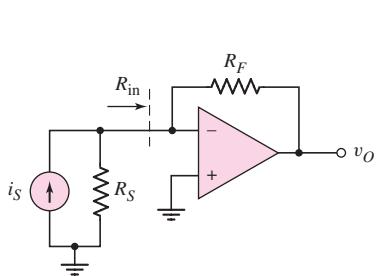


Figure P9.54

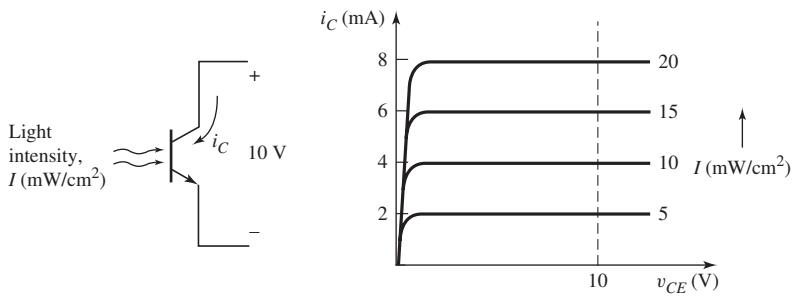


Figure P9.55

- *D9.55 Figure P9.55 shows a phototransistor that converts light intensity into an output current. The transistor must be biased as shown. The transistor output versus input characteristics are shown. Design a current-to-voltage converter to produce an output voltage between 0 and 8 V for an input light intensity between 0 and 20 mW/cm². Power supplies of +10 V and -10 V are available.
- D9.56 The circuit in Figure P9.56 is an analog voltmeter in which the meter reading is directly proportional to the input voltage v_I . Design the circuit such that a 1 mA full-scale reading corresponds to $v_I = 10$ V. Resistance R_2 corresponds to the meter resistance, and R_1 corresponds to the source resistance. How do these resistances influence the design?

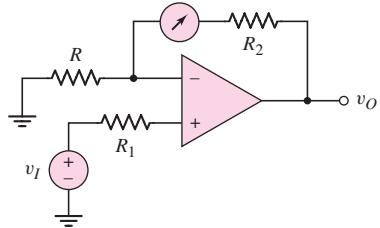


Figure P9.56

- D9.57 Consider the voltage-to-current converter in Figure 9.22 using an ideal op-amp. (a) Design the circuit such that the current in a 200 Ω load can be varied between 0 and 5 mA with an input voltage between 0 and -5 V. Assume the op-amp is biased at ± 15 V. (b) Using the results of part (a), determine voltage v_O ; currents i_2 , i_3 , i_4 ; and the output current of the op-amp for $v_I = -5$ V.
- D9.58 The circuit in Figure P9.58 is used to drive an LED with a voltage source. The circuit can also be thought of as a current amplifier in that, with the proper

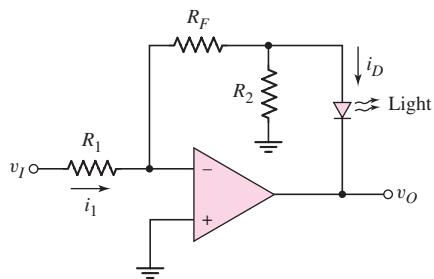


Figure P9.58

design, $i_D > i_1$. (a) Derive the expression for i_D in terms of i_1 and the resistors. (b) Design the circuit such that $i_D = 12 \text{ mA}$ and $i_1 = 1 \text{ mA}$ for $v_I = 5 \text{ V}$.

- *9.59 Figure P9.59 is used to calculate the resistance seen by the load in the voltage-to-current converter given in Figure 9.22. (a) Show that the output resistance is given by

$$R_o = \frac{R_1 R_2 R_3}{R_1 R_3 - R_2 R_F}$$

(b) Using the parameters given in Example 9.5, determine R_o . Is this result unexpected?

(c) Consider the design specification given by Equation (9.44). What is the expected value of R_o ?

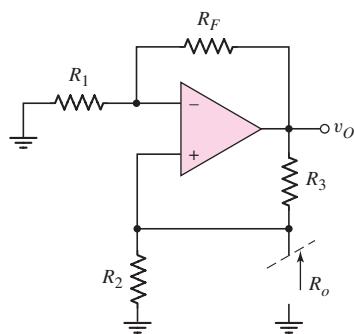


Figure P9.59

- D9.60 Consider the op-amp difference amplifier in Figure 9.24(a). Let $R_1 = R_3$ and $R_2 = R_4$. A load resistor $R_L = 10 \text{ k}\Omega$ is connected from the output terminal to ground. (a) Design the circuit such that the difference voltage gain is $A_d = 15$ and the minimum difference input resistance is $30 \text{ k}\Omega$. (b) If the load current is $i_L = 0.25 \text{ mA}$, what is the differential input voltage ($v_{I2} - v_{I1}$)? (c) If $v_{I1} = 1.5 \text{ V}$ and $v_{I2} = 1.2 \text{ V}$, determine i_L . (d) If $i_L = 0.5 \text{ mA}$ when $v_{I2} = 2.0 \text{ V}$, determine v_{I1} .

- D9.61 Consider the differential amplifier shown in Figure 9.24(a). Let $R_1 = R_3$ and $R_2 = R_4$. Design the amplifier such that the differential voltage gain is (a) 40, (b) 25, (c) 5, and (d) 0.5. In each case the differential input resistance should be as large as possible but under the condition that the largest resistor value is limited to $250 \text{ k}\Omega$.

- *9.62 Consider the differential amplifier shown in Figure 9.24(a). Assume that each resistor is $50(1 \pm x) \text{ k}\Omega$. (a) Determine the worst case common-mode gain $A_{CM} = v_O/v_{CM}$, where $v_{CM} = v_1 = v_2$. (b) Evaluate A_{CM} and CMRR(dB) for $x = 0.01, 0.02$, and 0.05 .

- 9.63 Let $R = 10 \text{ k}\Omega$ in the differential amplifier in Figure P9.63. Determine the voltages v_X , v_Y , v_O and the currents i_1 , i_2 , i_3 , i_4 for input voltages of (a) $v_1 = 1.80 \text{ V}$, $v_2 = 1.40 \text{ V}$; (b) $v_1 = 3.20 \text{ V}$, $v_2 = 3.60 \text{ V}$; and (c) $v_1 = -1.20 \text{ V}$, $v_2 = -1.35 \text{ V}$.

- 9.64 Consider the circuit shown in Figure P9.64. (a) The output current of the op-amp is 1.2 mA and the transistor current gain is $\beta = 75$. Determine the resistance R . (b) Repeat part (a) if the current is 0.2 mA and the transistor current gain is $\beta = 100$. (c) Using the results of part (a), determine

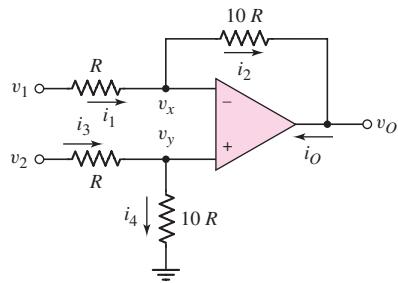


Figure P9.63

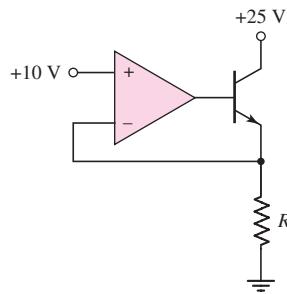


Figure P9.64

the op-amp output current if the input voltage is 6 V. (d) Using the results of part (b), determine the op-amp output current if the input voltage is 4 V.

- *9.65 The circuit in Figure P9.65 is a representation of the common-mode and differential-input signals to a difference amplifier. The output voltage can be written as

$$v_O = A_d v_d + A_{cm} v_{cm}$$

where A_d is the differential-mode gain and A_{cm} is the common-mode gain.

(a) Setting $v_d = 0$, show that the common-mode gain is given by

$$A_{cm} = \frac{\left(\frac{R_4}{R_3} - \frac{R_2}{R_1} \right)}{(1 + R_4/R_3)}$$

- (b) Determine A_{cm} if $R_1 = 10.4\text{ k}\Omega$, $R_2 = 62.4\text{ k}\Omega$, $R_3 = 9.6\text{ k}\Omega$, and $R_4 = 86.4\text{ k}\Omega$. (c) Determine the maximum value of $|A_{cm}|$ if $R_1 = 20\text{ k}\Omega \pm 1\%$, $R_2 = 80\text{ k}\Omega \pm 1\%$, $R_3 = 20\text{ k}\Omega \pm 1\%$, and $R_4 = 80\text{ k}\Omega \pm 1\%$.

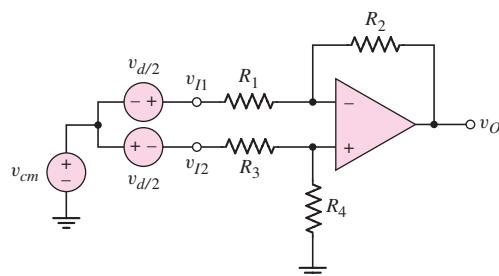


Figure P9.65

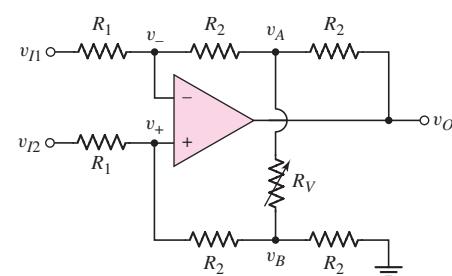


Figure P9.66

- *9.66 Consider the adjustable gain difference amplifier in Figure P9.66. Variable resistor R_V is used to vary the gain. Show that the output voltage v_O , as a function of v_{I1} and v_{I2} , is given by

$$v_O = \frac{2R_2}{R_1} \left(1 + \frac{R_2}{R_V} \right) (v_{I2} - v_{I1})$$

- 9.67 Assume the instrumentation amplifier in Figure 9.26 has ideal op-amps. The circuit parameters are $R_1 = 10\text{ k}\Omega$, $R_2 = 40\text{ k}\Omega$, $R_3 = 40\text{ k}\Omega$, and $R_4 = 120\text{ k}\Omega$. Determine v_{O1} , v_{O2} , v_O , and the current in R_1 for (a) $v_{I2} = 1.2 + 0.08 \sin \omega t$ (V), $v_{I1} = 1.2 - 0.08 \sin \omega t$ (V); and (b) $v_{I2} = -0.60 - 0.05 \sin \omega t$ (V), $v_{I1} = -0.65 + 0.05 \sin \omega t$ (V).
- 9.68 Consider the circuit in Figure P9.68. Assume ideal op-amps are used. The input voltage is $v_I = 0.5 \sin \omega t$. Determine the voltages (a) v_{OB} , (b) v_{OC} , and (c) v_O . (d) What is the voltage gain v_O/v_I ?

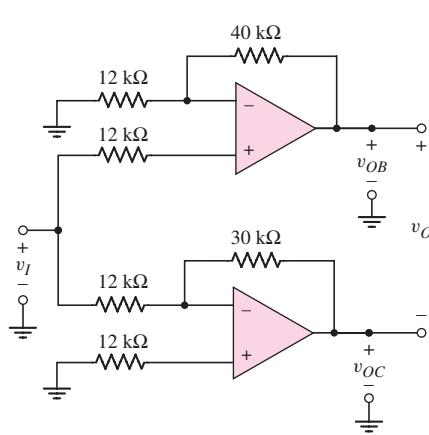


Figure P9.68

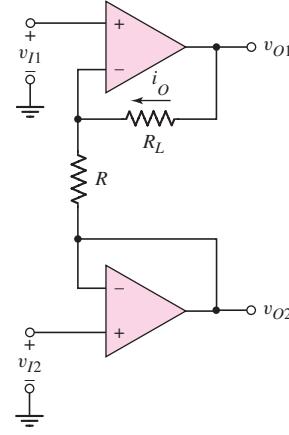


Figure P9.69

- 9.69 Consider the circuit in Figure P9.69. Assume ideal op-amps are used. (a) Derive the expression for the current i_O as a function of input voltages v_{I1} and v_{I2} . (b) Design the circuit such that $i_O = 5\text{ mA}$ for $v_{I1} = 0.25\text{ V}$ and $v_{I2} = -0.25\text{ V}$. (c) Using the results of part (b), determine v_{O1} and v_{O2} if $R_L = 1\text{ k}\Omega$. (d) Determine i_O , v_{O1} , and v_{O2} for $R = 500\text{ }\Omega$, $R_L = 3\text{ k}\Omega$, $v_{I1} = 1.25\text{ V}$, and $v_{I2} = 1.75\text{ V}$.
- 9.70 The instrumentation amplifier in Figure 9.26 has the same circuit parameters and input voltages as given in Problem 9.67, except that R_1 is replaced by a fixed resistance R_{1f} in series with a potentiometer, as shown in Figure 9.28. Determine the values of R_{1f} and the potentiometer resistance if the magnitude of the output has a minimum value of $|v_O| = 0.5\text{ V}$ and a maximum value of $|v_O| = 8\text{ V}$.
- D9.71 Design the instrumentation amplifier in Figure 9.26 such that the variable differential voltage gain covers the range of 5 to 200. Set the gain of the difference amplifier to 2.5. The maximum current in R_1 is to be limited to $50\text{ }\mu\text{A}$ for an output voltage of 10 V . What value of potentiometer is required?
- 9.72 All parameters associated with the instrumentation amplifier in Figure 9.26 are the same as given in Exercise Ex 9.8, except that resistor R_3 , which is connected to the inverting terminal of A3, is $R_3 = 30\text{ k}\Omega \pm 5\%$. Determine the maximum common-mode gain.
- 9.73 The parameters in the integrator circuit shown in Figure 9.30 are $R_1 = 20\text{ k}\Omega$ and $C_2 = 0.02\text{ }\mu\text{F}$. The input signal is $v_I = 0.25 \cos \omega t$ (V). (a) Determine the frequency at which the input and output signals have

equal amplitudes. At this frequency, what is the phase of the output signal with respect to the input? (b) At what frequency will the output signal amplitude be (i) $|v_O| = 1.5 \text{ V}$ and (ii) $|v_O| = 0.15 \text{ V}$?

- 9.74 Consider the ideal op-amp integrator. Assume the capacitor is initially uncharged. (a) The output voltage is $v_O = -5 \text{ V}$ at $t = 1.2 \text{ s}$ after a $+0.25 \text{ V}$ pulse is applied to the input. What is the RC time constant? (b) Use the results of part (a). At $t = 1.2 \text{ s}$, the input changes to -0.10 V . (i) At what time does $v_O = 0$? (ii) At what time does $v_O = +5 \text{ V}$.
- 9.75 The circuit in Figure P9.75 is a first-order low-pass active filter. (a) Show that the voltage transfer function is given by

$$A_v = \frac{-R_2}{R_1} \cdot \frac{1}{1 + j\omega R_2 C_2}$$

(b) What is the voltage gain at dc ($\omega = 0$)? (c) At what frequency is the magnitude of the voltage gain a factor of $\sqrt{2}$ less than the dc value? (This is the -3 dB frequency.)

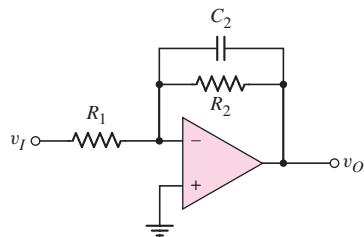


Figure P9.75

- D9.76 (a) Using the results of Problem 9.75, design the low-pass active filter in Figure P9.75 such that the input resistance is $20 \text{ k}\Omega$, the low-frequency gain is -15 , and the -3 dB frequency is 5 kHz . (b) Repeat part (a) such that the input resistance is $15 \text{ k}\Omega$, the low-frequency gain is -25 , and the -3 dB frequency is 15 kHz .
- 9.77 The circuit shown in Figure P9.77 is a first-order high-pass active filter. (a) Show that the voltage transfer function is given by

$$A_v = \frac{-R_2}{R_1} \cdot \frac{j\omega R_1 C_1}{1 + j\omega R_1 C_1}$$

(b) What is the voltage gain as the frequency becomes large? (c) At what frequency is the magnitude of the gain a factor of $\sqrt{2}$ less than the high-frequency limiting value?

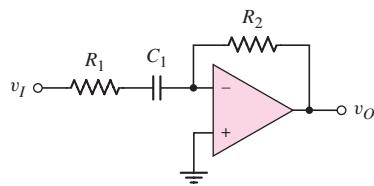


Figure P9.77

- D9.78 (a) Using the results of Problem 9.77, design the high-pass active filter in Figure P9.77 such that the high-frequency voltage gain is -15 and the -3 dB frequency is 20 kHz. The maximum resistance value is to be limited to $350\text{ k}\Omega$. (b) Repeat part (a) such that the high-frequency gain is -25 and the -3 dB frequency is 35 kHz. The minimum resistance value is to be limited to $20\text{ k}\Omega$.
- 9.79 Consider the voltage reference circuit shown in Figure P9.79. Determine v_O , i_2 , and i_Z .

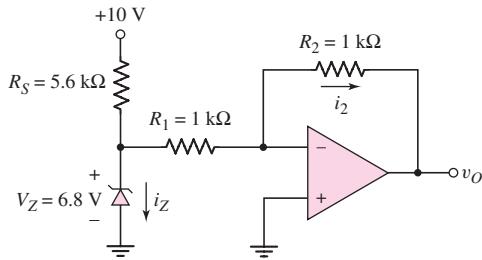


Figure P9.79

- 9.80 Consider the circuit in Figure 9.35. The diode parameter is $I_S = 10^{-14}\text{ A}$ and the resistance is $R_1 = 10\text{ k}\Omega$. Plot v_O versus v_I over the range $20\text{ mV} \leq v_I \leq 2\text{ V}$. (Plot v_I on a log scale.)
- *9.81 In the circuit in Figure P9.81, assume that Q_1 and Q_2 are identical transistors. If $T = 300\text{ K}$, show that the output voltage is

$$v_O = 1.0 \log_{10} \left(\frac{v_2 R_1}{v_1 R_2} \right)$$

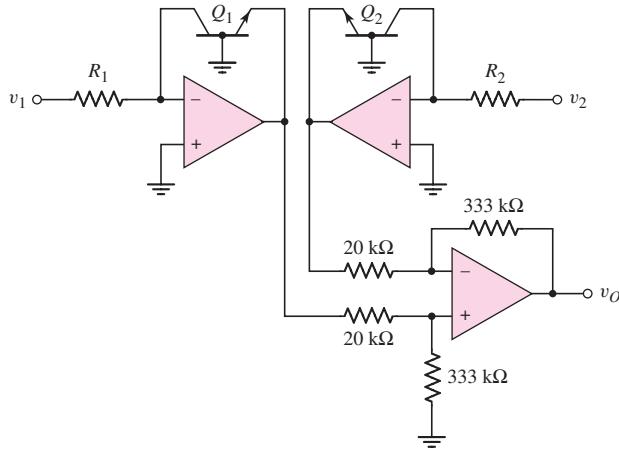


Figure P9.81

- 9.82 Consider the circuit in Figure 9.36. The diode parameter is $I_S = 10^{-14}\text{ A}$ and the resistance is $R_1 = 10\text{ k}\Omega$. Plot v_O versus v_I for $0.30 \leq v_I \leq 0.60\text{ V}$. (Plot v_O on a log scale.)

Section 9.7 Op-Amp Circuit Design

- *D9.83 Design an op-amp summer to produce the output voltage $v_O = 2v_{I1} - 10v_{I2} + 3v_{I3} - v_{I4}$. Assume the largest resistor value is $500\text{ k}\Omega$, and the input impedance seen by each source is the largest value possible.
- *D9.84 Design an op-amp summer to produce an output voltage of $v_O = 3v_{I1} + 1.5v_{I2} + 2v_{I3} - 4v_{I4} - 6v_{I5}$. The largest resistor value is to be $250\text{ k}\Omega$.
- *D9.85 Design a voltage reference source as shown in Figure 9.42 to have an output voltage of 12.0 V. A Zener diode with a breakdown voltage of 5.6 V is available. Assume the voltage regulation will be within specifications if the Zener diode current is within the range of $1.2 \leq I_Z \leq 1.35\text{ mA}$. The start-up voltage V_S is to be 10 V.
- *D9.86 Consider the voltage reference circuit in Figure P9.86. Using a Zener diode with a breakdown voltage of 5.6 V, design the circuit to produce an output voltage of 12.0 V. Assume the input voltage is 15 V and the Zener diode current is $I_Z = 2\text{ mA}$.

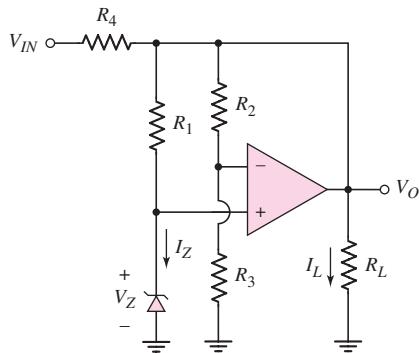


Figure P9.86

- *D9.87 Consider the bridge circuit in Figure P9.87. The resistor R_T is a thermistor with values of $20\text{ k}\Omega$ at $T = 300\text{ K}$ and $21\text{ k}\Omega$ at $T = 350\text{ K}$. Assume that the thermistor resistance is linear with temperature, and that the bridge is biased at $V^+ = 10\text{ V}$. Design an amplifier system with an output of 0 V at $T = 300\text{ K}$ and 5 V at $T = 350\text{ K}$.

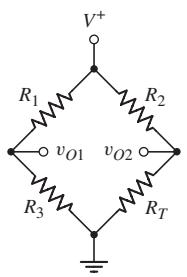


Figure P9.87

- *D9.88 Consider the bridge circuit in Figure 9.46. The resistance R is $20\text{ k}\Omega$ and the bias is $V^+ = 9\text{ V}$. (a) Determine v_{O1} as a function of ΔR . (b) Design an amplifier system such that the output varies from -5 V to $+5\text{ V}$ as ΔR varies from $+200\text{ }\Omega$ to $-200\text{ }\Omega$.



COMPUTER SIMULATION PROBLEMS

- 9.89 Using a computer simulation, verify the design in Example 9.4.
- 9.90 Using a computer simulation, verify the design in Example 9.8.
- 9.91 Using a computer simulation, verify the design in Problem 9.76(b). Plot v_o versus frequency over the range $2 \leq f \leq 50$ kHz.
- 9.92 Using a computer simulation, verify the design in Problem 9.78(a). Plot v_o versus frequency over the range $2 \leq f \leq 100$ kHz.



DESIGN PROBLEMS

See Design Problems 9.83 to 9.88.

[Note: Each design should be correlated with a computer analysis.]

Differential and Multistage Amplifiers

In this chapter, we introduce a special multitransistor circuit configuration called the differential amplifier, or diff-amp. We have encountered a diff-amp previously in our discussion of op-amp circuits. However, the diff-amp, in the context of this chapter, is at the basic transistor level.

The diff-amp is a fundamental building block of analog circuits. It is the input stage of virtually every op-amp, and is the basis of a high-speed digital logic circuit family, called emitter-coupled logic, which will be addressed in Chapter 17.

The design of diff-amps for integrated circuits, in general, incorporates current-source biasing and active loads, which were analyzed in the last chapter. At the end of this chapter, the reader should be able to design both BJT and MOSFET diff-amps to meet particular specifications.

Basic BiCMOS analog circuits are also considered. BiCMOS circuits combine bipolar and MOS transistors on the same semiconductor chip. The advantages of the MOSFET's high input impedance and the bipolar high gain can be utilized in the same circuit.

Up to this point, we have concentrated primarily on the analysis and design of single-stage amplifiers. However, these circuits have limited gain, input resistance, and output resistance characteristics. Multistage or cascaded-stage amplifiers can be designed to produce high gain and specified input and output resistance properties. In this chapter, we begin to consider these multistage amplifiers.

PREVIEW

In this chapter, we will:

- Describe the characteristics and terminology of the ideal differential amplifier.
- Analyze and determine the characteristics of the basic bipolar differential amplifier.
- Analyze and determine the characteristics of the basic MOSFET differential amplifier.
- Determine the characteristics of BJT and MOSFET differential amplifiers with active loads.
- Describe the characteristics of and analyze various BiCMOS circuits.
- Analyze an example of a gain stage and output stage of a multistage amplifier.
- Analyze a simplified multistage bipolar amplifier.
- Analyze the frequency response of the differential amplifier.
- As an application, design a CMOS diff-amp with an output gain stage to meet a set of specifications.



11.1 THE DIFFERENTIAL AMPLIFIER

Objective: • Describe the characteristics and terminology of the ideal differential amplifier.

In Chapters 4 and 6, we discussed the reasons linear amplifiers are necessary in analog electronic systems. In these chapters, we analyzed and designed several configurations of MOSFET and bipolar amplifiers. In these circuits, there was one input terminal and one output terminal.

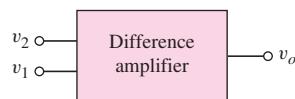


Figure 11.1 Difference amplifier block diagram

In this chapter, we introduce another basic transistor circuit configuration called the differential amplifier. This amplifier, also called a diff-amp, is the input stage to virtually all op-amps and is probably the most widely used amplifier building block in analog integrated circuits. Figure 11.1 is a block diagram of the diff-amp. There are two input terminals and one output terminal. Ideally, the output signal is proportional to only the difference between the two input signals.

The ideal output voltage can be written as

$$v_o = A_{vol}(v_1 - v_2) \quad (11.1)$$

where A_{vol} is called the open-loop voltage gain. In the ideal case, if $v_1 = v_2$, the output voltage is zero. We only obtain a nonzero output voltage if v_1 and v_2 are not equal.

We define the **differential-mode input voltage** as

$$v_d = v_1 - v_2 \quad (11.2)$$

and the **common-mode input voltage** as

$$v_{cm} = \frac{v_1 + v_2}{2} \quad (11.3)$$

These equations show that if $v_1 = v_2$, the differential-mode input signal is zero and the common-mode input signal is $v_{cm} = v_1 = v_2$.

If, for example, $v_1 = +10 \mu\text{V}$ and $v_2 = -10 \mu\text{V}$, then the differential-mode voltage is $v_d = 20 \mu\text{V}$ and the common-mode voltage is $v_{cm} = 0$. However, if $v_1 = 110 \mu\text{V}$ and $v_2 = 90 \mu\text{V}$, then the differential-mode input signal is still $v_d = 20 \mu\text{V}$, but the common-mode input signal is $v_{cm} = 100 \mu\text{V}$. If each pair of input voltages were applied to the ideal difference amplifier, the output voltage in each case would be exactly the same. However, amplifiers are not ideal, and the common-mode input signal does affect the output. One goal of the design of differential amplifiers is to minimize the effect of the common-mode input signal.



11.2 BASIC BJT DIFFERENTIAL PAIR

Objective: • Describe the characteristics of and analyze the basic bipolar differential amplifier.

In this section, we consider the basic bipolar **difference amplifier** or **diff-amp**. We introduce the terminology, qualitatively describe the operation of the circuit, and analyze the dc and small-signal characteristics of the diff-amp.

11.2.1 BJT Diff-Amp Operation—Qualitative Description

Figure 11.2 shows the basic BJT differential-pair configuration. Two identical transistors, Q_1 and Q_2 , whose emitters are connected together, are biased by a constant-current source I_Q , which is connected to a negative supply voltage V^- . The collectors of Q_1 and Q_2 are connected through resistors R_C to a positive supply voltage V^+ . By design, transistors Q_1 and Q_2 are to remain biased in the forward-active region. We assume that the two collector resistors R_C are equal, and that v_{B1} and v_{B2} are ideal sources, meaning that the output resistances of these sources are negligibly small.

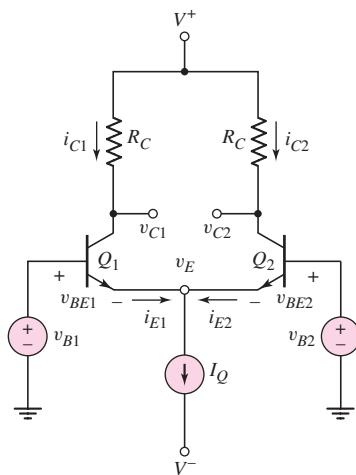


Figure 11.2 Basic BJT differential-pair configuration

Since both positive and negative bias voltages are used in the circuit, the need for coupling capacitors and voltage divider biasing resistors at the inputs of Q_1 and Q_2 has been eliminated. If the input signal voltages v_{B1} and v_{B2} in the circuit shown in Figure 11.2 are both zero, Q_1 and Q_2 are still biased in the active region by the current source I_Q . The common-emitter voltage v_E would be on the order of -0.7 V . This circuit, then, is referred to as a dc-coupled differential amplifier, so differences in dc input voltages can be amplified. Although the diff-amp contains two transistors, it is considered a single-stage amplifier. The analysis will show that it has characteristics similar to those of the common-emitter amplifier.

First, we consider the circuit in which the two base terminals are connected together and a common-mode voltage v_{cm} is applied as shown in Figure 11.3(a). The transistors are biased “on” by the constant-current source, and the voltage at the common emitters is $v_E = v_{cm} - V_{BE}(\text{on})$. Since Q_1 and Q_2 are matched or identical, current I_Q splits evenly between the two transistors, and

$$i_{E1} = i_{E2} = \frac{I_Q}{2} \quad (11.4)$$

If base currents are negligible, then $i_{C1} \cong i_{E1}$ and $i_{C2} \cong i_{E2}$, and

$$v_{C1} = V^+ - \frac{I_Q}{2} R_C = v_{C2} \quad (11.5)$$

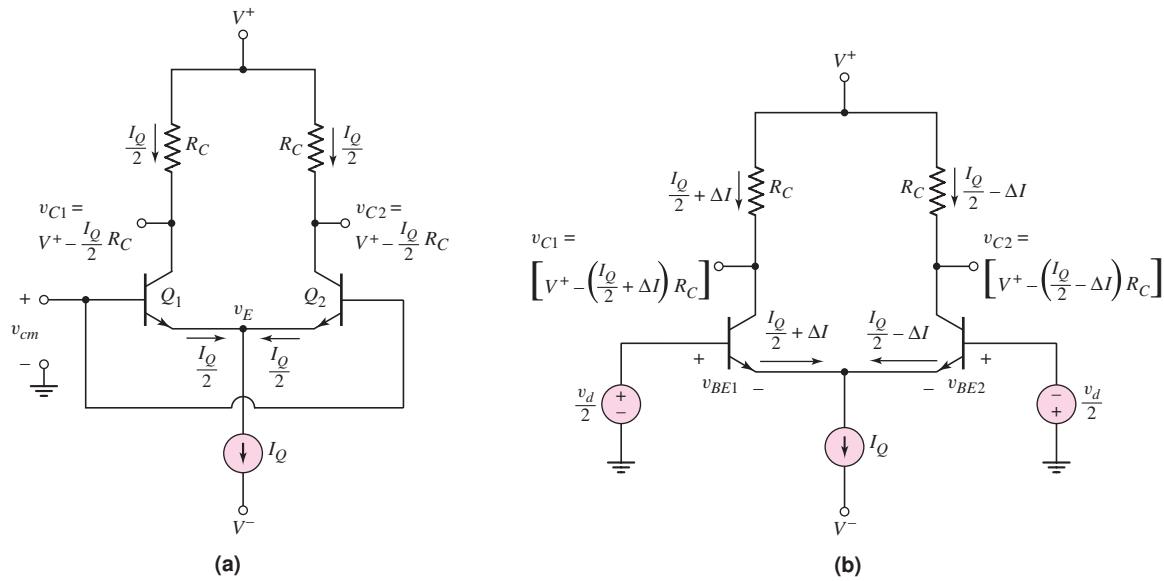


Figure 11.3 Basic diff-amp with applied common-mode voltage and (b) basic diff-amp with applied differential-mode voltage

We see from Equation (11.5) that, for an applied common-mode voltage, I_Q splits evenly between Q_1 and Q_2 and the difference between v_{C1} and v_{C2} is zero.

Now, if v_{B1} increases by a few millivolts and v_{B2} decreases by the same amount, or $v_{B1} = v_d/2$ and $v_{B2} = -v_d/2$, the voltages at the bases of Q_1 and Q_2 are no longer equal. Since the emitters are common, this means that the B-E voltages on Q_1 and Q_2 are no longer equal. Since v_{B1} increases and v_{B2} decreases, then $v_{BE1} > v_{BE2}$, which means that i_{C1} increases by ΔI above its quiescent value and i_{C2} decreases by ΔI below its quiescent value. This is shown in Figure 11.3(b). A potential difference now exists between the two collector terminals. We can write

$$\begin{aligned} v_{C2} - v_{C1} &= \left[V^+ - \left(\frac{I_{CQ}}{2} - \Delta I \right) R_C \right] \\ &\quad - \left[V^+ - \left(\frac{I_{CQ}}{2} + \Delta I \right) R_C \right] = 2\Delta I R_C \end{aligned} \quad (11.6)$$

A voltage difference is created between v_{C2} and v_{C1} when a differential-mode input voltage is applied.

EXAMPLE 11.1

Objective: Determine the quiescent collector current and collector-emitter voltage in a difference amplifier.

Consider the diff-amp in Figure 11.2, with circuit parameters: $V^+ = 10$ V, $V^- = -10$ V, $I_Q = 1$ mA, and $R_C = 10$ k Ω . The transistor parameters are: $\beta = \infty$ (neglect base currents), $V_A = \infty$, and $V_{BE(on)} = 0.7$ V. Determine i_{C1} and v_{CE1} for common-mode voltages $v_{B1} = v_{B2} = v_{CM} = 0$, -5 V, and $+5$ V.

Solution: We know that

$$i_{C1} = i_{C2} = \frac{I_Q}{2} = 0.5 \text{ mA}$$

therefore,

$$v_{C1} = v_{C2} = V^+ - i_{C1}R_C = 10 - (0.5)(10) = 5 \text{ V}$$

From $v_{CM} = 0$, $v_E = -0.7 \text{ V}$ and

$$v_{CE1} = v_{C1} - v_E = 5 - (-0.7) = 5.7 \text{ V}$$

For $v_{CM} = -5 \text{ V}$, $v_E = -5.7 \text{ V}$ and

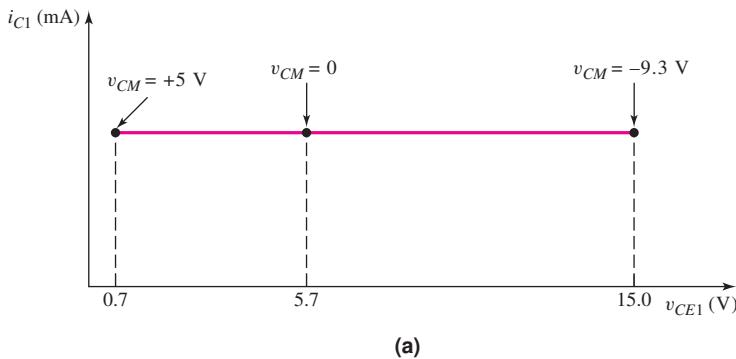
$$v_{CE1} = v_{C1} - v_E = 5 - (-5.7) = 10.7 \text{ V}$$

For $v_{CM} = +5 \text{ V}$, $v_E = 4.3 \text{ V}$ and

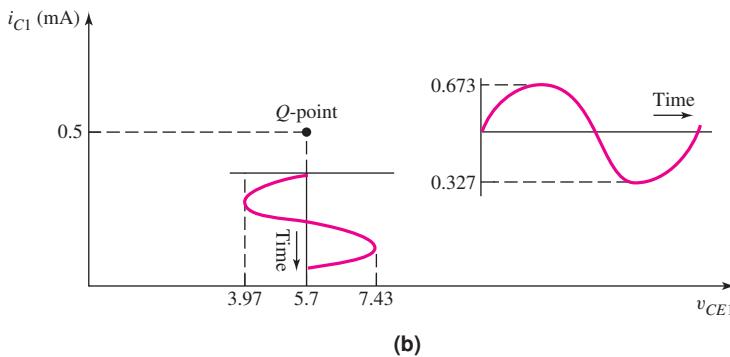
$$v_{CE1} = v_{C1} - v_E = 5 - 4.3 = 0.7 \text{ V}$$

Comment: As the common-mode input voltage varies, the ideal constant current I_Q still splits evenly between Q_1 and Q_2 , but the collector-emitter voltage varies, which means that the Q -point changes. The variation in Q -point as a function of common-mode input voltage is shown in Figure 11.4(a). In this example, if v_{CM} were to increase about $+5 \text{ V}$, then Q_1 and Q_2 would be driven into saturation. This demonstrates that there is a limited range of applied common-mode voltage over which Q_1 and Q_2 will remain biased in the forward-active mode.

Figure 11.4(b) shows the Q -point when $v_{CM} = 0$ and also shows the variation in i_{C1} and v_{CE1} when an 18 mV sinusoidal differential voltage is applied.



(a)



(b)

Figure 11.4 (a) Variation of Q -point for transistor Q_1 in the BJT diff-amp as the common-mode input voltage varies from $+5$ to -9.3 V ; (b) change in collector current and collector-emitter voltage versus time for transistor Q_1 in the BJT diff-amp when a sinusoidal 18 mV differential voltage is applied

EXERCISE PROBLEM

Ex 11.1: The circuit parameters for the differential amplifier shown in Figure 11.2 are $V^+ = 5\text{ V}$, $V^- = -5\text{ V}$, $I_Q = 0.3\text{ mA}$, and $R_C = 20\text{ k}\Omega$. The transistor parameters are $\beta = 180$, $V_{BE}(\text{on}) = 0.7\text{ V}$, and $V_A = \infty$. Determine the voltages v_E , v_{C1} , v_{C2} , v_{CE1} , and v_{CE2} for (a) $v_1 = v_2 = 0$, (b) $v_1 = v_2 = -1\text{ V}$, and (c) $v_1 = v_2 = +1\text{ V}$. (Ans. (a) $v_E = -0.7\text{ V}$, $v_{C1} = v_{C2} = 2\text{ V}$, $v_{CE1} = v_{CE2} = 2.7\text{ V}$; (b) $v_E = -1.7\text{ V}$, $v_{C1} = v_{C2} = 2\text{ V}$, $v_{CE1} = v_{CE2} = 3.7\text{ V}$; (c) $v_E = +0.3\text{ V}$, $v_{C1} = v_{C2} = 2\text{ V}$, $v_{CE1} = v_{CE2} = 1.7\text{ V}$)

11.2.2 DC Transfer Characteristics

We can perform a general analysis of the differential-pair configuration by using the exponential relationship between collector current and B-E voltage. To begin, we know that

$$i_{C1} = I_S e^{v_{BE1}/V_T} \quad (11.7\text{(a)})$$

and

$$i_{C2} = I_S e^{v_{BE2}/V_T} \quad (11.7\text{(b)})$$

We assume Q_1 and Q_2 are matched and are operating at the same temperature, so the coefficient I_S is the same in each expression.

Neglecting base currents and assuming I_Q is an ideal constant-current source, we have

$$I_Q = i_{C1} + i_{C2} \quad (11.8)$$

where i_{C1} and i_{C2} are the total instantaneous currents, which may include the signal currents. We then have

$$I_Q = I_S [e^{v_{BE1}/V_T} + e^{v_{BE2}/V_T}] \quad (11.9)$$

Taking the ratios of i_{C1} to I_Q and i_{C2} to I_Q , we obtain

$$\frac{i_{C1}}{I_Q} = \frac{1}{1 + e^{(v_{BE2}-v_{BE1})/V_T}} \quad (11.10\text{(a)})$$

and

$$\frac{i_{C2}}{I_Q} = \frac{1}{1 + e^{-(v_{BE2}-v_{BE1})/V_T}} \quad (11.10\text{(b)})$$

From Figure 11.3(b) we see that

$$v_{BE1} - v_{BE2} \equiv v_d \quad (11.11)$$

where v_d is the differential-mode input voltage. Equations (11.10(a)) and (11.10(b)) can then be written in terms of v_d , as follows:

$$i_{C1} = \frac{I_Q}{1 + e^{-v_d/V_T}} \quad (11.12\text{(a)})$$

and

$$i_{C2} = \frac{I_Q}{1 + e^{+v_d/V_T}} \quad (11.12\text{(b)})$$

Equations (11.12(a)) and (11.12(b)) describe the basic current-voltage characteristics of the differential amplifier. If the differential-mode input voltage is zero,

then the current I_Q splits evenly between i_{C1} and i_{C2} , as we discussed. However, when a differential-mode signal v_d is applied, a difference occurs between i_{C1} and i_{C2} which in turn causes a change in the collector terminal voltage. This is the fundamental operation of the diff-amp. If a common-mode signal $v_{CM} = v_{B1} = v_{B2}$ is applied, the bias current I_Q still splits evenly between the two transistors.

Figure 11.5 is the normalized plot of the **dc transfer characteristics** for the differential amplifier. We can make two basic observations. First, the gain of the differential amplifier is proportional to the slopes of the transfer curves about the point $v_d = 0$. In order to maintain a linear amplifier, the excursion of v_d about zero must be kept small.

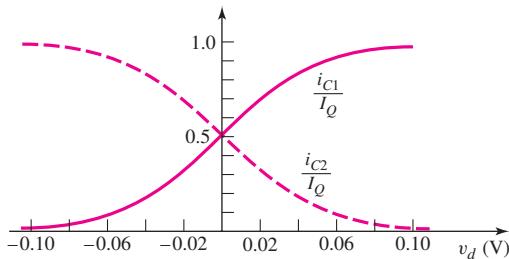


Figure 11.5 Normalized dc transfer characteristics for BJT differential amplifier

Second, as the magnitude of v_d becomes sufficiently large, essentially all of current I_Q goes to one transistor, and the second transistor effectively turns off. This particular characteristic is used in the emitter-coupled logic (ECL) family of digital logic circuits, which is discussed in Chapter 17.

EXAMPLE 11.2

Objective: Determine the maximum differential-mode input signal that can be applied and still maintain linearity in the differential amplifier.

Figure 11.6 shows an expanded view of the normalized i_{C1} versus v_d characteristic. A linear approximation that corresponds to the slope at $v_d = 0$ is superimposed on the curve. Determine $v_d(\max)$ such that the difference between the linear approximation and the actual curve is 1 percent.

Solution: The actual expression for i_{C1} versus v_d is, from Equation (11.12(a)),

$$i_{C1}(\text{actual}) = \frac{I_Q}{1 + e^{-v_d/V_T}}$$

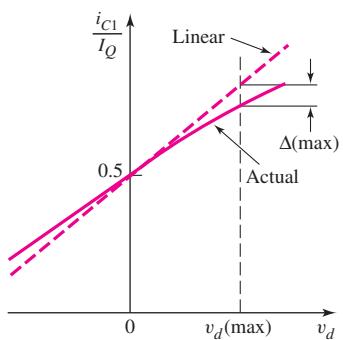


Figure 11.6 Expanded view, normalized i_{C1} versus v_d transfer characteristic

The slope at $v_d = 0$ is found to be

$$g_f = \frac{di_{C1}}{dv_d} \Big|_{v_d=0} = I_Q(-1)[1 + e^{-v_d/V_T}]^{-2} \left(\frac{-1}{V_T} \right) [e^{-v_d/V_T}] \Big|_{v_d=0}$$

or

$$g_f = \frac{I_Q}{4V_T} \quad (11.13)$$

where g_f is the **forward transconductance**. The linear approximation for i_{C1} versus v_d can be written

$$i_{C1}(\text{linear}) = 0.5I_Q + g_f v_d = 0.5I_Q + \left(\frac{I_Q}{4V_T} \right) v_d \quad (11.14)$$

The differential-mode input voltage $v_d(\text{max})$ that results in a 1 percent difference between the ideal linear curve and the actual curve is found from

$$\frac{i_{C1}(\text{linear}) - i_{C1}(\text{actual})}{i_{C1}(\text{linear})} = 0.01$$

or

$$\frac{\left[0.5I_Q + \left(\frac{I_Q}{4V_T} \right) v_d(\text{max}) \right] - \frac{I_Q}{1 + e^{-v_d(\text{max})/V_T}}}{\left[0.5I_Q + \left(\frac{I_Q}{4V_T} \right) v_d(\text{max}) \right]} = 0.01$$

If we rearrange terms, this expression becomes

$$0.99 \left[0.5 + \left(\frac{1}{4V_T} \right) v_d(\text{max}) \right] = \frac{1}{1 + e^{-v_d(\text{max})/V_T}}$$

Assuming $V_T = 26$ mV, and using trial and error, we find that

$$v_d(\text{max}) \cong 18 \text{ mV}$$

Comment: The differential-mode input voltage must be held to within ± 18 mV in order for the output signal of this diff-amp to be within 1 percent of a linear response.

EXERCISE PROBLEM

Ex 11.2: Consider the dc transfer characteristics shown in Figure 11.5. Determine the value of the differential-mode input voltage that results in (a) $i_{C1} = 0.25I_Q$ and (b) $i_{C2} = 0.9I_Q$. (Ans. (a) $v_d = -0.02856$ V, (b) $v_d = -0.05713$ V)

COMPUTER SIMULATION PROBLEM

PS 11.1 Plot the dc transfer characteristics in Figure 11.5 using a computer simulation.

We can now begin to consider the operation of the diff-amp in terms of the small-signal parameters. Figure 11.7 shows the differential-pair configuration with an

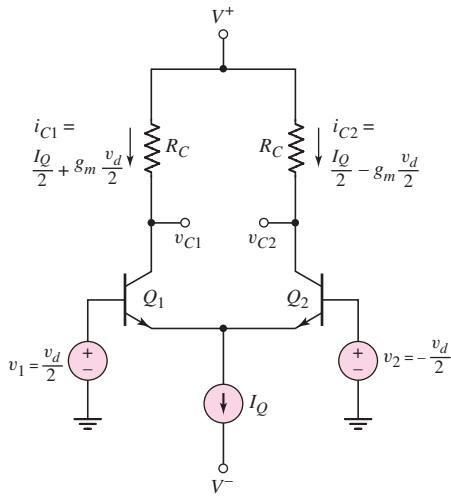


Figure 11.7 BJT differential amplifier with differential-mode input signal

applied differential-mode input signal. Note that the polarity of the input voltage at Q_1 is opposite to that at Q_2 . The forward-transconductance g_f can be written in terms of the individual transistor transconductances g_m . From Equation (11.13), we have

$$g_f = \frac{I_Q}{4V_T} = \frac{1}{2} \frac{I_Q/2}{V_T} = \frac{1}{2} g_m \quad (11.15)$$

where $(I_Q/2)$ is the quiescent collector current in Q_1 and Q_2 . The magnitude of the small-signal collector current in each transistor is then $(g_m v_d)/2$.

Figure 11.7 also shows the linear approximations for the collector currents in terms of the transistor transconductances g_m . The slope of i_{C1} versus v_d is the same magnitude as that of i_{C2} versus v_d , but it has the opposite sign. This is the reason for the negative sign in the expression for i_{C2} versus v_d .

We can define the output signal voltage as

$$v_o = v_{C2} - v_{C1} \quad (11.16)$$

When the output is defined as the difference between the two collector voltages, we have a **two-sided output**. From Figure 11.7, we can write the output voltage as

$$v_o = [V^+ - i_{C2}R_C] - [V^+ - i_{C1}R_C] = (i_{C1} - i_{C2})R_C \quad (11.17(a))$$

or

$$v_o = \left[\left(\frac{I_Q}{2} + \frac{g_m v_d}{2} \right) - \left(\frac{I_Q}{2} - \frac{g_m v_d}{2} \right) \right] R_C = g_m R_C v_d \quad (11.17(b))$$

Figure 11.8 shows the ac equivalent circuit of the diff-amp configuration, as well as the signal voltages and currents as functions of the transistor transconductances g_m . Since we are assuming an ideal current source, the output resistance looking into the current source is infinite (represented by the dashed line). Using the equivalent circuit in Figure 11.8(a), we find the signal output voltage to be

$$v_o = v_{C2} - v_{C1} = \left(\frac{g_m v_d}{2} \right) R_C - \left(\frac{-g_m v_d}{2} \right) R_C = g_m R_C v_d \quad (11.18)$$

which is the same as Equation (11.17(b)).

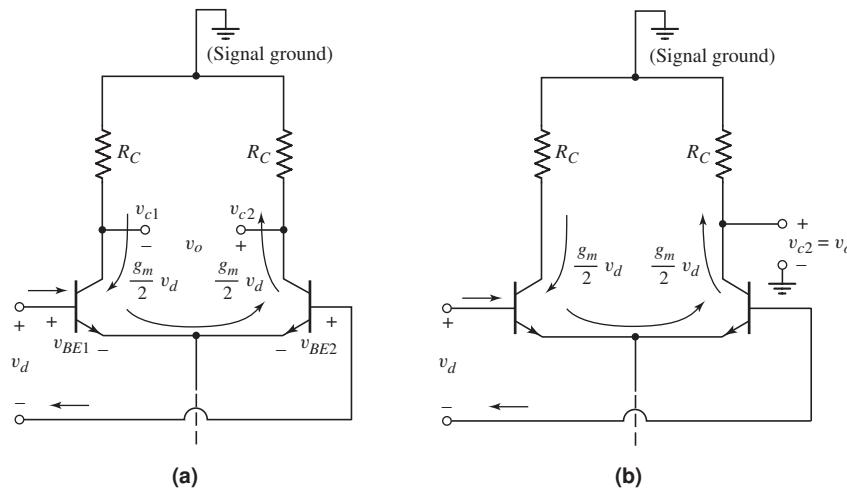


Figure 11.8 (a) Equivalent ac circuit, diff-amp with differential-mode input signal and two-sided output voltage and (b) ac equivalent circuit with one-sided output

The ratio of the output signal voltage to the differential-mode input signal is called the **differential-mode gain**, A_d , which is

$$A_d = \frac{v_o}{v_d} = g_m R_C = \frac{I_Q R_C}{2V_T} \quad (11.19)$$

If the output voltage is the difference between the two collector terminal voltages, then neither side of the output voltage is at ground potential. In many cases, the output voltage is taken at one collector terminal with respect to ground. The resulting voltage output is called a **one-sided output**. If we define the output to be v_{c2} , then from Figure 11.8(b), the signal output voltage is

$$v_o = \left(\frac{g_m v_d}{2} \right) R_C \quad (11.20)$$

The differential gain for the one-sided output is then given by

$$A_d = \frac{v_o}{v_d} = \frac{g_m R_C}{2} = \frac{I_Q R_C}{4V_T} \quad (11.21)$$

The differential gain for the one-sided output is one-half that of the two-sided output. However, as we will see in our discussion on active loads, only a one-sided output is available.

We have assumed that the transistors Q_1 and Q_2 , and the two collector resistors R_C , are matched. The effects of mismatched elements are discussed in the next section.

11.2.3 Small-Signal Equivalent Circuit Analysis

The dc transfer characteristics derived in the last section provide insight into the operation of the differential amplifier. Assuming we are operating in the linear range, we can also derive the gain and other characteristics of the diff-amp, using the small-signal equivalent circuit.

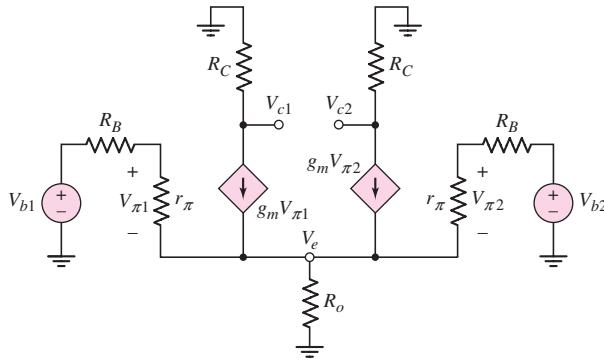


Figure 11.9 Small-signal equivalent circuit, bipolar differential amplifier

Figure 11.9 shows the small-signal equivalent circuit of the bipolar differential-pair configuration. We assume that the Early voltage is infinite for the two emitter-pair transistors, and that the constant-current source is not ideal but can be represented by a finite output impedance R_o . Resistances R_B are also included. These represent the output resistance of the signal voltage sources. All voltages are represented by their phasor components. Since the two transistors are biased at the same quiescent current, we have

$$r_{\pi 1} = r_{\pi 2} \equiv r_{\pi} \quad \text{and} \quad g_{m1} = g_{m2} \equiv g_m$$

Writing a KCL equation at node V_e , using phasor notation, we have

$$\frac{V_{\pi 1}}{r_{\pi}} + g_m V_{\pi 1} + g_m V_{\pi 2} + \frac{V_{\pi 2}}{r_{\pi}} = \frac{V_e}{R_o} \quad (11.22(a))$$

or

$$V_{\pi 1} \left(\frac{1 + \beta}{r_{\pi}} \right) + V_{\pi 2} \left(\frac{1 + \beta}{r_{\pi}} \right) = \frac{V_e}{R_o} \quad (11.22(b))$$

where $g_m r_{\pi} = \beta$. From the circuit, we see that

$$\frac{V_{\pi 1}}{r_{\pi}} = \frac{V_{b1} - V_e}{r_{\pi} + R_B} \quad \text{and} \quad \frac{V_{\pi 2}}{r_{\pi}} = \frac{V_{b2} - V_e}{r_{\pi} + R_B}$$

Solving for $V_{\pi 1}$ and $V_{\pi 2}$ and substituting into Equation (11.22(b)), we find

$$(V_{b1} + V_{b2} - 2V_e) \left(\frac{1 + \beta}{r_{\pi} + R_B} \right) = \frac{V_e}{R_o} \quad (11.23)$$

Solving for V_e , we obtain

$$V_e = \frac{V_{b1} + V_{b2}}{2 + \frac{r_{\pi} + R_B}{(1 + \beta)R_o}} \quad (11.24)$$

One-Sided Output

If we consider a one-sided output at the collector of Q_2 , then

$$V_o = V_{c2} = -(g_m V_{\pi 2}) R_C = -\frac{\beta R_C (V_{b2} - V_e)}{r_{\pi} + R_B} \quad (11.25)$$

Substituting Equation (11.24) into (11.25) and rearranging terms yields

$$V_o = \frac{-\beta R_C}{r_\pi + R_B} \left\{ \frac{V_{b2} \left[1 + \frac{r_\pi + R_B}{(1 + \beta)R_o} \right] - V_{b1}}{2 + \frac{r_\pi + R_B}{(1 + \beta)R_o}} \right\} \quad (11.26)$$

In an ideal constant-current source, the output resistance is $R_o = \infty$, and Equation (11.26) reduces to

$$V_o = -\frac{\beta R_C (V_{b2} - V_{b1})}{2(r_\pi + R_B)} \quad (11.27)$$

The differential-mode input is

$$V_d = V_{b1} - V_{b2}$$

and the differential-mode gain is

$$A_d = \frac{V_o}{V_d} = \frac{\beta R_C}{2(r_\pi + R_B)} \quad (11.28)$$

which for $R_B = 0$ is identical to Equation (11.21), which was developed from the voltage transfer characteristics.

Equation (11.26) includes a finite output resistance for the current source. We can see that when a common-mode signal $V_{cm} = V_{b1} = V_{b2}$ is applied, the output voltage is no longer zero.

Differential- and common-mode voltages are defined in Equations (11.2) and (11.3). Using phasor notation, we can solve these equations for V_{b1} and V_{b2} in terms of V_d and V_{cm} . We obtain

$$V_{b1} = V_{cm} + \frac{V_d}{2} \quad (11.29(a))$$

and

$$V_{b2} = V_{cm} - \frac{V_d}{2} \quad (11.29(b))$$

Since we are dealing with a linear amplifier, superposition applies. Equations (11.29(a)) and (11.29(b)) then simply state that the two input signals can be written as the sum of a differential-mode input signal component and a common-mode input signal component.

Substituting Equations (11.29(a)) and (11.29(b)) into Equation (11.26) and rearranging terms results in the following:

$$V_o = \frac{\beta R_C}{2(r_\pi + R_B)} \cdot V_d - \frac{\beta R_C}{r_\pi + R_B + 2(1 + \beta)R_o} \cdot V_{cm} \quad (11.30)$$

We can write the output voltage in the general form

$$V_o = A_d V_d + A_{cm} V_{cm} \quad (11.31)$$

where A_d is the differential-mode gain and A_{cm} is the common-mode gain. Comparing Equations (11.30) and (11.31), we see that the differential-mode gain is

$$A_d = \frac{\beta R_C}{2(r_\pi + R_B)} \quad (11.32(a))$$

and the common-mode gain is

$$A_{cm} = \frac{-\beta R_C}{r_\pi + R_B + 2(1+\beta)R_o} \quad (11.32(b))$$

We again observe that the common-mode gain goes to zero for an ideal current source in which $R_o = \infty$. For a nonideal current source, R_o is finite and the common-mode gain is not zero for this case of a one-sided output. A nonzero common-mode gain implies that the diff-amp is not ideal.

11.2.4 Common-Mode Rejection Ratio

The ability of a differential amplifier to reject a common-mode signal is described in terms of the common-mode rejection ratio (CMRR). The CMRR is a figure of merit for the diff-amp and is defined as

$$\text{CMRR} = \left| \frac{A_d}{A_{cm}} \right| \quad (11.33)$$

For an ideal diff-amp, $A_{cm} = 0$ and $\text{CMRR} = \infty$. Usually, the CMRR is expressed in decibels, as follows:

$$\text{CMRR}_{\text{dB}} = 20 \log_{10} \left| \frac{A_d}{A_{cm}} \right| \quad (11.34)$$

For the diff-amp in Figure 11.2, the one-sided differential- and common-mode gains are given by Equations (11.32(a)) and (11.32(b)). Using these equations, we can express the CMRR as

$$\text{CMRR} = \left| \frac{A_d}{A_{cm}} \right| = \frac{1}{2} \left[1 + \frac{2(1+\beta)R_o}{r_\pi + R_B} \right] \quad (11.35)$$

The common-mode gain decreases as R_o increases. Therefore, we see that the CMRR increases as R_o increases.

EXAMPLE 11.3

Objective: Determine the differential- and common-mode gains and the common-mode rejection ratio of a diff-amp.

Consider the circuit in Figure 11.2, with parameters: $V^+ = 10$ V, $V^- = -10$ V, $I_Q = 0.8$ mA, and $R_C = 12$ k Ω . The transistor parameters are $\beta = 100$ and $V_A = \infty$. Assume the output resistance looking into the constant-current source is $R_o = 25$ k Ω . Assume the source resistors R_B are zero. Use a one-sided output at v_{C2} .

Solution: From Equation (11.32(a)), the differential-mode gain can be written as

$$A_d = \frac{g_m R_C}{2} = \frac{I_{CQ} R_C}{2V_T} = \frac{I_{CQ} R_C}{4V_T} = \frac{(0.8)(12)}{4(0.026)} = 92.3$$

From Equation (11.32(b)), the common-mode gain can be written as

$$A_{cm} = \frac{-\left(\frac{I_{CQ} R_C}{2V_T}\right)}{1 + \frac{(1+\beta)I_{CQ} R_o}{V_T \beta}} = \frac{-\left[\frac{(0.8)(12)}{(2)(0.026)}\right]}{1 + \frac{(101)(0.8)(25)}{(0.026)(100)}} = -0.237$$

The common-mode rejection ratio is

$$\text{CMRR} = \left| \frac{A_d}{A_{cm}} \right| = \left| \frac{92.3}{-0.237} \right| = 389$$

In many cases, the value of CMRR is expressed in decibels, or

$$\text{CMRR}_{\text{dB}} = 20 \log_{10} \text{CMRR}$$

which, for this example, becomes

$$\text{CMRR}_{\text{dB}} = 20 \log_{10}(389) = 51.8 \text{ dB}$$

Comment: The common-mode gain is less than the differential-mode gain, but is not zero as determined for the ideal diff-amp with an ideal current source. In general, a common-mode rejection ratio of $\text{CMRR}_{\text{dB}} > 80 \text{ dB}$ is a design goal for a diff-amp. The aim, then, is to design a better diff-amp than considered in this example.

EXERCISE PROBLEM

Ex 11.3: Consider the diff-amp described in Example 11.3. Assume the same circuit and transistor parameters except for the current source output resistance R_o . Determine the required value of R_o to produce a CMRR_{dB} of (a) 75 dB and (b) 95 dB. (Ans. (a) $R_o = 362 \text{ k}\Omega$, (b) $R_o = 3.62 \text{ M}\Omega$)

DESIGN EXAMPLE 11.4

Objective: Design a differential amplifier to meet the specifications of an experimental system.

Specifications: Figure 11.10 shows a Hall-effect experiment to measure semiconductor material parameters. A Hall voltage V_H , which is perpendicular to both a current I_X and a magnetic field B_Z , is to be measured by using a diff-amp. The range of V_H is $-8 \leq V_H \leq +8 \text{ mV}$ and the desired range of the diff-amp output signal is to be $-0.8 \leq V_O \leq +0.8 \text{ V}$. The probes that make contact to the semiconductor have an effective resistance of 500Ω , and each probe has an induced 60 Hz signal with a magnitude of 100 mV. The diff-amp output 60 Hz signal is to be no larger than 10 mV. Typically, $V_X = 5 \text{ V}$, so that the quiescent or common-mode voltage of the Hall probes is 2.5 V.

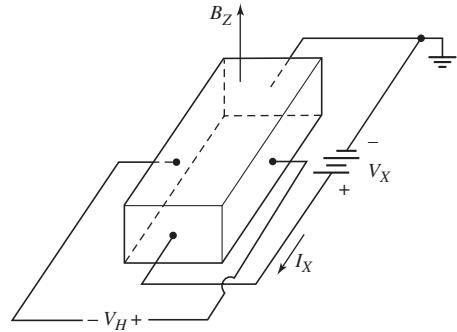


Figure 11.10 Experimental arrangement for measuring Hall voltage

Choices: The bipolar diff-amp with the configuration in Figure 11.7 is to be designed with bias voltages of ± 10 V. The diff-amp transistors are matched with $\beta = 100$ and matched integrated collector resistors of any value can be fabricated. Assume the transistors in the current source are matched with very large β values, $V_{BE(on)} = 0.7$ V, and $V_A = 80$ V. A bias current of $I_Q = 0.5$ mA is to be used.

Solution (Differential-Mode Gain): The differential-mode voltage gain requirement is

$$A_d = \frac{V_o}{V_d} = \frac{0.8}{0.008} = 100$$

The small-signal parameters are then

$$r_\pi = \frac{\beta V_T}{I_{CQ}} = \frac{(100)(0.026)}{0.25} = 10.4 \text{ k}\Omega$$

and

$$g_m = \frac{I_{CQ}}{V_T} = \frac{0.25}{0.026} = 9.62 \text{ mA/V}$$

The differential gain is

$$A_d = \frac{\beta R_C}{2(r_\pi + R_B)}$$

or

$$100 = \frac{(100)R_C}{2(10.4 + 0.5)}$$

which means that $R_C = 21.8$ k Ω . We may note that the voltage drop across R_C under quiescent conditions is 5.45 V. With a 2.5 V common-mode input voltage, the quiescent collector-emitter voltages of Q_1 and Q_2 are approximately 3.65 V. The two input transistors will then remain in the active region.

Solution (Common-Mode Gain): The common-mode voltage gain requirement is

$$A_{cm} = \frac{V_o}{V_{cm}} = \frac{10 \text{ mV}}{100 \text{ mV}} = 0.10$$

The common-mode gain is given by

$$|A_{cm}| = \frac{\beta R_C}{r_\pi + R_B + 2(1 + \beta)R_o}$$

or

$$0.10 = \frac{(100)(21.8)}{10.4 + 0.5 + 2(101)R_o}$$

which means that $R_o = 108$ k Ω . If we consider a simple two-transistor current source as discussed in the last chapter, the output resistance is $R_o = r_o = V_A/I_Q$, where V_A is the Early voltage. With $I_Q = 0.5$ mA, then $V_A = 54$ V is the Early voltage requirement. This specification is not difficult to achieve for most bipolar transistors.

Trade-offs: If the common-mode gain requirement had been more stringent, a different current source circuit might be required to provide a larger output resistance. The effects of mismatched devices and elements are considered in the next section.

Computer Simulation Verification: Figure 11.11 shows the circuit used in the computer simulation for this example. The bias current I_Q supplied by the Q_3 current source transistor is 0.568 mA. A 2.5 V common-mode input voltage is applied, a $500\ \Omega$ source (probe) resistance is included, and an 8 mV differential-mode input signal is applied. The differential output signal voltage measured at the collector of Q_2 is 0.84 V, which is just slightly larger than the designed value. The current gains of the standard 2N3904 transistors used in the computer simulation are larger than the values of 100 used in the hand analysis and design. A common-mode signal voltage of 100 mV replaced the differential-mode signals. The common-mode output signal is 7.11 mV, which is within the design specification.

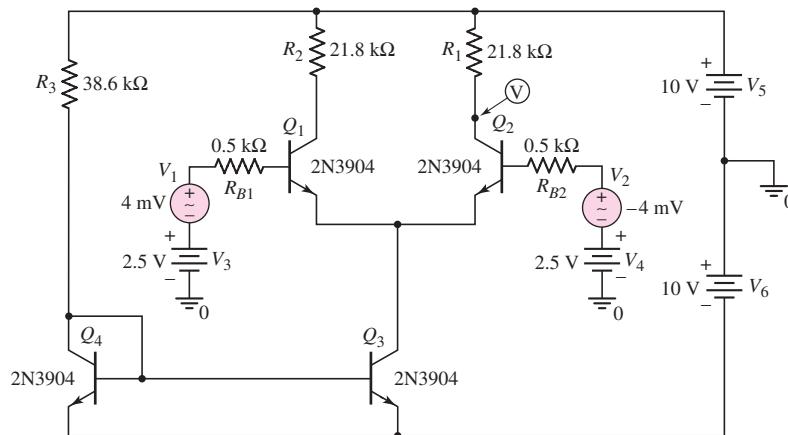


Figure 11.11 Circuit used in the computer simulation of Design Example 11.4

EXERCISE PROBLEM

Ex 11.4: Consider the diff-amp configuration shown in Figure 11.7. Assume Q_1 and Q_2 are matched, let $V_A = \infty$, and neglect base currents. Let $I_Q = 200\ \mu\text{A}$.
 (a) Design the circuit such that the differential-mode gain at v_{C1} is -150 , the differential-mode gain at v_{C2} is $+100$, and the common-mode voltage is in the range $-1.5 \leq v_{cm} \leq 1.5$ V. (b) Using the results of part (a), what are the minimum bias voltages $V^+ = -V^-$ such that the input transistors always remain biased in the forward-active region. (Ans. (a) $R_{C1} = 78.0\ \text{k}\Omega$, $R_{C2} = 52.0\ \text{k}\Omega$; (b) $V^+ = -V^- = 9.3$ V)

Test Your Understanding

TYU 11.1 Find the differential- and common-mode components of the input signal applied to a diff-amp for input voltages of (a) $v_1 = 2.100$ V and $v_2 = 2.120$ V; and (b) $v_1 = 0.25 - 0.002 \sin \omega t$ V and $v_2 = 0.50 + 0.002 \sin \omega t$ V. (Ans. (a) $v_d = -0.02$ V, $v_{cm} = 2.110$ V; (b) $v_d = -0.25 - 0.004 \sin \omega t$ V, $v_{cm} = 0.375$ V)

TYU 11.2 Consider the diff-amp in Figure 11.2, with parameters $V^+ = 5\text{ V}$, $V^- = -5\text{ V}$, and $I_Q = 0.4\text{ mA}$. (a) Redesign the circuit such that the common-mode input voltage is in the range $-3 \leq v_{cm} \leq 3\text{ V}$, while Q_1 and Q_2 remain biased in the forward-active region. (b) Using the results of part (a), find the differential-mode voltage gain $A_d = (v_{c2} - v_{c1})/v_d$. (Ans. (a) $R_C = 10\text{ k}\Omega$, (b) $A_d = 76.9$)

TYU 11.3 Assume the differential-mode gain of a diff-amp is $A_d = 80$ and the common-mode gain is $A_{cm} = -0.20$. Determine the output voltage for input signals of: (a) $v_1 = 0.995 \sin \omega t\text{ V}$ and $v_2 = 1.005 \sin \omega t\text{ V}$; and (b) $v_1 = 2 - 0.005 \sin \omega t\text{ V}$ and $v_2 = 2 + 0.005 \sin \omega t\text{ V}$. (Ans. (a) $v_o = -1.0 \sin \omega t\text{ V}$, (b) $v_o = -0.4 - 0.8 \sin \omega t\text{ V}$)

11.2.5 Two-Sided Output

If we consider the two-sided output of an ideal op-amp and define the output voltage as $V_o = V_{c2} - V_{c1}$, we can show that the differential-mode voltage gain is given by

$$A_d = \frac{\beta R_C}{r_\pi + R_B} \quad (11.36(a))$$

and the common-mode voltage gain is given by

$$A_{cm} = 0 \quad (11.36(b))$$

The result of $A_{cm} = 0$ for the two-sided output is a consequence of using matched devices and elements in the diff-amp circuit. We will reconsider a two-sided output and discuss the effects of mismatched elements.

Effect of R_C Mismatch—Two-Sided Output

We assume that R_{C1} and R_{C2} are the resistors in the collectors of Q_1 and Q_2 . If the two resistors are not matched, we assume that we can write $R_{C1} = R_C + \Delta R_C$ and $R_{C2} = R_C - \Delta R_C$. For simplicity, let $R_B = 0$.

From Figure 11.9, the output voltage for a two-sided output is given by

$$V_o = V_{c2} - V_{c1} = (-g_m V_{\pi 2} R_{C2}) - (-g_m V_{\pi 1} R_{C1}) \quad (11.37)$$

We also see from the figure (with $R_B = 0$) that $V_{\pi 1} = V_{b1} - V_e$ and $V_{\pi 2} = V_{b2} - V_e$. Using the expressions for V_e (Equation (11.24)), V_{b1} (Equation (11.29(a))), and V_{b2} (Equation (11.29(b))), we find the differential voltage gain as

$$A_d = g_m R_C \quad (11.38)$$

and the common-mode gain as

$$A_{cm} = g_m (2\Delta R_C) \cdot \frac{1}{\left[1 + \frac{2(1+\beta)R_o}{r_\pi} \right]} \quad (11.39)$$

In general, $2(1+\beta)R_o/r_\pi \gg 1$, so that

$$A_{cm} \cong g_m (2\Delta R_C) \cdot \frac{r_\pi}{2(1+\beta)R_o} \quad (11.40(a))$$

Noting that $g_m r_\pi = \beta$ and $\beta/(1+\beta) \cong 1$, we have the common-mode gain as

$$A_{cm} \cong \frac{\Delta R_C}{R_o} \quad (11.40(b))$$

The common-mode rejection ratio is then

$$\text{CMRR} = \left| \frac{A_d}{A_{cm}} \right| = \frac{g_m R_o}{(\Delta R_C / R_C)} \quad (11.41)$$

Effect of g_m Mismatch—Two-Sided Output

We can consider the effect of transistor mismatch by considering the effect of a mismatch in the transconductance g_m . We assume g_{m1} and g_{m2} are the transconductance parameters of the two transistors in the diff-amp. We will assume that we can write $g_{m1} = g_m + \Delta g_m$ and $g_{m2} = g_m - \Delta g_m$. Again, for simplicity, let $R_B = 0$.

Again, from Figure 11.9, the output voltage for a two-sided output is

$$V_o = V_{c2} - V_{c1} = (-g_{m2} V_{\pi 2} R_C) - (-g_{m1} V_{\pi 1} R_C) \quad (11.42)$$

Applying a differential input voltage, we find $V_{\pi 1} = V_d/2$ and $V_{\pi 2} = -V_d/2$. The differential voltage gain is then

$$A_d = \frac{V_o}{V_d} = g_m R_C \quad (11.43)$$

Applying a common-mode input voltage, we have $V_{\pi 1} = V_{\pi 2} = V_{cm} - V_e$. The output voltage is again given by

$$V_o = V_{c2} - V_{c1} = (-g_{m2} V_{\pi 2} R_C) - (-g_{m1} V_{\pi 1} R_C) \quad (11.44(a))$$

or

$$V_o = (V_{cm} - V_e) R_C (g_{m1} - g_{m2}) \quad (11.44(b))$$

Summing currents at the V_e node in Figure 11.9, we have

$$\frac{V_{\pi 1}}{r_{\pi 1}} + g_{m1} V_{\pi 1} + g_{m2} V_{\pi 2} + \frac{V_{\pi 2}}{r_{\pi 2}} = \frac{V_e}{R_o} \quad (11.45)$$

In general, we have $g_m \gg 1/r_\pi$. Then Equation (11.45) becomes

$$(V_{cm} - V_e)(g_{m1} + g_{m2}) = \frac{V_e}{R_o} \quad (11.46(a))$$

or

$$V_e = \frac{V_{cm}(g_{m1} + g_{m2})}{\frac{1}{R_o} + g_{m1} + g_{m2}} \quad (11.46(b))$$

The output voltage is then

$$V_o = \left[V_{cm} - \frac{V_{cm}(g_{m1} + g_{m2})}{(1/R_o) + g_{m1} + g_{m2}} \right] \cdot R_C (g_{m1} - g_{m2}) \quad (11.47)$$

Noting that $g_{m1} + g_{m2} = 2g_m$ and $g_{m1} - g_{m2} = 2(\Delta g_m)$, the common-mode gain is

$$A_{cm} = \frac{R_C (2\Delta g_m)}{1 + 2R_o g_m} \quad (11.48)$$

The common-mode rejection ratio now becomes

$$\text{CMRR} = \left| \frac{A_d}{A_{cm}} \right| = \frac{1 + 2R_o g_m}{2(\Delta g_m / g_m)} \quad (11.49)$$

11.2.6 Differential- and Common-Mode Gains—Further Observations

For greater insight into the mechanism that causes differential- and common-mode gains, we reconsider the diff-amp as pure differential- and common-mode signals are applied.

Figure 11.12(a) shows the ac equivalent circuit of the diff-amp with two sinusoidal input signals. The two input voltages are 180 degrees out of phase, so a pure differential-mode signal is being applied to the diff-amp. We see that $v_{b1} + v_{b2} = 0$. From Equation (11.24), we find $v_e = 0$, so the common emitters of Q_1 and Q_2 remain at signal ground. In essence, the circuit behaves like a balanced seesaw. As the base voltage of Q_1 goes into its positive-half cycle, the base voltage of Q_2 is in its negative half-cycle. Then, as the base voltage of Q_1 goes into its negative half-cycle, the base voltage of Q_2 is in its positive half-cycle. The signal current directions shown in the figure are valid for v_{b1} in its positive half-cycle.

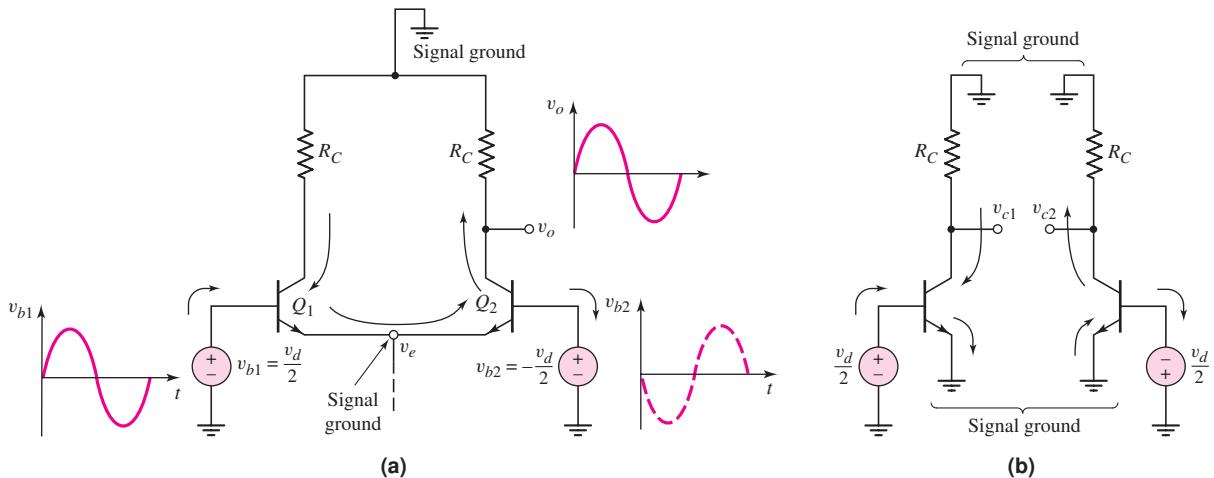


Figure 11.12 (a) Equivalent ac circuit, diff-amp with applied sinusoidal differential-mode input signal, and resulting signal current directions and (b) differential-mode half-circuits

Since v_e is always at ground potential, we can treat each half of the diff-amp as a common-emitter circuit. Figure 11.12(b) shows the differential half-circuits, clearly depicting the common-emitter configuration. The differential-mode characteristics of the diff-amp can be determined by analyzing the half-circuit. In evaluating the small-signal hybrid- π parameters, we must keep in mind that the half-circuit is biased at $I_Q/2$.

Figure 11.13(a) shows the ac equivalent circuit of the diff-amp with a pure common-mode sinusoidal input signal. In this case, the two input voltages are in phase. The current source is represented as an ideal source I_Q in parallel with its output resistance R_o . Current i_q is the time-varying component of the source current. As the two input signals increase, voltage v_e increases and current i_q increases. Since this current splits evenly between Q_1 and Q_2 , each collector current also increases. The output voltage v_o then decreases below its quiescent value.

As the two input voltages go through the negative half-cycle, all signal currents shown in the figure reverse direction, and v_o increases above its quiescent value.

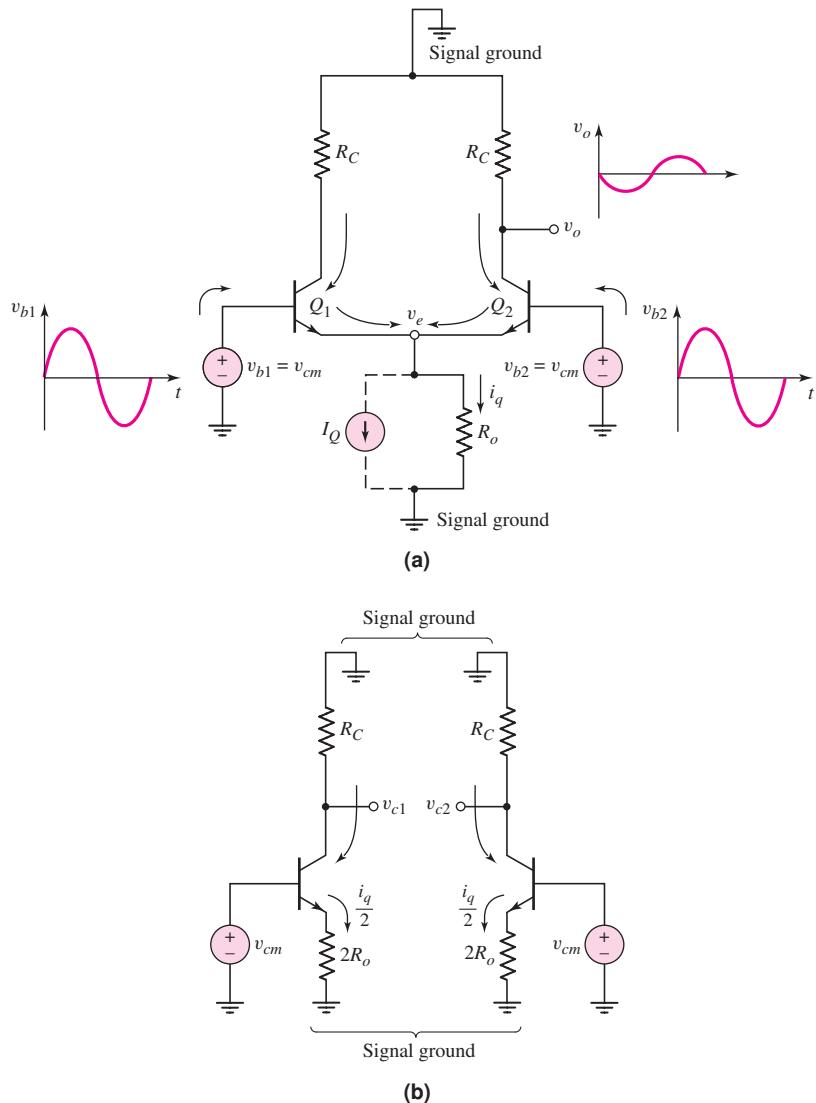


Figure 11.13 (a) Equivalent ac circuit of diff-amp with common-mode input signal, and resulting signal current directions and (b) common-mode half-circuits

Consequently, a common-mode sinusoidal input signal produces a sinusoidal output voltage, which means that the diff-amp has a nonzero common-mode voltage gain. If the value of R_o increases, the magnitude of i_q decreases for a given common-mode input signal, producing a smaller output voltage and hence a smaller common-mode gain.

With an applied common-mode voltage, the circuit shown in Figure 11.13(a) is perfectly symmetrical. The circuit can therefore be split into the identical common-mode half-circuits shown in Figure 11.13(b). The common-mode characteristics of the diff-amp can then be determined by analyzing the half-circuit, which is a common-emitter configuration with an emitter resistor. Each half-circuit is biased at $I_Q/2$.

The following examples further illustrate the effect of a nonzero common-mode gain on circuit performance.

EXAMPLE 11.5

Objective: Determine the output of a diff-amp when both differential- and common-mode signals are applied.

Consider the circuit shown in Figure 11.2. Use the transistor and circuit parameters described in Example 11.3. Assume that four sets of inputs are applied, as described in the following table, which also includes the differential- and common-mode voltages.

Input signal (μV)	Differential- and common-mode input signals (μV)	
Case 1 $v_1 = 10 \sin \omega t$ $v_2 = -10 \sin \omega t$	$v_d = 20 \sin \omega t$	$v_{cm} = 0$
Case 2 $v_1 = 20 \sin \omega t$ $v_2 = -20 \sin \omega t$	$v_d = 40 \sin \omega t$	$v_{cm} = 0$
Case 3 $v_1 = 210 \sin \omega t$ $v_2 = 190 \sin \omega t$	$v_d = 20 \sin \omega t$	$v_{cm} = 200 \sin \omega t$
Case 4 $v_1 = 220 \sin \omega t$ $v_2 = 180 \sin \omega t$	$v_d = 40 \sin \omega t$	$v_{cm} = 200 \sin \omega t$
Case 5 $v_1 = 200 \sin \omega t$ $v_2 = 200 \sin \omega t$	$v_d = 0$	$v_{cm} = 200 \sin \omega t$

Solution: The output voltage is given by Equation (11.31), as follows:

$$v_o = A_d v_d + A_{cm} v_{cm}$$

From Example 11.3, the differential- and common-mode gains are $A_d = 92.3$ and $A_{cm} = -0.237$. The output voltages for the four sets of inputs are:

Output signal (mV)
Case 1 $v_o = 1.846 \sin \omega t$
Case 2 $v_o = 3.692 \sin \omega t$
Case 3 $v_o = 1.799 \sin \omega t$
Case 4 $v_o = 3.645 \sin \omega t$
Case 5 $v_o = -0.0474 \sin \omega t$

Comment: In cases 1 and 2, the common-mode input is zero, and the output is directly proportional to the differential input signal. Comparing cases 1 and 3 and cases 2 and 4, we see that the output voltages are not equal, even though the differential input signals are the same. This shows that the common-mode signal affects the output. Also, even though the differential signal is doubled, in cases 4 and 3, the ratio of the output signals is not 2.0. If a common-mode signal is present, the output is not exactly linear with respect to the differential input signal.

For Case 5, the differential-input voltage is zero, but the output voltage is not zero, since a common-mode input voltage exists and $|A_{cm}| \neq 0$.

EXERCISE PROBLEM

Ex 11.5: Assume a diff-amp has a differential-mode gain of $A_d = 150$ and a common-mode rejection ratio of $\text{CMMR}_{\text{dB}} = 50 \text{ dB}$. Assume A_{cm} is positive. Determine the output voltage if the input voltages are (a) $v_1 = -10 \mu\text{V}$, $v_2 = +10 \mu\text{V}$ and (b) $v_1 = 190 \mu\text{V}$, $v_2 = 210 \mu\text{V}$. (Ans. (a) $v_o = -3.0 \text{ mV}$, (b) $v_o = -2.905 \text{ mV}$)

As mentioned previously, the common-mode gain is a function of the output resistance of the current source. If the required common-mode gain needs to be reduced, then the current source output resistance must be increased, which may require the design of a more sophisticated current source.

Problem-Solving Technique: Diff-Amps with Resistive Loads

1. To determine the differential-mode voltage gain, apply a pure differential-mode input voltage and use the differential-mode half-circuit in the analysis.
2. To determine the common-mode voltage gain, apply a pure common-mode input voltage and use the common-mode half-circuit in the analysis.

11.2.7 Differential- and Common-Mode Input Impedances

The input impedance, or resistance, of an amplifier is as important a property as the voltage gain. The input resistance determines the loading effect of the circuit on the signal source. We will look at two input resistances for the difference amplifier: the **differential-mode input resistance**, which is the resistance seen by a differential-mode signal source; and the **common-mode input resistance**, which is the resistance seen by a common-mode input signal source.

Differential-Mode Input Resistance

The differential-mode input resistance is the effective resistance between the two input base terminals when a differential-mode signal is applied. A diff-amp with a pure differential input signal is shown in Figure 11.14. The applicable differential-mode half-circuits were shown in Figure 11.12(b). For this circuit, we have

$$\frac{v_d/2}{i_b} = r_\pi \quad (11.50)$$

The differential-mode input resistance is therefore

$$R_{id} = \frac{v_d}{i_b} = 2r_\pi \quad (11.51)$$

Another common diff-amp configuration uses emitter resistors, as shown in Figure 11.15. With a pure applied differential-mode voltage, similar differential-mode half-circuits are applicable to this configuration. We can then use the resistance reflection rule to find the differential-mode input resistance. We have

$$\frac{v_d/2}{i_b} = r_\pi + (1 + \beta)R_E \quad (11.52)$$

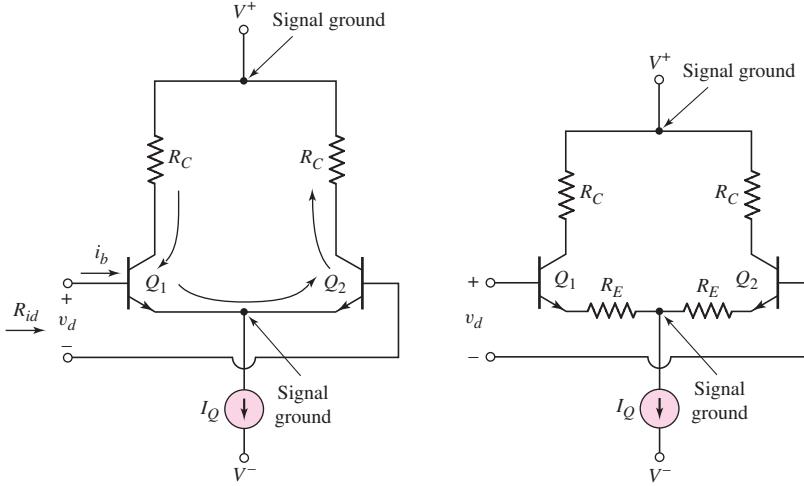


Figure 11.14 BJT differential amplifier with differential-mode input signal, showing differential input resistance

Figure 11.15 BJT differential amplifier with emitter resistors

Therefore,

$$R_{id} = \frac{v_d}{i_b} = 2[r_\pi + (1 + \beta)R_E] \quad (11.53)$$

Equation (11.53) implies that the differential-mode input resistance increases significantly when emitter resistors are included. We will see that the differential-mode gain decreases when emitter resistors are included in the same way that the voltage gain of a common-emitter amplifier decreases when an emitter resistor is included in the design. However, a larger differential-mode voltage (greater than 18 mV) may be applied to the diff-amp in Figure 11.15 and the amplifier remains linear.

Common-Mode Input Resistance

Figure 11.16(a) shows a diff-amp with an applied common-mode voltage. The small-signal output resistance R_o of the constant-current source is also shown. The equivalent common-mode half-circuits were given in Figure 11.13(b). Since the half-circuits are in parallel, we can write

$$2R_{icm} = r_\pi + (1 + \beta)(2R_o) \cong (1 + \beta)(2R_o) \quad (11.54)$$

Equation (11.54) is a first approximation for determining the common-mode input resistance.

Normally, R_o is large, and R_{icm} is typically in the megohm range. Therefore, the transistor output resistance r_o and the base-collector resistance r_μ may need to be included in the calculation. Figure 11.16(b) shows the more complete equivalent half-circuit model. For this model, we have

$$2R_{icm} = r_\mu \parallel [(1 + \beta)(2R_o)] \parallel [(1 + \beta)r_o] \quad (11.55(a))$$

Therefore,

$$R_{icm} = \left(\frac{r_\mu}{2} \right) \parallel [(1 + \beta)(R_o)] \parallel \left[(1 + \beta) \left(\frac{r_o}{2} \right) \right] \quad (11.55(b))$$

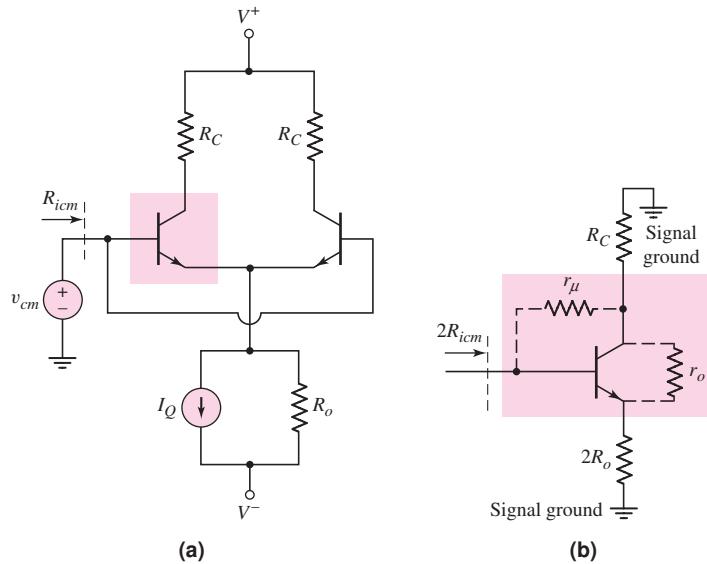


Figure 11.16 (a) BJT differential amplifier with common-mode input signal, including finite current source resistance and (b) equivalent common-mode half-circuit

EXAMPLE 11.6

Objective: Determine the differential- and common-mode input resistances of a differential amplifier.

Consider the circuit in Figure 11.17, with transistor parameters $\beta = 100$, $V_{BE(on)} = 0.7$ V, and $V_A = 100$ V. Determine R_{id} and R_{icm} .

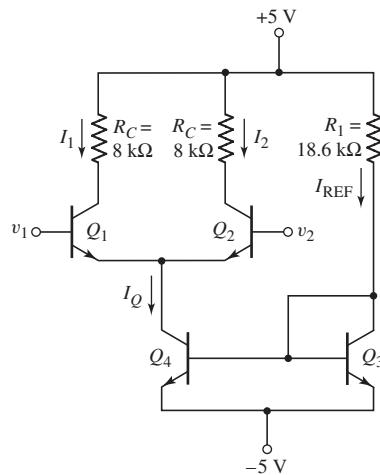


Figure 11.17 BJT differential amplifier for Example 11.6

Solution: From the circuit, we find

$$I_{REF} = 0.5 \text{ mA} \cong I_Q$$

and

$$I_1 = I_2 \cong I_Q/2 = 0.25 \text{ mA}$$

The small-signal parameters for Q_1 and Q_2 are then

$$r_\pi = \frac{\beta V_T}{I_{CQ}} = \frac{(100)(0.026)}{0.25} = 10.4 \text{ k}\Omega$$

and

$$r_o = \frac{V_A}{I_{CQ}} = \frac{100}{0.25} = 400 \text{ k}\Omega$$

and the output resistance of Q_4 is

$$R_o = \frac{V_A}{I_Q} = \frac{100}{0.5} = 200 \text{ k}\Omega$$

From Equation (11.51), the differential-mode input resistance is

$$R_{id} = 2r_\pi = 2(10.4) = 20.8 \text{ k}\Omega$$

From Equation (11.55(b)), neglecting the effect of r_μ , the common-mode input resistance is

$$R_{icm} = (1 + \beta) \left[(R_o) \left\| \left(\frac{r_o}{2} \right) \right\| \right] = (101) \left\{ 200 \left\| \left(\frac{400}{2} \right) \right\| \right\} \text{k}\Omega \rightarrow 10.1 \text{ M}\Omega$$

Comment: If a differential-mode input voltage with a peak value of 15 mV is applied, the source must be capable of supplying a current of $15 \times 10^{-3}/20.8 \times 10^{+3} = 0.72 \mu\text{A}$ without any severe loading effect. However, the input current from a 15 mV common-mode signal would only be approximately 1.5 nA.

EXERCISE PROBLEM

Ex 11.6: Consider the diff-amp shown in Figure 11.15. Assume the current source has a value of $I_Q = 0.5 \text{ mA}$, the transistor current gains are $\beta = 100$, and the emitter resistors are $R_E = 500 \Omega$. Find the differential input resistance.
(Ans. $R_{id} = 122 \text{ k}\Omega$)

Differential-Mode Voltage Gain with Emitter Degeneration

We may determine the differential-mode voltage gain of the circuit shown in Figure 11.15. Figure 11.18 shows the differential-mode half circuits. For a one-sided output and for matched elements, we have

$$V_o = V_{c2} = -g_m V_{\pi2} R_C \quad (11.56)$$

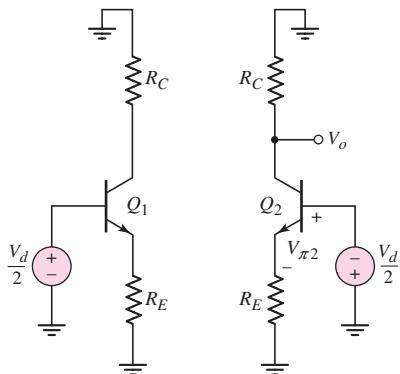


Figure 11.18 Differential half-circuits with emitter degeneration

Writing a KVL equation around the B-E loop, we have

$$\frac{V_d}{2} + V_{\pi 2} + g_m V_{\pi 2} R_E = 0 \quad (11.57)$$

which yields

$$V_{\pi 2} = \frac{-(V_d/2)}{1 + g_m R_E} \quad (11.58)$$

Substituting Equation (11.58) into (11.56), we find the differential-mode voltage gain as

$$A_d = \frac{V_o}{V_d} = \frac{g_m R_C}{2(1 + g_m R_E)} \quad (11.59)$$

EXAMPLE 11.7

Objective: Determine the one-sided differential-mode voltage gain of the circuit shown in Figure 11.15.

Assume $I_Q = 0.5$ mA, $\beta = 100$, and $R_C = 10$ k Ω . Find the differential-mode voltage gain for (a) $R_E = 0$ and (b) $R_E = 500$ Ω .

Solution: The small-signal transconductance is found to be $g_m = 9.62$ mA/V. We find the differential-mode voltage gain to be (a) for $R_E = 0$:

$$A_d = \frac{g_m R_C}{2} = \frac{(9.62)(10)}{2} = 48.1$$

and (b) for $R_E = 500$ Ω :

$$A_d = \frac{g_m R_C}{2(1 + g_m R_E)} = \frac{(9.62)(10)}{2[1 + (9.62)(0.5)]} = 8.28$$

Comment: As with any design problem, there are trade-offs. Including an emitter resistor R_E decreases the voltage gain but increases the input differential-mode resistance.

EXERCISE PROBLEM

Ex 11.7: Consider the diff-amp described in Example 11.7. Assume the same parameters except the value of R_E . Determine the value of R_E that results in a differential-mode voltage gain of $A_d = 10$. What is the corresponding value of differential-input resistance? (Ans. $R_E = 0.396$ k Ω , $R_{id} = 100.8$ k Ω)

Test Your Understanding

TYU 11.4 Consider the effect of a mismatch in collector resistors. Assume that $g_m = 3.86$ mA/V², $R_o = 100$ k Ω , and a nominal collector resistor of $R_C = 10$ k Ω . Determine the minimum mismatch in the collector resistor ΔR_C such that the common-mode rejection ratio is 75 dB. (Ans. $\Delta R_C = 0.686$ k Ω)

TYU 11.5 Consider the effect of a mismatch in the transconductance of the transistors. Assume $R_o = 100$ k Ω and the nominal transconductance is $g_m = 3.86$ mA/V. Determine the minimum mismatch in the transconductance Δg_m such that the common-mode rejection ratio is 90 dB. (Ans. $\Delta g_m = 0.0472$ mA/V or $\Delta g_m/g_m \rightarrow 1.22\%$)

TYU 11.6 The parameters of the diff-amp shown in Figure 11.2 are $V^+ = 5\text{ V}$, $V^- = -5\text{ V}$, $I_Q = 0.4\text{ mA}$, and $R_C = 10\text{ k}\Omega$. The output resistance of the constant-current source is $R_o = 100\text{ k}\Omega$. The transistor parameters are $\beta = 150$, $V_{BE(\text{on})} = 0.7\text{ V}$, and $V_A = \infty$. (a) Determine the dc input base currents. (b) Determine the differential signal input currents if a differential-mode input voltage $v_d = 10 \sin \omega t\text{ mV}$ is applied. (c) If a common-mode input voltage $v_{cm} = 3 \sin \omega t\text{ V}$ is applied, determine the common-mode signal input base currents. (Ans. (a) $I_{B1} = I_{B2} = 1.32\text{ }\mu\text{A}$, (b) $I_b = 0.256 \sin \omega t\text{ }\mu\text{A}$, (c) $I_b = 0.0993 \sin \omega t\text{ }\mu\text{A}$)

P 11.3 BASIC FET DIFFERENTIAL PAIR

Objective: • Describe the characteristics of and analyze the basic FET differential amplifier.

In this section, we will evaluate the basic FET differential amplifier, concentrating on the MOSFET diff-amp. As we did for the bipolar diff-amp, we will develop the dc transfer characteristics, and determine the differential- and common-mode gains.

Differential amplifiers using JFETs are also available. Since the analysis is almost identical to that for the MOSFET diff-amp, we will only briefly consider the JFET differential pair. A few of the problems at the end of this chapter are based on these circuits.

11.3.1 DC Transfer Characteristics

Figure 11.19 shows the basic MOSFET differential pair, with matched transistors M_1 and M_2 biased with a constant current I_Q . We assume that M_1 and M_2 are always biased in the saturation region.

Like the basic bipolar configuration, the basic MOSFET diff-amp uses both positive and negative bias voltages, thereby eliminating the need for coupling capacitors and voltage divider biasing resistors at the gate terminals. Even with $v_{G1} = v_{G2} = 0$, the transistors M_1 and M_2 can be biased in the saturation region by the current source I_Q . This circuit, then, is also a **dc-coupled** diff-amp.

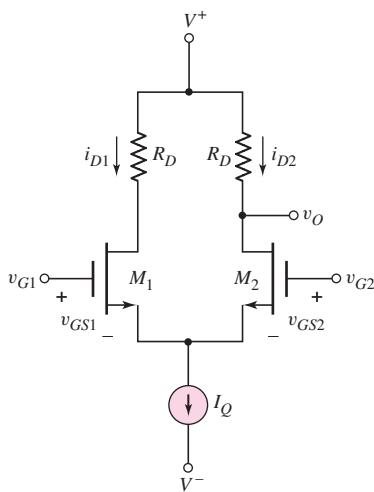


Figure 11.19 Basic MOSFET differential pair configuration

EXAMPLE 11.8

Objective: Calculate the dc characteristics of a MOSFET diff-amp.

Consider the differential amplifier shown in Figure 11.20. The transistor parameters are: $K_{n1} = K_{n2} = 0.1 \text{ mA/V}^2$, $K_{n3} = K_{n4} = 0.3 \text{ mA/V}^2$, and for all transistors, $\lambda = 0$ and $V_{TN} = 1 \text{ V}$. Determine the maximum range of common-mode input voltage.

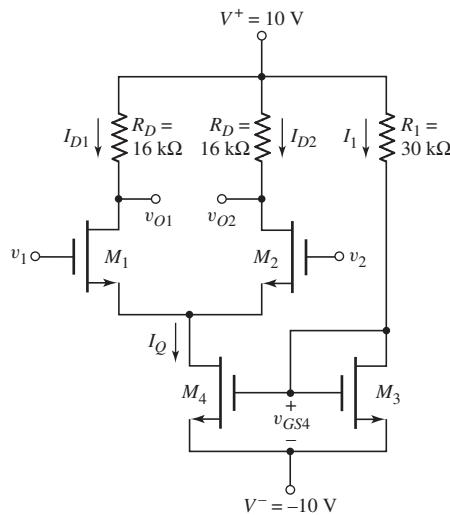


Figure 11.20 MOSFET differential amplifier for Example 11.8

Solution: The reference current can be determined from

$$I_1 = \frac{20 - V_{GS4}}{R_1}$$

and from

$$I_1 = K_{n3}(V_{GS4} - V_{TN})^2$$

Combining these two equations and substituting the parameter values, we obtain

$$9V_{GS4}^2 - 17V_{GS4} - 11 = 0$$

which yields

$$V_{GS4} = 2.40 \text{ V} \quad \text{and} \quad I_1 = 0.587 \text{ mA}$$

Since M_3 and M_4 are identical, we also find

$$I_Q = 0.587 \text{ mA}$$

The quiescent drain currents in M_1 and M_2 are

$$I_{D1} = I_{D2} = I_Q/2 \cong 0.293 \text{ mA}$$

The gate-to-source voltages are then

$$V_{GS1} = V_{GS2} = \sqrt{\frac{I_{D1}}{K_{n1}}} + V_{TN} = \sqrt{\frac{0.293}{0.1}} + 1 = 2.71 \text{ V}$$

The quiescent values of v_{O1} and v_{O2} are

$$v_{O1} = v_{O2} = 10 - I_{D1}R_D = 10 - (0.293)(16) = 5.31 \text{ V}$$

The maximum common-mode input voltage is the value when M_1 and M_2 reach the transition point, or

$$V_{DS1} = V_{DS2} = V_{DS1}(\text{sat}) = V_{GS1} - V_{TN} = 2.71 - 1 = 1.71 \text{ V}$$

Therefore,

$$v_{CM}(\text{max}) = v_{O1} - V_{DS1}(\text{sat}) + V_{GS1} = 5.31 - 1.71 + 2.71$$

or

$$v_{CM}(\text{max}) = 6.31 \text{ V}$$

The minimum common-mode input voltage is the value when M_4 reaches the transition point, or

$$V_{DS4} = V_{DS4}(\text{sat}) = V_{GS4} - V_{TN} = 2.4 - 1 = 1.4 \text{ V}$$

Therefore,

$$v_{CM}(\text{min}) = V_{GS1} + V_{DS4}(\text{sat}) - 10 = 2.71 + 1.4 - 10$$

or

$$v_{CM}(\text{min}) = -5.89 \text{ V}$$

Comment: For this circuit the maximum range for the common-mode input voltage is $-5.89 \leq v_{CM} \leq 6.31 \text{ V}$.

EXERCISE PROBLEM

***Ex 11.8:** For the differential amplifier in Figure 11.20, the parameters are: $V^+ = 5 \text{ V}$, $V^- = -5 \text{ V}$, $R_1 = 80 \text{ k}\Omega$, and $R_D = 40 \text{ k}\Omega$. The transistor parameters are $\lambda = 0$ and $V_{TN} = 0.8 \text{ V}$ for all transistors, and $K_{n3} = K_{n4} = 100 \mu\text{A/V}^2$ and $K_{n1} = K_{n2} = 50 \mu\text{A/V}^2$. Determine the range of the common-mode input voltage. (Ans. $-2.18 \leq v_{cm} \leq 3.76 \text{ V}$)

The dc transfer characteristics of the MOSFET differential pair can be determined from the circuit in Figure 11.19. Neglecting the output resistances of M_1 and M_2 , and assuming the two transistors are matched, we can write

$$i_{D1} = K_n(v_{GS1} - V_{TN})^2 \quad (11.60(a))$$

and

$$i_{D2} = K_n(v_{GS2} - V_{TN})^2 \quad (11.60(b))$$

Taking the square roots of Equations (11.60(a)) and (11.60(b)), and subtracting the two equations, we obtain

$$\sqrt{i_{D1}} - \sqrt{i_{D2}} = \sqrt{K_n}(v_{GS1} - v_{GS2}) = \sqrt{K_n} \cdot v_d \quad (11.61)$$

where $v_d = v_{G1} - v_{G2} = v_{GS1} - v_{GS2}$ is the differential-mode input voltage. If $v_d > 0$, then $v_{G1} > v_{G2}$ and $v_{GS1} > v_{GS2}$, which implies that $i_{D1} > i_{D2}$. Since

$$i_{D1} + i_{D2} = I_Q \quad (11.62)$$

then Equation (11.61) becomes

$$(\sqrt{i_{D1}} - \sqrt{I_Q - i_{D1}})^2 = (\sqrt{K_n} \cdot v_d)^2 = K_n v_d^2 \quad (11.63)$$

when both sides of the equation are squared. After the terms are rearranged, Equation (11.63) becomes

$$\sqrt{i_{D1}(I_Q - i_{D1})} = \frac{1}{2}(I_Q - K_n v_d^2) \quad (11.64)$$

If we square both sides of this equation, we develop the quadratic equation

$$i_{D1}^2 - I_Q i_{D1} + \frac{1}{4}(I_Q - K_n v_d^2)^2 = 0 \quad (11.65)$$

Applying the quadratic formula, rearranging terms, and noting that $i_{D1} > I_Q/2$ and $v_d > 0$, we obtain

$$i_{D1} = \frac{I_Q}{2} + \sqrt{\frac{K_n I_Q}{2}} \cdot v_d \sqrt{1 - \left(\frac{K_n}{2 I_Q}\right) v_d^2} \quad (11.66)$$

Using Equation (11.62), we find that

$$i_{D2} = \frac{I_Q}{2} - \sqrt{\frac{K_n I_Q}{2}} \cdot v_d \sqrt{1 - \left(\frac{K_n}{2 I_Q}\right) v_d^2} \quad (11.67)$$

The normalized drain currents are

$$\frac{i_{D1}}{I_Q} = \frac{1}{2} + \sqrt{\frac{K_n}{2 I_Q}} \cdot v_d \sqrt{1 - \left(\frac{K_n}{2 I_Q}\right) v_d^2} \quad (11.68)$$

and

$$\frac{i_{D2}}{I_Q} = \frac{1}{2} - \sqrt{\frac{K_n}{2 I_Q}} \cdot v_d \sqrt{1 - \left(\frac{K_n}{2 I_Q}\right) v_d^2} \quad (11.69)$$

These equations describe the dc transfer characteristics for this circuit. They are plotted in Figure 11.21 as a function of a normalized differential input voltage $v_d / \sqrt{(2 I_Q / K_n)}$.

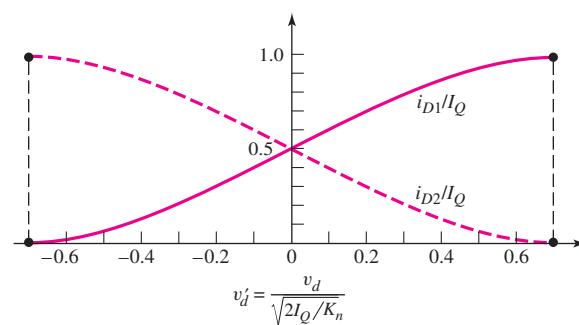


Figure 11.21 Normalized dc transfer characteristics, MOSFET differential amplifier

We can see from Equations (11.68) and (11.69) that, at a specific differential input voltage, bias current I_Q is switched entirely to one transistor or the other. This occurs when

$$|v_d|_{\max} = \sqrt{\frac{I_Q}{K_n}} \quad (11.70)$$

The forward transconductance is defined as the slope of the i_{D1} versus v_d transfer characteristic evaluated at $v_d = 0$, or

$$g_f(\max) = \left. \frac{di_{D1}}{dv_d} \right|_{v_d=0} \quad (11.71)$$

Using Equation (11.66), we find that

$$g_f(\max) = \sqrt{\frac{K_n I_Q}{2}} = \frac{g_m}{2} \quad (11.72)$$

where g_m is the transconductance of each transistor. The slope of the i_{D2} characteristic curve at $v_d = 0$ is the same, except it is negative.

We can perform an analysis similar to that in Example 11.2 to determine the maximum differential-mode input signal that can be applied and still maintain linearity. If we let $I_Q = 1$ mA and $K_n = 1$ mA/V², then for differential input voltages less than 0.34 V, the difference between the linear approximation and the actual curve is less than 1 percent. The maximum differential input signal for the MOSFET diff-amp is much larger than for the bipolar diff-amp. The principal reason is that the gain of the MOSFET diff-amp, as we will see, is much smaller than the gain of the bipolar diff-amp.

Figure 11.22 is the ac equivalent circuit of the diff-amp configuration, showing only the differential voltage and signal currents as a function of the transistor transconductance g_m . We assume that the output resistance looking into the current source is infinite. Using this equivalent circuit, we find the one-sided output voltage at v_{o2} , as follows:

$$v_{o2} \equiv v_o = +\left(\frac{g_m v_d}{2} \right) R_D \quad (11.73)$$

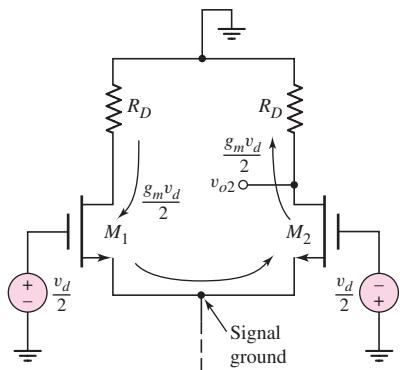


Figure 11.22 AC equivalent circuit, MOSFET differential amplifier

The differential voltage gain is then

$$A_d = \frac{v_o}{v_d} = \frac{g_m R_D}{2} = \sqrt{\frac{K_n I_Q}{2}} \cdot R_D \quad (11.74)$$

11.3.2 Differential- and Common-Mode Input Impedances

At low frequencies, the input impedance of a MOSFET is essentially infinite, which means that both the differential- and common-mode input resistances of a MOSFET diff-amp are infinite. Also, we know that the differential input resistance of a bipolar pair can be in the low kilohm range. A design trade-off, then, would be to use a MOSFET diff-amp with infinite input resistance, and sacrifice the differential-mode voltage gain.

11.3.3 Small-Signal Equivalent Circuit Analysis

We can determine the basic relationships for the differential-mode gain, common-mode gain, and common-mode rejection ratio from an analysis of the small-signal equivalent circuit.

Figure 11.23 shows the small-signal equivalent circuit of the MOSFET differential pair configuration. We assume the transistors are matched, with $\lambda = 0$ for each transistor, and that the constant-current source is represented by a finite output resistance R_o . All voltages are represented by their phasor components. The two transistors are biased at the same quiescent current, and $g_{m1} = g_{m2} \equiv g_m$.

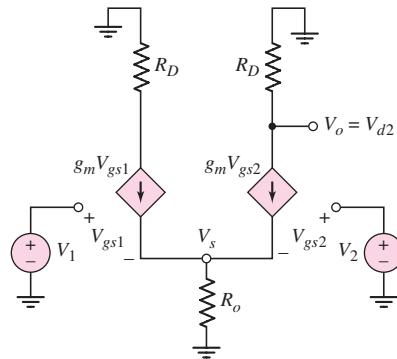


Figure 11.23 Small-signal equivalent circuit, MOSFET differential amplifier

Writing a KCL equation at node V_s , we have

$$g_m V_{gs1} + g_m V_{gs2} = \frac{V_s}{R_o} \quad (11.75)$$

From the circuit, we see that $V_{gs1} = V_1 - V_s$ and $V_{gs2} = V_2 - V_s$. Equation (11.75) then becomes

$$g_m(V_1 + V_2 - 2V_s) = \frac{V_s}{R_o} \quad (11.76)$$

Solving for V_s we obtain

$$V_s = \frac{V_1 + V_2}{2 + \frac{1}{g_m R_o}} \quad (11.77)$$

For a one-sided output at the drain of M_2 , we have

$$V_o = V_{d2} = -(g_m V_{gs2}) R_D = -(g_m R_D)(V_2 - V_s) \quad (11.78)$$

Substituting Equation (11.77) into (11.78) and rearranging terms yields

$$V_o = -g_m R_D \left[\frac{V_2 \left(1 + \frac{1}{g_m R_o} \right) - V_1}{2 + \frac{1}{g_m R_o}} \right] \quad (11.79)$$

Based on the relationships between the input voltages V_1 and V_2 and the differential- and common-mode voltages, as given by Equation (11.29), Equation (11.79) can be written

$$V_o = \frac{g_m R_D}{2} V_d - \frac{g_m R_D}{1 + 2g_m R_o} V_{cm} \quad (11.80)$$

The output voltage, in general form, is

$$V_o = A_d V_d + A_{cm} V_{cm} \quad (11.81)$$

The transconductance g_m of the MOSFET is

$$g_m = 2\sqrt{K_n I_{DQ}} = \sqrt{2K_n I_Q}$$

Comparing Equations (11.80) and (11.81), we develop the relationships for the differential-mode gain,

$$A_d = \frac{g_m R_D}{2} = \sqrt{2K_n I_Q} \left(\frac{R_D}{2} \right) = \sqrt{\frac{K_n I_Q}{2}} \cdot R_D \quad (11.82(a))$$

and the common-mode gain

$$A_{cm} = \frac{-g_m R_D}{1 + 2g_m R_o} = \frac{-\sqrt{2K_n I_Q} \cdot R_D}{1 + 2\sqrt{2K_n I_Q} \cdot R_o} \quad (11.82(b))$$

We again see that for an ideal current source, the common-mode gain is zero since $R_o = \infty$.

From Equations (11.82(a)) and (11.82(b)), the common-mode rejection ratio, $\text{CMRR} = |A_d/A_{cm}|$, is found to be

$$\text{CMRR} = \frac{1}{2} [1 + 2\sqrt{2K_n I_Q} \cdot R_o] \quad (11.83)$$

This demonstrates that the CMRR for the MOSFET diff-amp is also a strong function of the output resistance of the constant-current source.

EXAMPLE 11.9

Objective: Determine the differential-mode voltage gain, common-mode voltage gain, and CMRR for a MOSFET diff-amp.

Consider a MOSFET diff-amp with the configuration in Figure 11.20. Assume the same transistor parameters as given in Example 11.8 except assume $\lambda = 0.01 \text{ V}^{-1}$ for M_4 .

Solution: From Example 11.8, we found the bias current to be $I_Q = 0.587$ mA. The output resistance of the current source is then

$$R_o = \frac{1}{\lambda I_Q} = \frac{1}{(0.01)(0.587)} = 170 \text{ k}\Omega$$

The differential-mode voltage gain is

$$A_d = \sqrt{\frac{K_n I_Q}{2}} \cdot R_D = \sqrt{\frac{(1)(0.587)}{2}} \cdot (16) = 8.67$$

and the common-mode voltage gain is

$$A_{cm} = -\frac{\sqrt{2K_n I_Q} \cdot R_D}{1 + 2\sqrt{2K_n I_Q} \cdot R_o} = -\frac{\sqrt{2(1)(0.587)} \cdot (16)}{1 + 2\sqrt{2(1)(0.587)} \cdot (170)} = -0.0469$$

The common-mode rejection ratio is then

$$\text{CMRR}_{\text{dB}} = 20 \log_{10} \left(\frac{8.67}{0.0469} \right) = 45.3 \text{ dB}$$

Comment: As mentioned earlier, the differential-mode voltage gain of the MOSFET diff-amp is considerably less than that of the bipolar diff-amp, since the value of the MOSFET transconductance is, in general, much smaller than that of the BJT.

EXERCISE PROBLEM

Ex 11.9: The parameters of the circuit shown in Figure 11.19 are $V^+ = 3$ V, $V^- = -3$ V, $I_Q = 0.2$ mA, and $R_D = 15 \text{ k}\Omega$. Assume M_1 and M_2 are matched with parameters $V_{TN} = 0.4$ V, $k'_n = 100 \mu\text{A/V}^2$, and $\lambda = 0$. (a) Design the width-to-length ratios of the transistors such that the one-sided differential voltage gain is $A_d = 15$. (b) Using the results of part (a), what is the value of $g_f(\text{max})$? (Ans. (a) $W/L = 200$, (b) $g_f(\text{max}) = 1.0 \text{ mA/V}$)

The value of the common-mode rejection ratio can be increased by increasing the output resistance of the current source. An increase in the output resistance can be accomplished by using a more sophisticated current source circuit. Figure 11.24

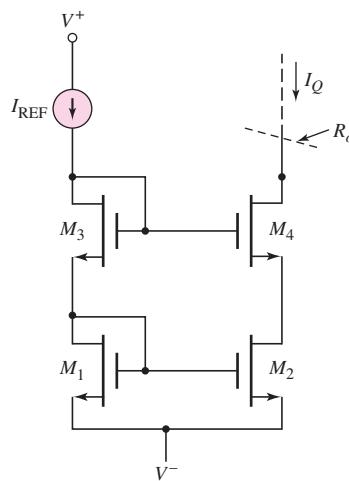


Figure 11.24 MOSFET cascode current source

shows a MOSFET cascode current mirror that was discussed in the last chapter. The output resistance, as given by Equation (10.57), is $R_o = r_{o2} + r_{o4}(1 + g_m r_{o4})$. For the parameters of Example 11.9, $r_{o2} = r_{o4} = 170 \text{ k}\Omega$ and $g_m = 2\sqrt{K_n I_Q} = 1.53 \text{ mA/V}$. Then

$$R_o = 170 + 170[1 + (1.53)(170)] \Rightarrow 44.6 \text{ M}\Omega$$

Again, using the parameters of Example 11.9, the common-mode voltage gain of the diff-amp with a cascode current mirror would be

$$A_{cm} = -\frac{\sqrt{2K_n I_Q} \cdot R_D}{1 + 2\sqrt{2K_n I_Q} \cdot R_o} = -\frac{\sqrt{2(1)(0.587)} \cdot (16)}{1 + 2\sqrt{2(1)(0.587)} \cdot (44600)} = -0.000179$$

so that the CMRR would be

$$\text{CMRR}_{\text{dB}} = 20 \log_{10} \left(\frac{8.67}{0.000179} \right) = 93.7 \text{ dB}$$

We increased the common-mode rejection ratio dramatically by using the cascode current mirror instead of the single two-transistor current source. Note, however, that the differential-mode voltage gain is unchanged.

To gain an appreciation of the difference in CMRR between 45.3 dB and 93.7 dB, we can reconsider the linear scale. For a $\text{CMRR}_{\text{dB}} = 45.3 \text{ dB}$, the differential gain is a factor of 185 times larger than the common-mode gain, while for a $\text{CMRR}_{\text{dB}} = 93.7 \text{ dB}$, the differential gain is a factor of 48,436 times larger than the common-mode gain.

11.3.4 Two-Sided Output

If we consider the two-sided output of an ideal MOSFET op-amp and define the output voltage as $V_o = V_{d2} - V_{d1}$, we can show that the differential-mode voltage gain is given by

$$A_d = g_m R_D \quad (11.84(\text{a}))$$

and the common-mode voltage gain is given by

$$A_{cm} = 0 \quad (11.84(\text{b}))$$

The result of $A_{cm} = 0$ for the two-sided output is a consequence of using matched devices and elements in the diff-amp circuit. We will reconsider a two-sided output and discuss the effects of mismatched elements in the next section.

Effect of R_D Mismatch—Two-Sided Output

We assume that R_{D1} and R_{D2} are the resistors in the drains of M_1 and M_2 . If the two resistors are not matched, we assume that we can write $R_{D1} = R_D + \Delta R_D$ and $R_{D2} = R_D - \Delta R_D$. Using the small-signal equivalent circuit in Figure 11.23, we can find

$$A_d = g_m R_D \quad (11.85(\text{a}))$$

and

$$A_{cm} \cong \frac{\Delta R_D}{R_o} \quad (11.85(\text{b}))$$

The common-mode rejection ratio is then

$$\text{CMRR} = \left| \frac{A_d}{A_{cm}} \right| = \frac{g_m R_D}{(\Delta R_D / R_o)} \quad (11.86)$$

This result is essentially the same as the BJT diff-amp.

Effect of g_m Mismatch—Two-Sided Output

We can consider the effect of transistor mismatch by considering the effect of a mismatch in the transconductance g_m . We assume g_{m1} and g_{m2} are the transconductance parameters of the two transistors in the diff-amp. We will assume that we can write $g_{m1} = g_m + \Delta g_m$ and $g_{m2} = g_m - \Delta g_m$. Again, using the small-signal equivalent circuit shown in Figure 11.23, we find the differential-mode voltage gain is

$$A_d = g_m R_D \quad (11.87(\text{a}))$$

and the common-mode gain is

$$A_{cm} = \frac{R_D(2\Delta g_m)}{1 + 2R_o g_m} \quad (11.87(\text{b}))$$

The common-mode rejection ratio now becomes

$$\text{CMRR} = \left| \frac{A_d}{A_{cm}} \right| = \frac{1 + 2R_o g_m}{2(\Delta g_m / g_m)} \quad (11.88)$$

The CMRR of mismatched elements in the MOSFET diff-amp is identical with the results of mismatched elements in the BJT diff-amp.

11.3.5 JFET Differential Amplifier

Figure 11.25 shows a basic JFET differential pair biased with a constant-current source. If a pure differential-mode input signal is applied such that $v_{G1} = +v_d/2$ and $v_{G2} = -v_d/2$, then drain currents I_{D1} and I_{D2} increase and decrease, respectively, in exactly the same way as in the MOSFET diff-amp.

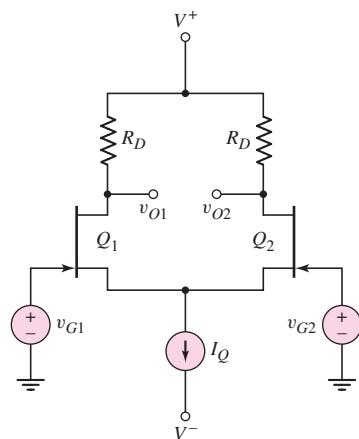


Figure 11.25 Basic JFET differential pair configuration

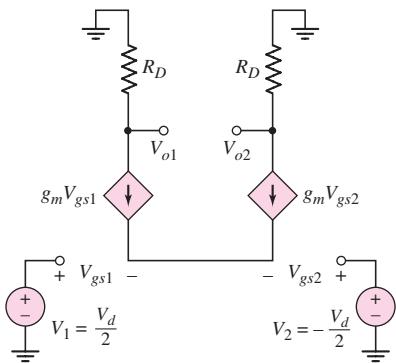


Figure 11.26 Small-signal equivalent circuit, JFET differential amplifier

We can determine the differential-mode voltage gain by analyzing the small-signal equivalent circuit. Figure 11.26 shows the equivalent circuit, with the output resistance of the constant-current source and the small-signal resistances of Q_1 and Q_2 assumed to be infinite. The small-signal equivalent circuit of the JFET diff-amp is identical to that of the MOSFET diff-amp in Figure 11.23 for the case when the current-source output resistance is infinite. A KCL equation at the common-source node, in phasor notation, is

$$g_m V_{gs1} + g_m V_{gs2} = 0 \quad (11.89(a))$$

or

$$V_{gs1} = -V_{gs2} \quad (11.89(b))$$

The differential-mode input voltage is

$$V_d \equiv V_1 - V_2 = V_{gs1} - V_{gs2} = -2V_{gs2} \quad (11.90)$$

A one-sided output at V_{o2} is given by

$$V_{o2} = -g_m V_{gs2} R_D = -g_m \left(\frac{-V_d}{2} \right) R_D \quad (11.91)$$

and the differential-mode voltage gain is

$$A_d = \frac{V_{o2}}{V_d} = +\frac{g_m R_D}{2} \quad (11.92)$$

The expression for the differential-mode voltage gain for the JFET diff-amp (Equation (11.92)) is exactly the same as that of the MOSFET diff-amp (Equation 11.82(a)). If the constant-current source output resistance is finite, then the JFET diff-amp will also have a nonzero common-mode voltage gain.

Test Your Understanding

TYU 11.7 The circuit parameters of the diff-amp shown in Figure 11.19 are $V^+ = 3$ V, $V^- = -3$ V, $I_Q = 0.40$ mA, and $R_D = 7.5$ k Ω . The transistor parameters are $V_{TN} = 0.5$ V, $k'_n = 100$ μ A/V 2 , and $\lambda = 0$. (a) Design the transistor W/L ratio such that the differential voltage gain is $A_d = 12$. (b) What is the maximum positive common-mode voltage that can be applied such that the transistors remain biased in the saturation region. (Ans. (a) $W/L = 256$, (b) $v_{cm} = 2$ V)

TYU 11.8 Consider the differential amplifier in Figure 11.20. The transistor parameters are given in Example 11.8, except that $\lambda = 0.02 \text{ V}^{-1}$ for M_3 and M_4 . Determine the differential voltage gain $A_d = v_{o2}/v_d$, the common-mode gain $A_{cm} = v_{o2}/v_{cm}$, and the CMRR_{dB}. (Ans. $A_d = 2.74$, $A_{cm} = -0.0925$, CMRR_{dB} = 29.4 dB)

TYU 11.9 The diff-amp in Figure 11.19 is biased at $I_Q = 100 \mu\text{A}$. The transistor parameters are $k'_n = 100 \mu\text{A/V}^2$ and $W/L = 10$. (a) If the output resistance of the current source is $R_o = 1 \text{ M}\Omega$, determine the common-mode rejection ratio CMRR_{dB}. (b) Determine the required value of R_o such that CMRR_{dB} = 80 dB. (Ans. (a) CMRR_{dB} = 50 dB, (b) $R_o = 31.6 \text{ M}\Omega$)

***TYU 11.10** The differential amplifier in Figure 11.20 is to be redesigned. The current-source biasing is to be replaced with the cascode current source in Figure 11.24. The reference current is $I_{REF} = 100 \mu\text{A}$ and λ for transistors in the current source circuit is 0.01 V^{-1} . The parameters of the differential pair M_1 and M_2 are the same as described in Example 11.8. The range of the common-mode input voltage is to be $-4 \leq v_{cm} \leq +4 \text{ V}$. Redesign the diff-amp to achieve the highest possible differential-mode voltage gain. Determine the values of A_d , A_{cm} , and CMRR_{dB}. (Ans. $A_d = 9.90$, $A_{cm} = 0.0003465$, CMRR_{dB} = 89.1 dB)

11.4 DIFFERENTIAL AMPLIFIER WITH ACTIVE LOAD

Objective: • Analyze the characteristics of BJT and FET differential amplifiers with active loads.

In Chapter 10, we considered an active load in conjunction with a simple transistor amplifier. Active loads can also be used in diff-amp circuits to increase the differential-mode gain.

Active loads are essentially transistor current sources used in place of resistive loads. The transistors in the active load circuit are biased at a Q -point in the forward-active mode as shown in Figure 11.27. A change in collector current is induced by the differential-pair, which, in turn, produces a change in the emitter–collector voltage as shown in the figure. The relation between the change in current and change in voltage is proportional to the small-signal output resistance r_o of the transistor. The value of r_o is, in general, much larger than that of a discrete resistive load, so the small-signal voltage gain will be larger with the active load.

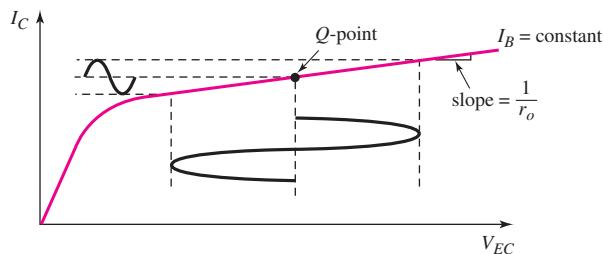


Figure 11.27 Current–voltage characteristic of active load device

11.4.1 BJT Diff-Amp with Active Load

Figure 11.28 shows a differential amplifier with an active load. Transistors Q_1 and Q_2 are the differential pair biased with a constant current I_Q , and transistors Q_3 and Q_4 form the load circuit. From the collectors of Q_2 and Q_4 , we obtain a one-sided output.

If we assume all transistors are matched, then a pure applied common-mode voltage means that $v_{B1} = v_{B2} = v_{CM}$, and current I_Q splits evenly between Q_1 and Q_2 . Neglecting base currents, $I_4 = I_3$ through the current-source circuit and $I_1 = I_2 = I_3 = I_4 = I_Q/2$ with no load connected at the output.

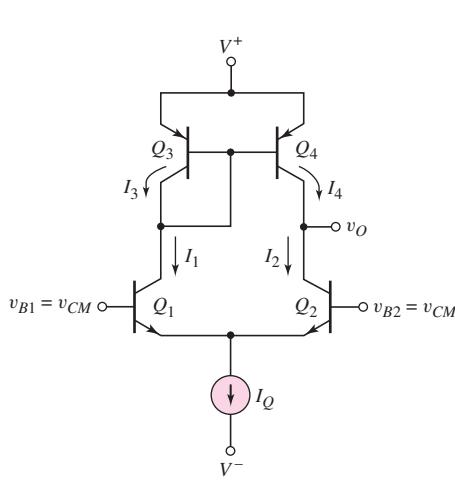


Figure 11.28 BJT differential amplifier with active load

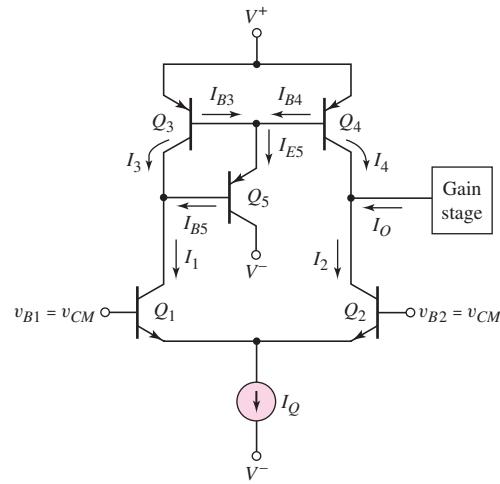


Figure 11.29 BJT differential amplifier with three-transistor active load and second stage

In actual diff-amp circuits, base currents are not zero. In addition, a second amplifier stage is connected at the diff-amp output. Figure 11.29 shows a diff-amp with an active load circuit, corresponding to a three-transistor current source, as well as a second amplifying stage. In general, the common-emitter current gain β is a function of collector current, as was shown in Figure 6.21(c). However, for simplicity, we assume all transistor current gains are equal, even though the current level in Q_5 is much smaller than in the other transistors. Current I_O is the dc bias current from the gain stage. Assuming all transistors are matched and $v_{B1} = v_{B2} = v_{CM}$, current I_Q splits evenly and $I_1 = I_2$. To ensure that Q_2 and Q_4 are biased in the forward-active mode, the dc currents must be balanced, or $I_3 = I_4$. We see that

$$I_{E5} = I_{B3} + I_{B4} = \frac{I_3}{\beta} + \frac{I_4}{\beta} \quad (11.93)$$

Then

$$I_{B5} = \frac{I_{E5}}{1 + \beta} = \frac{I_3 + I_4}{\beta(1 + \beta)} \quad (11.94)$$

If the base currents and I_O are small, then

$$I_3 + I_4 \cong I_Q \quad (11.95)$$

Therefore,

$$I_{B5} \cong \frac{I_Q}{\beta(1 + \beta)} \quad (11.96)$$

For the circuit to be balanced, that is, for $I_1 = I_2$ and $I_3 = I_4$, we must have

$$I_Q = I_{B5} = \frac{I_Q}{\beta(1 + \beta)} \quad (11.97)$$

Equation (11.97) implies that the second amplifying stage must be designed and biased such that the direction of the dc bias current is as shown and is equal to the result of Equation (11.97). To illustrate this condition, we will analyze a second amplifying stage using a Darlington pair, later in this chapter.

11.4.2 Small-Signal Analysis of BJT Active Load

Figure 11.30 shows a diff-amp with a three-transistor active load circuit. The resistance R_L represents the small-signal input resistance of the gain stage. We will assume that a pure differential-mode input voltage is applied as shown in the figure. From previous results, we know that the common-emitter terminals of Q_1 and Q_2 are at signal ground. The signal voltage at the base of Q_1 produces a signal collector current $i_1 = (g_m v_d)/2$, where g_m is the transistor transconductance for both Q_1 and Q_2 . Assuming the base currents are negligible, a signal current $i_3 = i_1$ is induced in Q_3 , and the current mirror produces a signal current i_4 equal to i_3 . The signal voltage at the base of Q_2 produces a signal collector current $i_2 = (g_m v_d)/2$, with the direction shown. The two signal currents, i_2 and i_4 , add to produce a signal current in the load resistance R_L . The discussion is a first-order evaluation of the circuit operation.

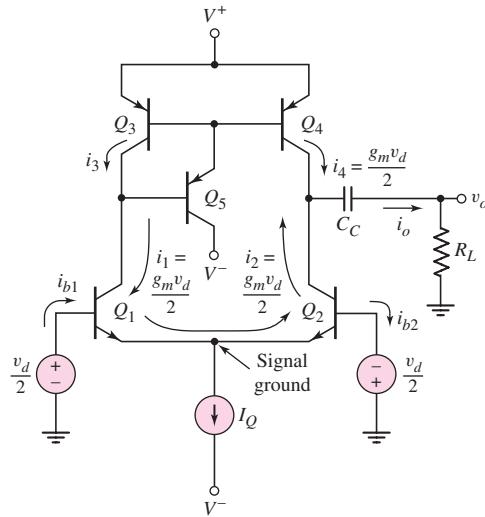


Figure 11.30 BJT differential amplifier with three-transistor active load, showing the signal currents

From the above discussion, we know the induced currents in Q_2 and Q_4 . To more accurately determine the output voltage, we need to consider the equivalent small-signal collector-emitter output circuit of the two transistors. Figure 11.31(a)

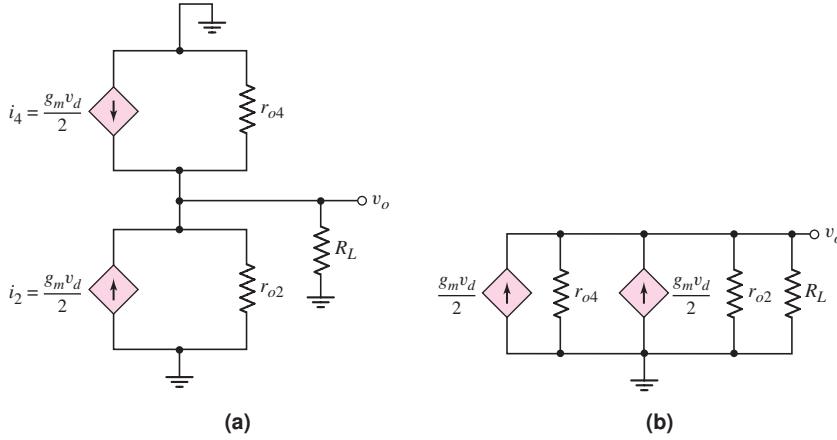


Figure 11.31 (a) Small-signal equivalent circuit BJT differential amplifier with active load and (b) rearrangement of small-signal equivalent circuit

shows the small-signal equivalent circuit at the collector nodes of Q_2 and Q_4 . The circuit can be rearranged to combine the signal grounds at a common point, as in Figure 11.31(b). From this figure, we determine that

$$v_o = 2\left(\frac{g_m v_d}{2}\right)(r_{o2} \| r_{o4} \| R_L) \quad (11.98)$$

and the small-signal differential-mode voltage gain is

$$A_d = \frac{v_o}{v_d} = g_m(r_{o2} \| r_{o4} \| R_L) \quad (11.99)$$

Equation (11.99) can be rewritten in the form

$$A_d = \frac{\frac{g_m}{r_{o2}} + \frac{1}{r_{o4}} + \frac{1}{R_L}}{\frac{1}{r_{o2}} + \frac{1}{r_{o4}} + \frac{1}{R_L}} = \frac{g_m}{g_{o2} + g_{o4} + G_L} \quad (11.100)$$

We recall that $g_m = I_Q/2V_T$, $r_{o2} = V_{A2}/I_2$, and $r_{o4} = V_{A4}/I_4$. The parameters g_{o2} , g_{o4} , and G_L are the corresponding conductances. Assuming $I_2 = I_4 = I_Q/2$, we can write Equation (11.100) in the form

$$A_d = \frac{\frac{I_Q}{2V_T}}{\frac{I_Q}{2V_{A2}} + \frac{I_Q}{2V_{A4}} + \frac{1}{R_L}} \quad (11.101)$$

This expression of the differential-mode voltage gain of the diff-amp with an active load is very similar to that obtained in the last chapter for a simple amplifier with an active load.

The output resistance looking back into the common collector node is $R_o = r_{o2} \| r_{o4}$. To minimize loading effects, we need $R_L > R_o$. However, since R_o is generally large for active loads, we may not be able to satisfy this condition. We can determine the severity of the loading effect by comparing R_L and R_o .

EXAMPLE 11.10

Objective: Determine the differential-mode gain of a diff-amp with an active load, taking loading effects into account.

Consider the diff-amp in Figure 11.30, biased with $I_Q = 0.20$ mA. Assume an Early voltage of $V_A = 100$ V for all transistors. Determine the open-circuit ($R_L = \infty$) differential-mode voltage gain, as well as the differential-mode voltage gain when $R_L = 100$ k Ω .

Solution: From Equation (11.101), the open-circuit voltage gain becomes

$$A_d = \frac{\frac{1}{V_T}}{\frac{1}{V_{A2}} + \frac{1}{V_{A4}}} = \frac{\frac{1}{0.026}}{\frac{1}{100} + \frac{1}{100}} = 1923$$

When $R_L = 100$ k Ω , the voltage gain is

$$A_d = \frac{\frac{0.20 \times 10^{-3}}{2(0.026)}}{\frac{0.20 \times 10^{-3}}{2(100)} + \frac{0.20 \times 10^{-3}}{2(100)} + \frac{1}{100 \times 10^3}}$$

which can be written

$$A_d = \frac{\frac{0.20}{2(0.026)}}{\frac{0.20}{2(100)} + \frac{0.20}{2(100)} + \frac{1}{100}} = \frac{3.85}{0.001 + 0.001 + 0.01} = 321$$

An inspection of this last equation shows that the external load factor, $1/R_L$, dominates the denominator term and thus has a tremendous influence on the gain.

Comment: The open-circuit differential-mode voltage gain, for a diff-amp with an active load, is large. However, a finite load resistance R_L causes severe loading effects, as shown in this example. A 100 k Ω load caused almost an order of magnitude decrease in the gain.

EXERCISE PROBLEM

Ex 11.10: The diff-amp circuit in Figure 11.30 is biased at $I_Q = 0.4$ mA. The transistor parameters are $\beta = 120$, $V_{A1} = V_{A2} = 150$ V, and $V_{A3} = V_{A4} = 90$ V. (a) Determine the open-circuit ($R_L = \infty$) differential-mode voltage gain. (b) Find the differential-mode voltage gain when $R_L = 250$ k Ω . (c) Determine the differential-mode input resistance. (d) Find the output resistance looking back from the load R_L . (Ans. (a) 2163, (b) 1018, (c) 31.2 k Ω , (d) 281 k Ω)

11.4.3 MOSFET Differential Amplifier with Active Load

We can use an active load in conjunction with a MOSFET differential pair, as we did for the bipolar differential amplifier. Figure 11.32 shows a MOSFET diff-amp with an active load. Transistors M_1 and M_2 are n-channel devices and form the differential

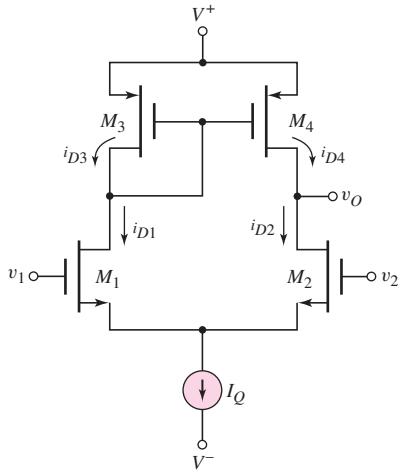


Figure 11.32 MOSFET differential amplifier with active load

pair biased with I_Q . The load circuit consists of transistors M_3 and M_4 , both p-channel devices, connected in a current mirror configuration. A one-sided output is taken from the common drains of M_2 and M_4 . When a common-mode voltage of $v_1 = v_2 = v_{cm}$ is applied, the current I_Q splits evenly between M_1 and M_2 , and $i_{D1} = i_{D2} = I_Q/2$. There are no gate currents; therefore, $i_{D3} = i_{D1}$ and $i_{D4} = i_{D2}$.

If a small differential-mode input voltage $v_d = v_1 - v_2$ is applied, then from Equation (11.66) and (11.67), we can write

$$i_{D1} = \frac{I_Q}{2} + i_d \quad (11.102(a))$$

and

$$i_{D2} = \frac{I_Q}{2} - i_d \quad (11.102(b))$$

where i_d is the signal current. For small values of v_d , we have $i_d = (g_m v_d)/2$. Since M_1 and M_3 are in series, we see that

$$i_{D3} = i_{D1} = \frac{I_Q}{2} + i_d \quad (11.103)$$

Finally, the current mirror consisting of M_3 and M_4 produces

$$i_{D4} = i_{D3} = \frac{I_Q}{2} + i_d \quad (11.104)$$

Figure 11.33 is the ac equivalent circuit of the diff-amp with active load, showing the signal currents. The negative sign for i_{D2} in Equation (11.102(b)) shows up as a change in current direction in M_2 , as indicated in the figure.

Figure 11.34(a) shows the small-signal equivalent circuit at the drain node of M_2 and M_4 . If the output is connected to the gate of another MOSFET, which is equivalent to an infinite impedance at low frequency, the output terminal is effectively an open circuit. The circuit can be rearranged by combining the signal grounds at a common point, as shown in Figure 11.34(b). Then,

$$v_o = 2\left(\frac{g_m v_d}{2}\right)(r_{o2} \| r_{o4}) \quad (11.105)$$

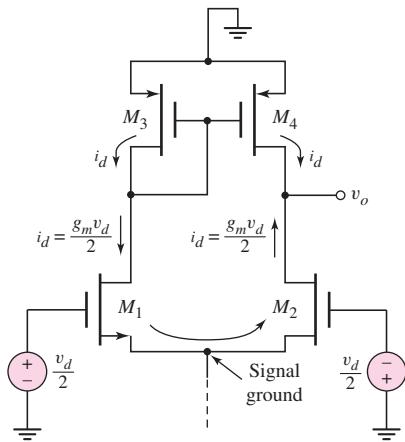


Figure 11.33 The ac equivalent circuit, MOSFET differential amplifier with active load

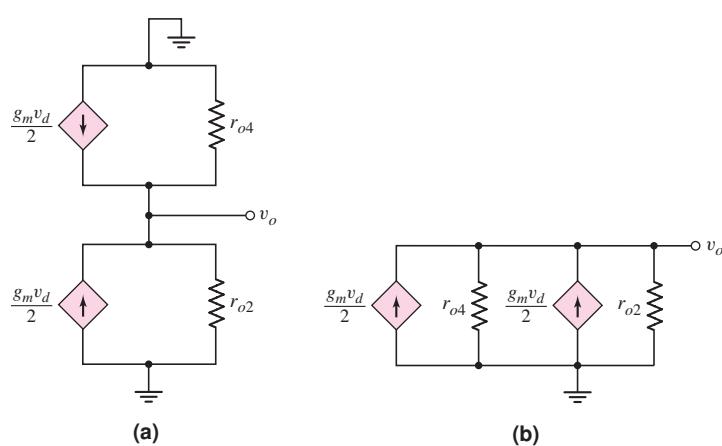


Figure 11.34 (a) Small-signal equivalent circuit, MOSFET differential amplifier with active load and (b) rearranged small-signal equivalent circuit

and the small-signal differential-mode voltage gain is

$$A_d = \frac{v_o}{v_d} = g_m (r_{o2} \| r_{o4}) \quad (11.106)$$

Equation (11.106) can be rewritten in the form

$$A_d = \frac{\frac{g_m}{1 + \frac{1}{r_{o2}}}}{\frac{1}{r_{o2}}} = \frac{g_m}{g_{o2} + g_{o4}} \quad (11.107)$$

If we recall that $g_m = 2\sqrt{K_n I_D} = \sqrt{2K_n I_Q}$, $g_{o2} = \lambda_2 I_{DQ2} = (\lambda_2 I_Q)/2$, and $g_{o4} = \lambda_4 I_{DQ4} = (\lambda_4 I_Q)/2$, then Equation (11.107) becomes

$$A_d = \frac{2\sqrt{2K_n I_Q}}{I_Q(\lambda_2 + \lambda_4)} = 2\sqrt{\frac{2K_n}{I_Q}} \cdot \frac{1}{\lambda_2 + \lambda_4} \quad (11.108)$$

DESIGN EXAMPLE 11.11

Objective: Design a MOSFET diff-amp with the configuration in Figure 11.32 to meet the specifications of the experimental system in Example 11.4.

Design Approach: We need not only to try to obtain the necessary differential-mode gain and minimize the common-mode gain in our design, but we must also be cognizant of the swing in the output voltage. In the circuit in Figure 11.32, if the corresponding PMOS and NMOS transistors are matched, then the quiescent value of V_{SD4} is equal to $V_{SG4} = V_{SG3}$. As the signal output voltage increases, the source-to-drain voltage of M_4 decreases. The minimum value of this voltage such that M_4 remains biased in the saturation region is $V_{SD4}(\min) = V_{SD4}(\text{sat}) = V_{SG} + V_{TP}$. This means that the maximum swing in the output voltage is equal to the magnitude of the threshold voltage of M_4 . In this example, the maximum swing in the output voltage is 0.8 V, so that the magnitude of the threshold voltages of the PMOS devices must be greater than 0.8 V. Assume that NMOS devices are available with the following parameters: $V_{TN} = 0.5$ V, $k'_n = 80 \mu\text{A}/\text{V}^2$, and $\lambda_n = 0.02 \text{ V}^{-1}$. Assume that

PMOS devices are available with the following parameters: $V_{TP} = -1.0$ V, $k'_p = 40 \mu\text{A}/\text{V}^2$, and $\lambda_p = 0.02 \text{ V}^{-1}$. Choose supply voltages of ± 5 V and choose a bias current of approximately $I_Q = 200 \mu\text{A}$.

Figure 11.35 is the diff-amp and current-source network used for the design in this example.

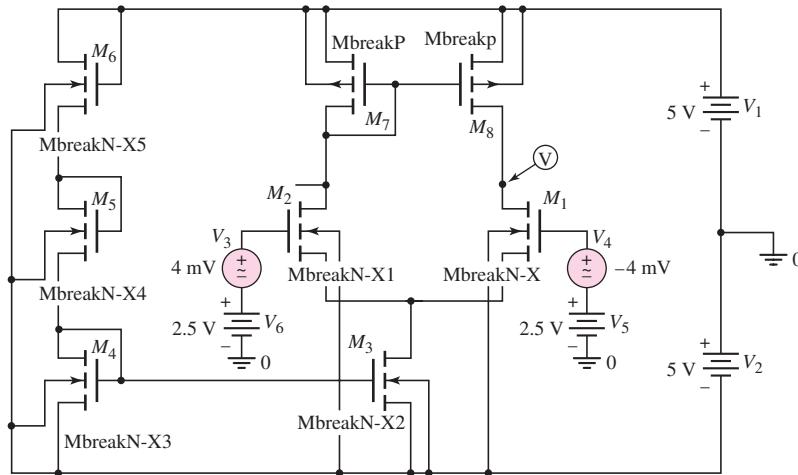


Figure 11.35 CMOS differential amplifier and current source network for Example 11.11

Design, Differential Amplifier: Differential-Mode Gain: From Equation (11.108), the differential-mode gain is

$$A_d = 2\sqrt{2\left(\frac{k'_n}{2}\right)\left(\frac{W}{L}\right)_n \frac{1}{I_Q}} \cdot \frac{1}{\lambda_n + \lambda_p}$$

or

$$100 = 2\sqrt{2\left(\frac{80}{2}\right)\left(\frac{W}{L}\right)_n \frac{1}{200}} \cdot \frac{1}{0.02 + 0.02}$$

which yields a width-to-length ratio of $(W/L)_n = 10$ for the NMOS differential pair. Since the width-to-length ratios of the other transistors do not directly affect the gain of the diff-amp, we may arbitrarily choose width-to-length ratios of 10 for all other transistors except M_5 and M_6 . The W/L ratio of 10 means that the other devices are reasonably small and do not lead to a large circuit area.

Design, Current-Source Network: For the transistor M_3 in the current source, we have

$$I_Q = \frac{k'}{2} \cdot \frac{W}{L} \cdot (V_{GS3} - V_{TN})^2$$

or

$$200 = \frac{80}{2}(10)(V_{GS3} - 0.5)^2$$

which means that the required gate-to-source voltage of M_3 is $V_{GS3} = 1.21$ V. We may choose M_4 and M_3 to be identical so that the current in the reference portion of

the circuit is also $200 \mu\text{A}$. Assuming that M_5 and M_6 are identical, then each transistor must have a gate-to-source voltage of

$$V_{GS5} = V_{GS6} = (10 - 1.21)/2 \cong 4.4 \text{ V}$$

The width-to-length of these transistors is now found from

$$I_{\text{REF}} = I_Q = \frac{k'_n}{2} \cdot \left(\frac{W}{L}\right)_5 (V_{GS5} - V_{TN})^2$$

or

$$200 = \frac{80}{2} \cdot \left(\frac{W}{L}\right)_5 (4.4 - 0.5)^2$$

which yields

$$(W/L)_5 = (W/L)_6 = 0.33$$

Computer Simulation Verification: The circuit in Figure 11.35 was used in the computer simulation verification. In the hand design, the finite output resistance (lambda parameter) was neglected in the dc calculations. These parameters became important in the actual design and in the actual currents developed in the circuit. For $(W/L)_5 = (W/L)_6 = 0.75$, the reference current is $I_{\text{REF}} = 231 \mu\text{A}$ and the bias current is $I_Q = 208 \mu\text{A}$.

The differential-mode voltage gain is approximately 102 so that the signal output voltage is 0.82 V for a differential-mode input signal voltage of 8 mV . The common-mode output signal is approximately 0.86 mV , which is well within the specified 10 mV maximum value.

Design Pointer: The body effect has been neglected in this design. In actual integrated circuits, the differential pair transistors may actually be fabricated within their own p-type substrate region (for NMOS devices). This p-type substrate region is then directly connected to the source terminals so that the body effect in the NMOS differential pair devices can be neglected.

EXERCISE PROBLEM

Ex 11.11: Determine I_{REF} , I_Q , and A_d of the diff-amp designed in Example 11.11 for the case when the bias voltages are changed to $V^+ = +3 \text{ V}$ and $V^- = -3 \text{ V}$.
(Ans. $I_{\text{REF}} = I_Q = 56.37 \mu\text{A}$, $A_d = 188$)

11.4.4

MOSFET Diff-Amp with Cascode Active Load

The differential-mode voltage gain is proportional to the output resistance looking into the active load transistor. The voltage gain can be increased, therefore, if the output resistance can be increased. An increase in output resistance can be achieved by using, for example, a cascode active load. This configuration is shown in Figure 11.36.

The output resistance R_o was considered in the last section in the discussion of the cascode current source. As applied to Figure 11.36, the output resistance is given by

$$R_o = r_{o4} + r_{o6}(1 + g_m r_{o4}) \cong g_m r_{o4} r_{o6} \quad (11.109)$$

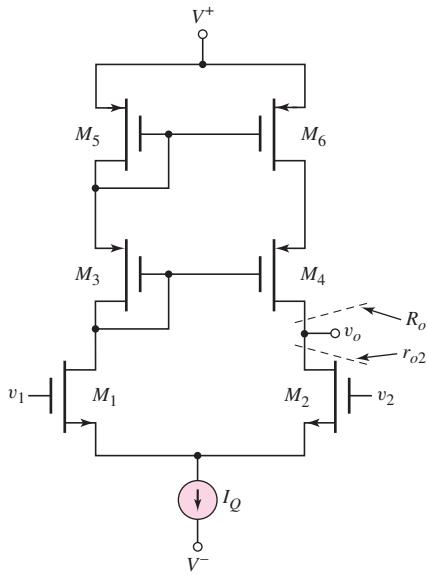


Figure 11.36 MOSFET diff-amp with cascode active load

The small-signal differential-mode voltage gain is then

$$A_d = \frac{v_o}{v_d} = g_m(r_{o2} \| R_o) \quad (11.110)$$

EXAMPLE 11.12

Objective: Calculate the differential-mode voltage gain of a MOSFET diff-amp with a cascode active load.

Consider the diff-amp shown in Figure 11.36. Assume the circuit and transistor parameters are the same as in Example 11.11.

Solution: The transistor transconductance is

$$g_m = 2\sqrt{K_n I_{DQ}} = 2\sqrt{\left(\frac{0.08}{2}\right)(10)(0.1)} = 0.40 \text{ mA/V}$$

The output resistance of the individual transistors is

$$r_o = \frac{1}{\lambda I_{DQ}} = \frac{1}{(0.02)(0.1)} = 500 \text{ k}\Omega$$

The output resistance of the cascode active load is then

$$R_o = r_{o4} + r_{o6}(1 + g_m r_{o4}) = 0.5 + 0.5[1 + (0.40)(500)] = 101 \text{ M}\Omega$$

The differential-mode voltage gain is then found to be

$$A_d = g_m(r_{o2} \| R_o) = (0.40)(500 \| 101000) = 200$$

Comment: Since $R_o \gg r_{o2}$, the voltage gain is now essentially equal to $A_d = g_m r_{o2}$ which is twice as large as the gain calculated in Example 11.11.

EXERCISE PROBLEM

Ex 11.12: The parameters of the circuit and transistors in Figure 11.36 are the same as described in Example 11.12 except for M_1 and M_2 . Assume $k'_{n1} = k'_{n2} = 80 \mu\text{A/V}^2$. Determine $(W/L)_1 = (W/L)_2$ such that $A_d = 400$. (Ans. $(W/L)_1 = (W/L)_2 = 40$)

The differential-mode voltage gain can be further increased by incorporating a cascode configuration in the differential pair as well as in the active load. One such example is shown in Figure 11.37. Transistors M_3 and M_4 are the cascode transistors for the differential pair M_1 and M_2 . The differential-mode voltage gain is now

$$A_d = \frac{v_o}{v_d} = g_m(R_{o4} \parallel R_{o6})$$

where $R_{o4} \cong g_m r_{o2} r_{o4}$ and $R_{o6} \cong g_m r_{o6} r_{o8}$. The small-signal differential-mode voltage gain of this type of amplifier can be on the order of 10,000.

Other types of MOSFET differential amplifiers will be considered in Chapter 13 when operational amplifier circuits are discussed.

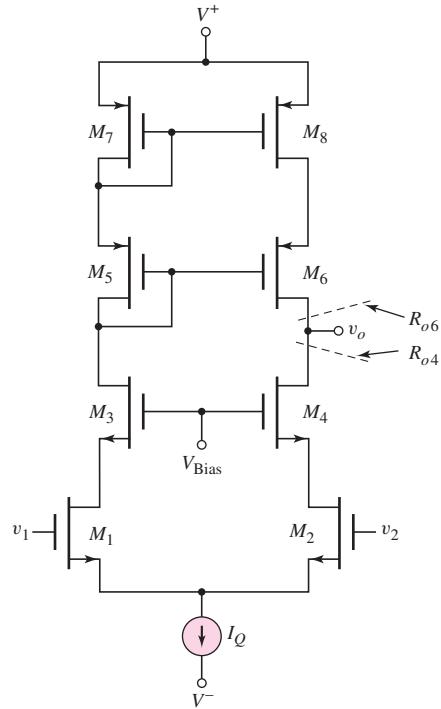


Figure 11.37 A MOSFET cascode diff-amp with a cascode active load

Test Your Understanding

TYU 11.11 Consider the diff-amp in Figure 11.29, with parameters: $V^+ = 10 \text{ V}$, $V^- = -10 \text{ V}$, and $I_Q = 0.5 \text{ mA}$. The transistor parameters are: $\beta = 180$, $V_{BE(\text{on})} = 0.7 \text{ V}$, and $V_A = 100 \text{ V}$. (a) Find I_O such that the circuit is balanced. (b) For the balanced condition, what are the values of V_{EC4} and V_{CE2} , for $v_1 = v_2 = 0$? (Ans. (a) $I_O = 15.3 \text{ nA}$ (b) $V_{EC4} = 1.4 \text{ V}$, $V_{CE2} = 9.3 \text{ V}$)

TYU 11.12 The circuit parameters of the diff-amp in Figure 11.28 are $V^+ = 5\text{ V}$, $V^- = -5\text{ V}$, and $I_Q = 0.1\text{ mA}$. The npn transistor parameters are $\beta_{npn} = 180$, $V_{AN} = 120\text{ V}$, and $V_{BE(on)} = 0.7\text{ V}$; and the pnp transistor parameters are $\beta_{pnp} = 120$, $V_{AP} = 80\text{ V}$, and $V_{EB(on)} = 0.7\text{ V}$. Determine the differential-mode voltage gain. (Ans. $A_d = 1846$)

TYU 11.13 Redesign the circuit in Figure 11.30 using a Widlar current source and bias voltages of $\pm 5\text{ V}$. The bias current I_Q is to be no less than $100\text{ }\mu\text{A}$ and the total power dissipated in the circuit (including the current-source circuit) is to be no more than 10 mW . The diff-amp transistor parameters are the same as in Exercise Ex11.10. The circuit is to provide a minimum loading effect when a second stage with an input resistance of $R = 90\text{ k}\Omega$ is connected to the diff-amp. Determine the differential-mode voltage gain for this circuit. (Ans. $R_1 = 10.3\text{ k}\Omega$, $R_E = 0.571\text{ k}\Omega$, $A_d = 158$)

TYU 11.14 Consider the diff-amp in Figure 11.28, using the parameters described in Exercise TYU11.12. (a) For a differential-mode input signal, determine the output resistance R_o at the output terminal. (b) Determine the load resistance R_L that would reduce the differential-mode voltage gain to one-half the open-circuit value. (Ans. (a) $R_o = 0.96\text{ M}\Omega$, (b) $R_L = 0.96\text{ M}\Omega$)

TYU 11.15 The circuit parameters of the diff-amp in Figure 11.32 are $V^+ = 5\text{ V}$, $V^- = -5\text{ V}$, and $I_Q = 0.2\text{ mA}$. The NMOS transistor parameters are $K_n = 180\text{ }\mu\text{A/V}^2$, $V_{TN} = 0.5\text{ V}$, and $\lambda_n = 0.015\text{ V}^{-1}$ and the PMOS transistor parameters are $K_p = 120\text{ }\mu\text{A/V}^2$, $V_{TP} = -0.5\text{ V}$, and $\lambda_p = 0.025\text{ V}^{-1}$. Determine the differential-mode voltage gain $A_d = v_o/v_d$. (Ans. $A_d = 67.1$)



11.5 BiCMOS CIRCUITS

Objective: • Describe the characteristics of and analyze various BiCMOS circuits.

Thus far, we have considered two basic amplifier design technologies: the bipolar technology, which uses npn and pnp bipolar junction transistors; and the MOS technology, which uses NMOS and PMOS field-effect transistors. We showed that bipolar transistors have a larger transconductance than MOSFETs biased at the same current levels, and that, in general, bipolar amplifiers have larger voltage gains. We also showed that MOSFET circuits have an essentially infinite input impedance at low frequencies, which implies a zero input bias current.

These advantages of the two technologies can be exploited by combining bipolar and MOS transistors in the same integrated circuit. The technology is called **BiCMOS**. BiCMOS technology is especially useful in digital circuit design, but also has applications in analog circuits. In this section, we will examine basic BiCMOS analog circuit configurations.

11.5.1 Basic Amplifier Stages

A bipolar multitransistor circuit previously studied is the Darlington pair configuration. Figure 11.38(a) shows a modified Darlington pair configuration, in which the bias current I_{BIAS} , or some equivalent element, is used to control the quiescent

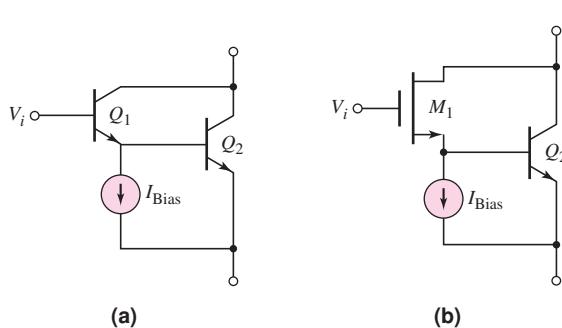


Figure 11.38 (a) Bipolar Darlington pair configuration and (b) BiCMOS Darlington pair configuration

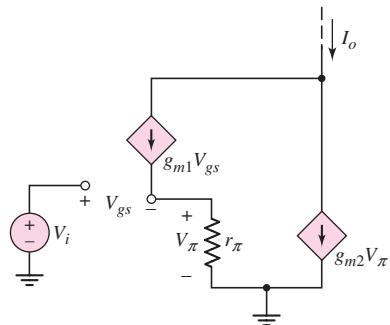


Figure 11.39 Small-signal equivalent circuit, BiCMOS Darlington pair configuration

current in \$Q_1\$. This Darlington pair circuit is used to boost the effective current gain of bipolar transistors. There is no comparable configuration in FET circuits.

A potentially useful BiCMOS circuit is shown in Figure 11.38(b). Transistor \$Q_1\$ in the Darlington pair is replaced with a MOSFET. The advantages of this configuration are an infinite input resistance, and a large transconductance due to the bipolar transistor \$Q_2\$.

To analyze the circuit, we consider the small-signal equivalent circuit shown in Figure 11.39. We assume that \$r_o = \infty\$ in both transistors.

The output signal current is

$$I_o = g_{m1}V_{gs} + g_{m2}V_\pi \quad (11.111)$$

We see that

$$V_i = V_{gs} + V_\pi \quad (11.112)$$

and

$$V_\pi = g_{m1}V_{gs}r_\pi \quad (11.113)$$

Combining Equations (11.112) and (11.113) produces

$$V_{gs} = \frac{V_i}{1 + g_{m1}r_\pi} \quad (11.114)$$

From Equation (11.111), the output current can now be written

$$I_o = g_{m1}V_{gs} + g_{m2}(g_{m1}r_\pi)V_{gs} = (g_{m1} + g_{m2}g_{m1}r_\pi)V_{gs} \quad (11.115)$$

Substituting Equation (11.114) into (11.115), we obtain

$$I_o = \frac{g_{m1}(1 + g_{m2}r_\pi)}{(1 + g_{m1}r_\pi)} \cdot V_i = g_m^c \cdot V_i \quad (11.116)$$

where \$g_m^c\$ is the **composite transconductance**. Since \$g_{m2}\$ of the bipolar transistor is usually at least an order of magnitude greater than \$g_{m1}\$ of the MOSFET, the composite transconductance is approximately an order of magnitude larger than that of the MOSFET alone. We now have the advantages of a large transconductance and an infinite input resistance.

A bipolar cascode circuit is shown in Figure 11.40(a); a corresponding BiCMOS configuration is shown in Figure 11.40(b). The output resistance of the cascode circuit is very high, as we saw in Chapter 10. Also, the cascode amplifier has a wider

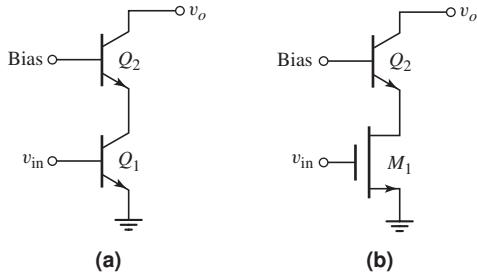


Figure 11.40 (a) Bipolar cascode configuration and (b) BiCMOS cascode configuration

frequency bandwidth than the common-emitter circuit, since the input resistance looking into the emitter of \$Q_2\$ is very low, thereby minimizing the Miller multiplication effect. This effect was observed in Chapter 7.

Again, the advantage of the BiCMOS circuit is the infinite input resistance of \$M_1\$. The equivalent resistance looking into the emitter of a bipolar transistor is much less than the resistance looking into the source of a MOSFET; therefore, the frequency response of a BiCMOS cascode circuit is superior to that of an all-MOSFET cascode circuit.

11.5.2 Current Sources

In our previous discussions of constant-current sources, we mentioned that cascode current sources increase the output resistance, as well as the stability of the bias current. Figure 11.41 shows a bipolar cascode configuration in which the output resistance is \$R_o \cong \beta r_{o4}\$. The bias current in this circuit is much more stable against variations in output voltage than the basic two-transistor current source.

A BiCMOS double cascode constant-current source is shown in Figure 11.42. The small-signal equivalent circuit for determining output resistance is shown in

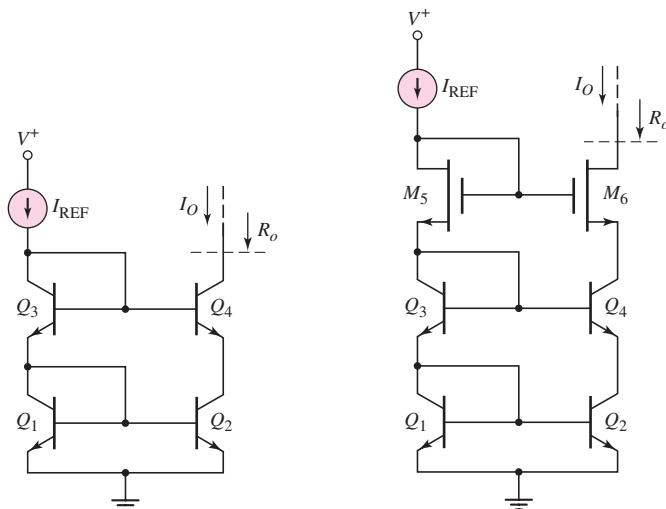


Figure 11.41 Bipolar cascode constant-current source

Figure 11.42 BiCMOS double cascode constant-current source

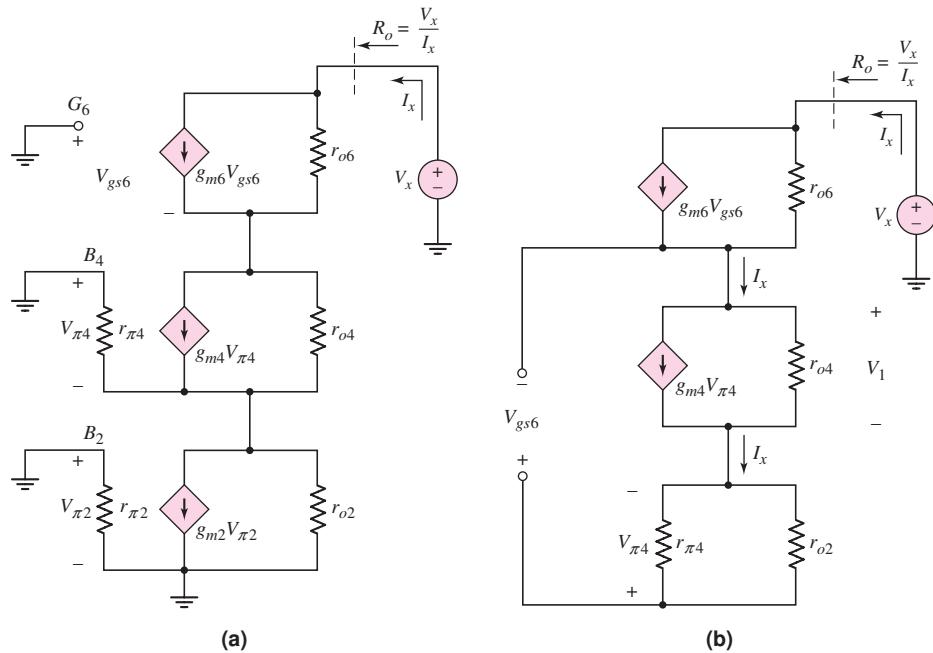


Figure 11.43 (a) Equivalent circuit for determining output impedance of BiCMOS double cascode current source and (b) rearranged equivalent circuit

Figure 11.43(a). The gate voltage to M_6 and the base voltages to Q_2 and Q_4 are constants, equivalent to signal ground. Also, since $V_{\pi2} = 0$, then $g_{m6}V_{\pi2} = 0$, and the equivalent circuit can be rearranged as shown in Figure 11.43(b).

The output resistance of this circuit is extremely large. A detailed analysis shows that the output resistance is given approximately by

$$R_o \cong (g_{m6}r_{o6})(\beta r_{o4}) \quad (11.117)$$

The output resistance is increased by a factor $(g_{m6}r_{o6})$ compared to the bipolar cascode circuit in Figure 11.41. If a bipolar transistor were to be used in place of M_6 , then a resistance $r_{\pi6}$ would be connected across the terminals indicated by V_{gs6} . This resistance would effectively eliminate the multiplying constant $(g_{m6}r_{o6})$, and the output resistance would be essentially the same as that of the circuit in Figure 11.41. The BiCMOS circuit, then, increases the output resistance compared to an all-bipolar circuit.

11.5.3 BiCMOS Differential Amplifier

A basic BiCMOS differential amplifier, with a constant-current source bias and a bipolar active load, is shown in Figure 11.44. Again, the primary advantages are the infinite input resistance and the zero input bias current. One disadvantage of a MOSFET input stage is a relatively high offset voltage compared to that of a bipolar input circuit. Offset voltages occur when the differential-pair input transistors are mismatched. In Chapter 14, we will examine the effect of offset voltages, as well as nonzero bias currents, in op-amp circuits.

We will consider additional BiCMOS op-amp circuits in Chapter 13, when we discuss the analysis and design of full op-amp circuits.

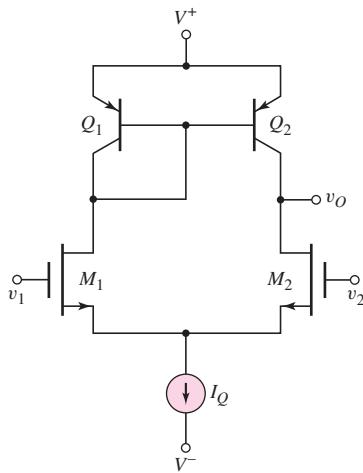


Figure 11.44 Basic BiCMOS differential amplifier

Test Your Understanding

TYU 11.16 Consider the BiCMOS Darlington pair in Figure 11.45. The NMOS transistor parameters are $K_n = 50 \mu\text{A/V}^2$, $V_{TN} = 0.5 \text{ V}$, and $\lambda = 0$. The BJT parameters are $\beta = 150$, $V_{BE(\text{on})} = 0.7 \text{ V}$, and $V_A = \infty$. Determine the small-signal parameters for each transistor, as well as the composite transconductance. (Ans. $g_{m1} = 71.4 \mu\text{A/V}$, $g_{m2} = 2.865 \text{ mA/V}$, $r_{\pi 2} = 52.3 \text{ k}\Omega$, $g_m^c = 2.275 \text{ mA/V}$)

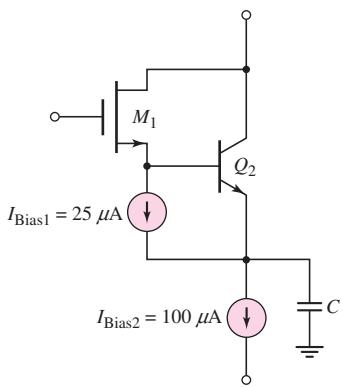


Figure 11.45 Figure for Exercise TYU 11.16

TYU 11.17 The reference current in each of the constant-current source circuits shown in Figures 11.41 and 11.42 is $I_{\text{REF}} = 0.5 \text{ mA}$. All bipolar transistor parameters are $\beta = 150$ and $V_A = 80 \text{ V}$, and all MOSFET parameters are: $K_n = 500 \mu\text{A/V}^2$, $V_{TN} = 1 \text{ V}$, and $\lambda = 0.0125 \text{ V}^{-1}$. Neglecting bipolar base currents, determine the output resistance R_o of each constant-current source. (Ans. For Figure 11.41, $R_o \cong 24 \text{ M}\Omega$; for Figure 11.42, $R_o = 3840 \text{ M}\Omega$)

11.6 GAIN STAGE AND SIMPLE OUTPUT STAGE

Objective: • Analyze an example of a gain stage and output stage of a multistage amplifier.

A diff-amp, including those previously discussed, is the input stage of virtually all op-amps. The second op-amp stage, or gain stage, is often a Darlington pair configuration, and the third, or output, stage is normally an emitter follower.

11.6.1 Darlington Pair and Simple Emitter-Follower Output

Figure 11.46 shows a BJT diff-amp with a three-transistor active load, a Darlington pair connected to the diff-amp output, and a simple emitter-follower output stage.

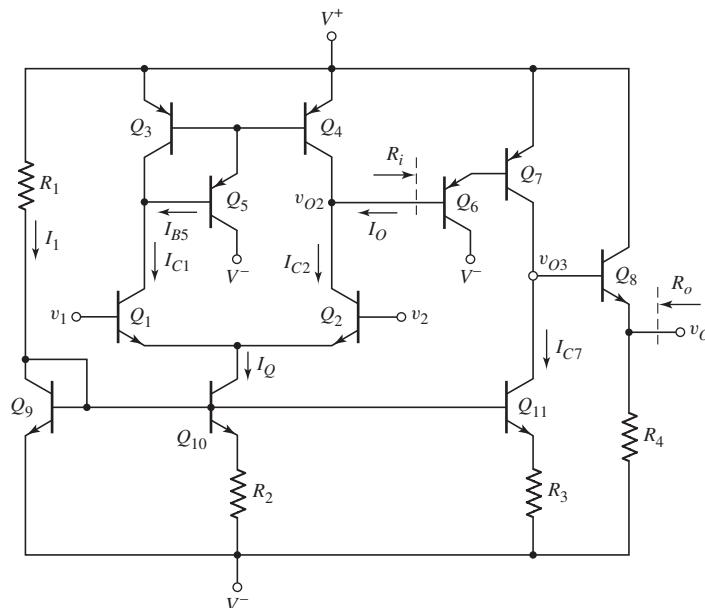


Figure 11.46 BJT diff-amp with three-transistor active load, Darlington pair gain stage, and simple emitter-follower output stage

The differential-pair transistors are biased with a Widlar current source at a bias current I_Q . We noted previously that, for the diff-amp dc currents to be balanced, we must have

$$I_O = I_{B5} = \frac{I_Q}{\beta(1 + \beta)} \quad (11.118)$$

From the figure, we see that

$$I_O = \frac{I_{E6}}{(1 + \beta)} = \frac{I_{C7}}{\beta(1 + \beta)} \quad (11.119)$$

In order for $I_O = I_{B5}$, we must require that $I_{C7} = I_Q$. This means that the emitter resistors of Q_{10} and Q_{11} should have the same value. Transistor Q_{11} also acts as an active load for the Darlington pair gain stage.

Transistor Q_8 and resistor R_4 form the simple emitter-follower output stage. The emitter-follower amplifier minimizes loading effects because its output resistance is small.

Ideally, when the diff-amp input is a pure common-mode signal, the output v_o is zero. The combination of Q_7 and Q_{11} allows the dc level to shift. By slightly changing the bias current I_{C7} , we can vary voltages V_{EC7} and V_{CE11} such that $v_o = 0$. The small variation of I_{C7} required to achieve the necessary dc level shift will not significantly change the balance between I_O and I_{B5} . As we will see in later chapters, other forms of level shifters could also be used.

11.6.2 Input Impedance, Voltage Gain, and Output Impedance

The input resistance of the Darlington pair determines the loading effect on the basic diff-amp. In addition, the gain of the Darlington pair affects the overall gain of the op-amp circuit, and the output resistance of the emitter follower determines any loading effects on the output signal.

Figure 11.47(a) is the ac equivalent circuit of the Darlington pair, where R_{L7} is the effective resistance connected between the collector of Q_7 and signal ground. Figure 11.47(b) shows the simple hybrid- π model of the Darlington pair. We see that the equivalent circuits for Q_6 and Q_7 have been effectively turned upside down compared to the transistors in Figure 11.47(a).

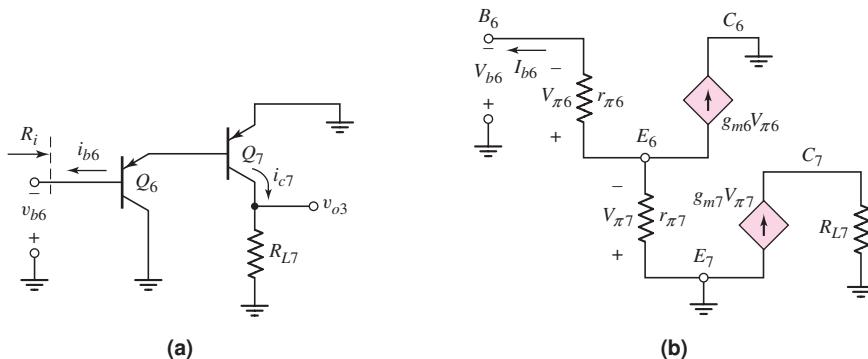


Figure 11.47 (a) The ac equivalent circuit, Darlington pair, and (b) small-signal equivalent circuit, Darlington pair

Writing a KVL equation around the B-E loop of Q_6 and Q_7 , we have

$$V_{b6} = V_{\pi6} + V_{\pi7} \quad (11.120)$$

We can also write that

$$V_{\pi6} = I_{b6}r_{\pi6} \quad (11.121)$$

and the KCL equation is

$$\frac{V_{\pi 7}}{r_{\pi 7}} = \frac{V_{\pi 6}}{r_{\pi 6}} + g_{m6} V_{\pi 6} \quad (11.122(a))$$

or

$$V_{\pi 7} = r_{\pi 7} \left[\frac{(1 + \beta)}{r_{\pi 6}} \right] V_{\pi 6} = r_{\pi 7}(1 + \beta) I_{b6} \quad (11.122(b))$$

where $r_{\pi 6} g_{m6} = \beta$. Substituting Equations (11.122(b)) and (11.121) into Equation (11.120), we obtain

$$V_{b6} = I_{b6} r_{\pi 6} + r_{\pi 7}(1 + \beta) I_{b6} \quad (11.123)$$

The input resistance is therefore

$$R_i = \frac{V_{b6}}{I_{b6}} = r_{\pi 6} + r_{\pi 7}(1 + \beta) \quad (11.124)$$

Assuming $I_{C7} = I_Q$, the hybrid- π parameters are

$$r_{\pi 7} = \frac{\beta V_T}{I_{C7}} = \frac{\beta V_T}{I_Q} \quad (11.125(a))$$

and

$$r_{\pi 6} = \frac{\beta V_T}{I_{C6}} = \frac{(1 + \beta)\beta V_T}{I_Q} \quad (11.125(b))$$

Combining Equations (11.125(a)), (11.125(b)), and Equation (11.124) yields an expression for the input resistance, as follows:

$$R_i = \frac{(1 + \beta)\beta V_T}{I_Q} + \frac{(1 + \beta)\beta V_T}{I_Q} = \frac{2(1 + \beta)\beta V_T}{I_Q} \quad (11.126)$$

We can determine the small-signal voltage gain of the Darlington pair circuit by using the small-signal equivalent circuit in Figure 11.47(b). We see that

$$v_{o3} = i_{c7} R_{L7} = (\beta i_{b7}) R_{L7} = \beta(1 + \beta) i_{b6} R_{L7} \quad (11.127)$$

and

$$i_{b6} = \frac{v_{b6}}{R_i} \quad (11.128)$$

The small-signal voltage gain is therefore

$$A_v = \frac{v_{o3}}{v_{b6}} = \frac{\beta(1 + \beta) R_{L7}}{R_i} \quad (11.129)$$

Substituting Equation (11.126) into (11.129), we find that

$$A_v = \frac{\beta(1 + \beta) R_{L7}}{\frac{2(1 + \beta)\beta V_T}{I_Q}} = \left(\frac{I_Q}{2V_T} \right) R_{L7} \quad (11.130)$$

In Figure 11.46, we see that resistance R_{L7} is the parallel combination of the resistance looking into the collector of Q_{11} and the resistance looking into the base of Q_8 . From Chapter 10, the resistance looking into the collector of Q_{11} is

$$R_{c11} = r_{o11}(1 + g_{m11} R'_E) \quad (11.131)$$

where $R'_E = r_{\pi11} \parallel R_3$. The resistance looking into the base of Q_8 is

$$R_{b8} = r_{\pi8} + (1 + \beta)R_4. \quad (11.132)$$

Equations (11.131) and (11.132) indicate that resistances R_{c11} and R_{b8} are large, which means that the effective resistance R_{L7} is also large.

EXAMPLE 11.13

Objective: Calculate the input resistance and the small-signal voltage gain of a Darlington pair.

Consider the circuit shown in Figure 11.46, with parameters $I_{C7} = I_Q = 0.2 \text{ mA}$, $I_{C8} = 1 \text{ mA}$, $R_4 = 10 \text{ k}\Omega$, and $R_3 = 0.2 \text{ k}\Omega$. Assume $\beta = 100$ for all transistors, and the Early voltage for Q_{11} is 100 V.

Solution: The input resistance, given by Equation (11.126), is

$$R_i = \frac{2(1 + \beta)\beta V_T}{I_Q} = \frac{2(101)(100)(0.026)}{0.2} \Rightarrow 2.63 \text{ M}\Omega$$

The small-signal voltage gain is a function of R_{L7} , which in turn is a function of R_{c11} and R_{b8} . We can find that

$$r_{\pi11} = \beta V_T / I_Q = (100)(0.026) / 0.2 = 13 \text{ k}\Omega$$

such that

$$R'_E = 13 \parallel 0.2 = 0.197 \text{ k}\Omega$$

Also

$$g_{m11} = I_Q / V_T = 0.2 / 0.026 = 7.69 \text{ mA/V}$$

and

$$r_{o11} = V_A / I_Q = 100 / 0.2 = 500 \text{ k}\Omega$$

Therefore,

$$R_{c11} = r_{o11}(1 + g_{m11}R'_E) = 500[1 + (7.69)(0.197)] \Rightarrow 1.26 \text{ M}\Omega$$

We can determine that

$$r_{\pi8} = \beta V_T / I_{C8} = (100)(0.026) / 1 = 2.6 \text{ k}\Omega$$

Then

$$R_{b8} = r_{\pi8} + (1 + \beta)R_4 = 2.6 + (101)(10) \Rightarrow 1.01 \text{ M}\Omega$$

Consequently, resistance R_{L7} is

$$R_{L7} = R_{c11} \parallel R_{b8} = 1.26 \parallel 1.01 = 0.561 \text{ M}\Omega$$

Finally, from Equation (11.130), the small-signal voltage gain is

$$A_v = \left(\frac{I_Q}{2V_T} \right) R_{L7} = \left[\frac{0.2}{2(0.026)} \right] (561) = 2158$$

Comment: The input resistance of the Darlington pair is in the megohm range, which should minimize severe loading effects on the diff-amp. In addition, the small-signal gain is large because of the active load (Q_{11}) and the large input resistance of the emitter-follower output stage.

EXERCISE PROBLEM

Ex 11.13: Consider the Darlington pair Q_6 and Q_7 in Figure 11.46. Determine the current gain of the Darlington pair, I_{c7}/I_{b6} . Use the parameters described in Example 11.13. (Ans. $(101)(100) = 1.01 \times 10^4$)

We can use the results of Chapter 6 to determine the output resistance of the emitter follower. The output resistance is

$$R_o = R_4 \left\| \left(\frac{r_{\pi8} + Z}{(1 + \beta)} \right) \right\| \quad (11.133)$$

where Z is the equivalent impedance, or resistance, in the base of Q_8 . In this case, $Z = R_{c11} \parallel R_{c7}$, where R_{c7} is the resistance looking into the collector of Q_7 . Because of the factor $(1 + \beta)$ in the denominator, the output resistance of the emitter follower is normally small, as previously determined.

EXAMPLE 11.14

Objective: Calculate the output resistance of the circuit in Figure 11.46.

Consider the same circuit and transistor parameters described in Example 11.13. Assume the Early voltage of Q_7 is 100 V.

Solution: From Example 11.13, we have that $R_{c11} = 1.26 \text{ M}\Omega$ and $r_{\pi8} = 2.6 \text{ k}\Omega$. We can then determine that

$$R_{c7} = \frac{V_A}{I_Q} = \frac{100}{0.2} = 500 \text{ k}\Omega$$

Then,

$$Z = R_{c11} \parallel R_{c7} = 1260 \parallel 500 = 358 \text{ k}\Omega$$

Therefore,

$$R_o = R_4 \left\| \left(\frac{r_{\pi8} + Z}{(1 + \beta)} \right) \right\| = 10 \left\| \left(\frac{2.6 + 358}{101} \right) \right\| = 2.63 \text{ k}\Omega$$

Comment: The output resistance is obviously less than R_4 and is substantially less than the equivalent resistance Z in the base of Q_8 . In a later chapter, we will examine a Darlington pair emitter-follower output stage in which the output resistance is on the order of 100 Ω .

EXERCISE PROBLEM

Ex 11.14: The circuit shown in Figure 11.48 is an ac equivalent circuit of a Darlington pair output stage. Assume the transistor current gains are $\beta_A = 90$ and $\beta_B = 180$. Assuming transistor Q_B is biased at $I_{CQB} = 0.5$ mA, determine the output resistance R_o . (Ans. $R_o = 120 \Omega$)

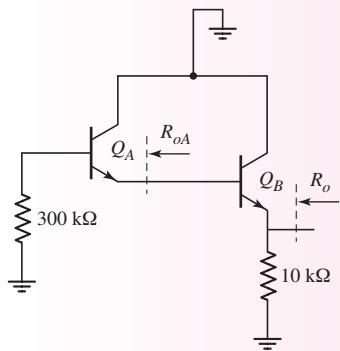


Figure 11.48 Figure for Exercise Ex11.14

A BJT diff-amp with an active load can produce a small-signal differential-mode voltage gain on the order of 10^3 , and the Darlington pair can also provide a voltage gain on the order of 10^3 . Since the emitter follower has a gain of essentially unity, the overall voltage gain of the op-amp circuit is on the order of 10^6 . This value is typical for the low-frequency, open-loop gain of op-amp circuits.

Test Your Understanding

TYU 11.18 Consider the Darlington pair and emitter-follower portions of the circuit in Figure 11.46. The parameters are: $I_{C7} = I_Q = 0.5$ mA, $I_{C8} = 2$ mA, $R_4 = 5$ k Ω , and $R_3 = 0.1$ k Ω . For all transistors, the current gain is $\beta = 120$, and for Q_{11} and Q_7 , the Early voltage is $V_A = 120$ V. Calculate the input resistance and small-signal voltage gain of the Darlington pair, and the output resistance of the emitter follower. (Ans. $R_i = 1.51$ M Ω , $A_v = 3115$, $R_o = 1.14$ k Ω)

TYU 11.19 In the circuit in Figure 11.46, the Darlington pair and emitter-follower transistor parameters are the same as in Exercise TYU 11.18. Determine the effective resistance R_{L7} (see Figure 11.47(a)) such that the small-signal voltage gain is 10^3 . (Ans. $R_{L7} = 104$ k Ω)



11.7 SIMPLIFIED BJT OPERATIONAL AMPLIFIER CIRCUIT

Objective: • Analyze a simplified multistage bipolar amplifier.

An operational amplifier (op-amp) is a multistage circuit composed of a differential amplifier input stage, a gain stage, and an output stage. In this section, we will consider a simplified BJT op-amp circuit.

Although active load devices increase the gain of an amplifier, in this discussion, we will consider resistive loads, in order to simplify the analysis and design. For the bipolar circuit, all component values are given; we will analyze both the dc and ac circuit characteristics.

Figure 11.49. depicts a simple bipolar operational amplifier. The differential amplifier stage is biased with a Widlar current source, and a one-sided output is connected to the Darlington pair gain stage. An emitter bypass capacitor C_E is included to increase the small-signal voltage gain. The output stage is an emitter follower. In general, we want the dc value of the output voltage to be zero when the input voltage is zero. To accomplish this, we need to insert a dc level shifting circuit between the voltage v_{O3} and the output voltage v_O .

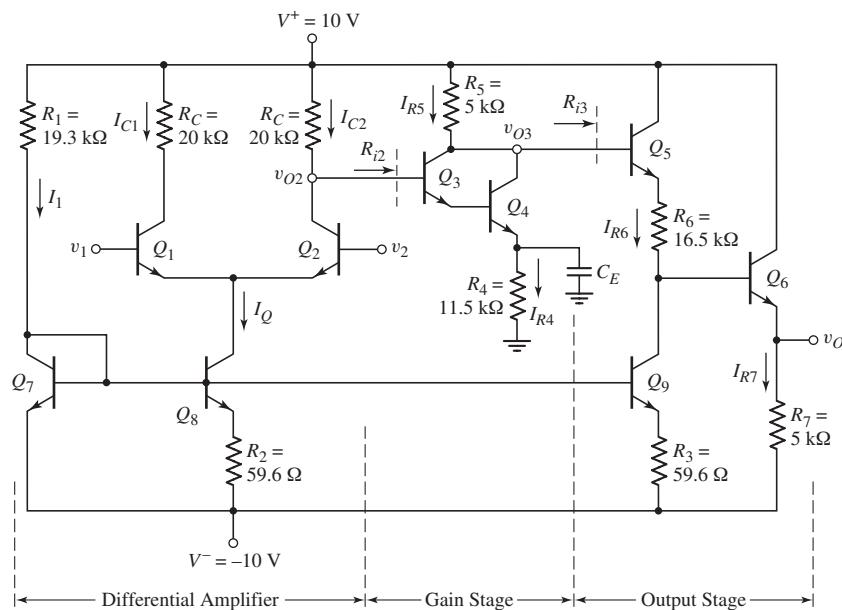


Figure 11.49 Bipolar operational amplifier circuit

EXAMPLE 11.15

Objective: Analyze the dc characteristics of the bipolar op-amp circuit.

Consider the circuit in Figure 11.49. Neglect base currents and, as a simplification, assume $V_{BE(on)} = 0.7$ V for all transistors except Q_8 and Q_9 in the Widlar circuit.

Solution: The reference current I_1 is

$$I_1 = \frac{10 - 0.7 - (-10)}{19.3} = 1 \text{ mA}$$

The bias current I_Q is determined from

$$I_Q R_2 = V_T \ln \left(\frac{I_1}{I_Q} \right)$$

and is

$$I_Q = 0.4 \text{ mA}$$

The collector currents are then

$$I_{C1} = I_{C2} = 0.2 \text{ mA}$$

The dc voltage at the collector of Q_2 is

$$V_{O2} = 10 - I_{C2}R_C = 10 - (0.2)(20) = 6 \text{ V}$$

With these circuit parameters, the common-mode input voltage is limited to the range $-8.6 \leq v_{CM} \leq 6 \text{ V}$, which will keep all transistors biased in the forward-active mode.

The current I_{R4} is determined to be

$$I_{R4} = \frac{V_{O2} - 2V_{BE}(\text{on})}{R_4} = \frac{6 - 1.4}{11.5} = 0.4 \text{ mA}$$

Since base currents are assumed negligible, the current I_{R5} is $I_{R5} \cong I_{R4}$.

The dc voltage at the collectors of Q_3 and Q_4 is then

$$V_{O3} = 10 - I_{R5}R_5 = 10 - (0.4)(5) = 8 \text{ V}$$

This shows us that the dc voltage V_{O3} is midway between the 10 V supply voltage and the dc input voltage $V_{O2} = 6 \text{ V}$ to Q_3 . This allows a maximum symmetrical swing in the time-varying voltage at v_{o3} .

Transistor Q_5 and resistor R_6 form the dc voltage level shifting function. Since $R_3 = R_2$, we have

$$I_{R6} = I_Q = 0.4 \text{ mA}$$

The dc voltage at the base of Q_6 is found to be

$$V_{B6} = V_{O3} - V_{BE}(\text{on}) - I_{R6}R_6 = 8 - 0.7 - (0.4)(16.5) = 0.7 \text{ V}$$

This relationship produces a zero dc output voltage when a zero differential-mode voltage is applied at the input.

Finally, current I_{R7} is

$$I_{R7} = \frac{v_o - (-10)}{R_7} = \frac{10}{5} = 2 \text{ mA}$$

Comment: The dc analysis of this simplified op-amp circuit proceeds in much the same way as in previous examples. We observe that all transistors are biased in the forward-active mode.

EXERCISE PROBLEM

Ex 11.15: Consider the simple bipolar op-amp circuit in Figure 11.49. The transistor parameters are: $\beta = 100$, $V_{BE}(\text{on}) = 0.7 \text{ V}$ (except for Q_8 and Q_9), and $V_A = \infty$. Redesign the circuit such that $I_{C1} = I_{C2} = 0.1 \text{ mA}$, $I_{R7} = 5 \text{ mA}$, $I_1 = I_{R4} = I_{R6} = 0.6 \text{ mA}$, $V_{CE1} = V_{CE2} = 4 \text{ V}$, $V_{CE4} = 3 \text{ V}$, and $v_o = 0$. (Ans. $R_1 = 32.2 \text{ k}\Omega$, $R_2 = 143 \Omega$, $R_3 = 0$, $R_C = 67 \text{ k}\Omega$, $R_4 = 3.17 \text{ k}\Omega$, $R_5 = 8.5 \text{ k}\Omega$, $R_6 = 5.83 \text{ k}\Omega$, and $R_7 = 2 \text{ k}\Omega$)

EXAMPLE 11.16

Objective: Determine the small-signal differential-mode voltage gain of the bipolar op-amp circuit.

Consider the circuit in Figure 11.49, with transistor parameters $\beta = 100$ and $V_A = \infty$.

Solution: The overall differential-mode voltage gain can be written

$$A_d = A_{d1} \cdot A_2 \cdot A_3 = \left(\frac{v_{o2}}{v_1 - v_2} \right) \cdot \left(\frac{v_{o3}}{v_{o2}} \right) \cdot \left(\frac{v_o}{v_{o3}} \right)$$

The overall small-signal voltage gain is the product of the individual stage gains *only* if the load resistance of the following stage is taken into account.

We will rely on previous results to determine the individual voltage gains. The input resistances to the Darlington pair R_{i2} and to the output stage R_{i3} are indicated in Figure 11.49. The one-sided differential-mode voltage gain of the diff-amp is given by

$$A_{d1} = \frac{V_{o2}}{v_d} = \frac{g_m}{2} (R_C \parallel R_{i2})$$

where R_{i2} is the input resistance of the Darlington pair, as follows:

$$R_{i2} = r_{\pi3} + (1 + \beta)r_{\pi4}$$

where

$$r_{\pi4} = \beta V_T / I_{R4} = (100)(0.026)/0.4 = 6.5 \text{ k}\Omega$$

and

$$r_{\pi3} \cong \beta^2 V_T / I_{R4} = (100)^2(0.026)/0.4 = 650 \text{ k}\Omega$$

Therefore,

$$R_{i2} = 650 + (101)(6.5) = 1307 \text{ k}\Omega$$

The transistor transconductance is

$$g_m = \frac{I_Q}{2V_T} = \frac{0.4}{2(0.026)} = 7.70 \text{ mA/V}$$

The gain of the differential amplifier stage is therefore

$$A_{d1} = \frac{g_m}{2} (R_C \parallel R_{i2}) = \left(\frac{7.70}{2} \right) [20 \parallel 1307] = 75.8$$

Since the load resistance $R_{i2} \gg R_C$, there is no significant loading effect of the second stage on the diff-amp stage.

From previous results, we know the voltage gain of the Darlington pair is given by

$$A_2 = \left(\frac{I_{R4}}{2V_T} \right) (R_5 \parallel R_{i3})$$

where

$$R_{i3} = r_{\pi5} + (1 + \beta)[R_6 + r_{\pi6} + (1 + \beta)R_7]$$

We find that

$$r_{\pi5} = \beta V_T / I_{R6} = (100)(0.026)/0.4 = 6.5 \text{ k}\Omega$$

and

$$r_{\pi6} = \beta V_T / I_{R7} = (100)(0.026)/2 = 1.3 \text{ k}\Omega$$

Therefore

$$R_{i3} = 6.5 + (101)[16.5 + 1.3 + (101)(5)] \Rightarrow 52.8 \text{ M}\Omega$$

Since $R_{i3} \gg R_5$, the output stage does not load down the gain stage, and the small-signal voltage gain is approximately

$$A_2 \cong \left(\frac{I_{R4}}{2V_T} \right) R_5 = \left[\frac{0.4}{2(0.026)} \right] (5) = 38.5$$

The combination of Q_5 and Q_6 forms an emitter follower, and the gain of the output stage is

$$A_3 = v_o/v_{o3} \approx 1$$

The overall small-signal voltage gain is therefore

$$A_d = A_{d1} \cdot A_2 \cdot A_3 = (75.8)(38.5)(1) = 2918$$

Comment: From our previous discussion, we know that the overall gain can be increased substantially by using active loads. Yet, the analysis of this simplified circuit provides some insight into the design of multistage circuits, as well as the overall small-signal voltage gain of op-amp circuits.

Computer Correlation: A PSpice analysis was performed on the bipolar op-amp circuit in Figure 11.49. The dc output voltage from this analysis was $V_O = -0.333$ V, rather than the desired value of zero. This occurred because the B-E voltages were not exactly 0.7 V, as assumed in the hand analysis. A zero output voltage can be obtained by slightly adjusting R_6 . The differential voltage gain was $A_d = 2932$, which agrees very well with the hand analysis.

EXERCISE PROBLEM

Ex 11.16: Consider the simple bipolar op-amp circuit in Figure 11.49 with circuit and transistor parameters given in Exercise Problem Ex11.15. Determine the input resistances R_{i2} and R_{i3} , and the differential-mode voltage gain $A_d = v_o/v_d$. (Ans. $R_{i2} = 870$ k Ω , $R_{i3} = 21.0$ M Ω , $A_d = 11,729$)

Problem-Solving Technique: Multistage Circuits

1. Perform the dc analysis of the circuit to determine the small-signal parameters of the transistors. In most cases BJT base currents can be neglected. This assumption will normally provide sufficient accuracy for a hand analysis.
2. Perform the ac analysis on each stage of the circuit, *taking into account the loading effect of the following stage*. (In many cases, previous results of small-signal analyses can be used directly.)
3. The overall small-signal voltage gain or current gain is the product of the gains of the individual stages *as long as the loading effect of each stage is taken into account*.



11.8 DIFF-AMP FREQUENCY RESPONSE

Objective: • Analyze the frequency response of the differential amplifier.

In Chapter 7, we considered the frequency responses of the three basic amplifier configurations. In this section, we will analyze the frequency response of the differential amplifier. Since the diff-amp is a linear circuit, we can determine the frequency

response due to: (a) a pure differential-mode input signal, (b) a pure common-mode input signal, and (c) the total or net result, using superposition.

11.8.1 Due to Differential-Mode Input Signal

Consider the basic bipolar diff-amp shown in Figure 11.50(a). The input is a pure differential-mode input signal. We know from Equation (11.24) that the small-signal voltage v_e is at signal ground when a differential-mode input signal is applied. To determine the frequency response, we evaluate the equivalent common-emitter half-circuit in Figure 11.50(b).

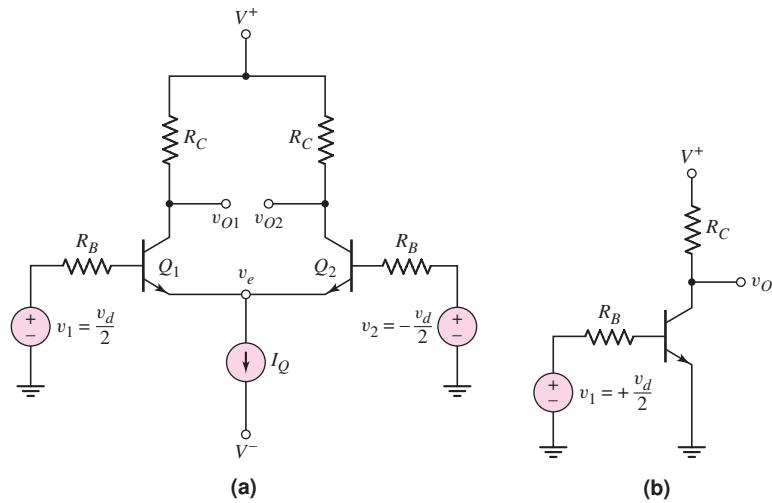


Figure 11.50 (a) BJT differential amplifier with differential-mode input signal and (b) equivalent common-emitter half-circuit of differential amplifier

Since the diff-amp is a direct-coupled amplifier, the midband voltage gain extends to zero frequency. This one-sided midband gain is

$$A_{v1} = \frac{V_{o1}}{V_d/2} = -g_m R_C \left(\frac{r_\pi}{r_\pi + R_B} \right) \quad (11.134(a))$$

or

$$A_{v1} = \frac{-\beta R_C}{r_\pi + R_B} \quad (11.134(b))$$

From the high-frequency common-emitter characteristics determined in Chapter 7 we know that the upper 3 dB frequency is

$$f_H = \frac{1}{2\pi[r_\pi \| R_B](C_\pi + C_M)} \quad (11.135)$$

where C_M is the equivalent Miller capacitance given by

$$C_M = C_\mu(1 + g_m R_C) \quad (11.136)$$

Equation (11.136) implies that, if the value of R_C is fairly large, the Miller capacitance will significantly affect the bandwidth of the differential amplifier.

11.8.2 Due to Common-Mode Input Signal

Figure 11.51(a) shows the basic diff-amp with a pure common-mode input signal. The circuit is symmetrical, which means that resistors R_B , resistors R_C , and the transistors are effectively in parallel. Figure 11.51(b) is the small-signal equivalent circuit, with the constant-current source replaced by its output resistance R_o and capacitance C_o .

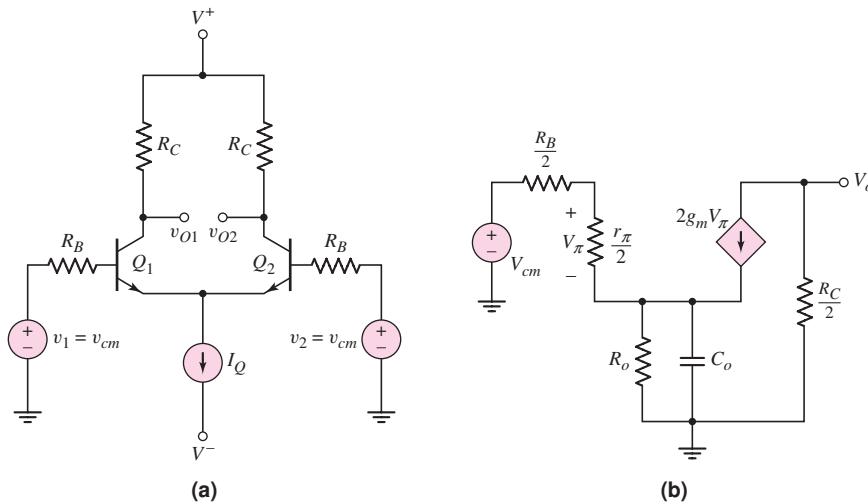


Figure 11.51 (a) BJT differential amplifier with common-mode input signal and (b) small-signal equivalent circuit, common-mode configuration

We will justify neglecting the transistor parameters C_π and C_μ . The output voltage is

$$V_o = -(2g_m V_\pi) \left(\frac{R_C}{2} \right) \quad (11.137)$$

A KVL equation around the B-E loop produces

$$V_{cm} = \left(\frac{V_\pi}{r_\pi/2} \right) \left(\frac{R_B}{2} \right) + V_\pi + \left(\frac{V_\pi}{r_\pi/2} + 2g_m V_\pi \right) \left[R_o \parallel \left(\frac{1}{sC_o} \right) \right] \quad (11.138(a))$$

or

$$V_{cm} = V_\pi \left\{ \frac{R_B}{r_\pi} + 1 + 2 \left(\frac{1+\beta}{r_\pi} \right) \left(\frac{R_o}{1+sR_oC_o} \right) \right\} \quad (11.138(b))$$

Solving for V_π and substituting the result into Equation (11.137) yields the common-mode gain, which is

$$A_{cm} = \frac{V_o}{V_{cm}} = \frac{-g_m R_C}{\frac{R_B}{r_\pi} + 1 + \frac{2(1+\beta)}{r_\pi} \left(\frac{R_o}{1+sR_oC_o} \right)} \quad (11.139(a))$$

or

$$A_{cm} = \frac{-g_m R_C (1 + s R_o C_o)}{\left(1 + \frac{R_B}{r_\pi}\right) (1 + s R_o C_o) + \frac{2(1 + \beta) R_o}{r_\pi}} \quad (11.139(b))$$

Equation (11.139(b)) shows that there is a zero in the common-mode gain. To explain, capacitor C_o is in parallel with R_o , and it acts as a bypass capacitor. At very low frequency, C_o is effectively an open circuit and the common-mode signal “sees” R_o . As the frequency increases, the impedance of the capacitor decreases and R_o is effectively bypassed; hence, the zero in Equation (11.139(b)). The frequency analysis of an emitter bypass capacitor also showed the presence of a zero in the voltage gain expression.

The common-mode gain frequency response is shown in Figure 11.52. The frequency of the zero is

$$f_z = \frac{1}{2\pi R_o C_o} \quad (11.140)$$

Since the output resistance R_o of a constant-current source is normally large, a small capacitance C_o can result in a small f_z . For frequencies greater than f_z , the common-mode gain increases at the rate of 6 dB/octave.

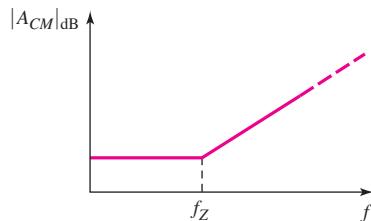


Figure 11.52 Frequency response of common-mode gain

Equation (11.139(b)) also shows that there is a pole associated with the common-mode gain. Rearranging the terms in that equation, we see that the frequency of the pole is

$$f_p = \frac{1}{2\pi R_{eq} C_o} \quad (11.141)$$

where

$$R_{eq} = \frac{R_o \left(1 + \frac{R_B}{r_\pi}\right)}{1 + \frac{R_B}{r_\pi} + \frac{2(1 + \beta) R_o}{r_\pi}} \quad (11.142)$$

The denominator of Equation (11.142) is very large, because of the term $(1 + \beta) R_o$. This implies that R_{eq} is small, which means that the frequency f_p of the pole is very large.

The differential-mode gain is shown in Figure 11.53. The frequency response of the common-mode rejection ratio is found by combining Figures 11.52 and 11.53, and is shown in Figure 11.54.

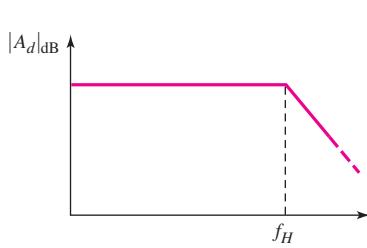


Figure 11.53 Frequency response of differential-mode gain

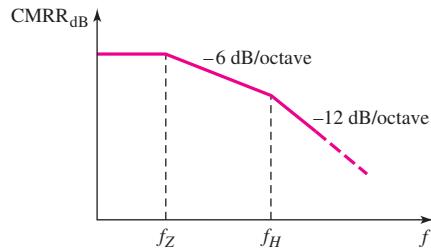


Figure 11.54 Frequency response of common-mode rejection ratio

EXAMPLE 11.17

Objective: Determine the zero and pole frequencies in the common-mode gain.

Consider a diff-amp biased with a constant-current source. The output resistance is $R_o = 10 \text{ M}\Omega$ and the output capacitance is $C_o = 1 \text{ pF}$. Assume the circuit and transistor parameters are $R_B = 0.5 \text{ k}\Omega$, $r_\pi = 10 \text{ k}\Omega$, and $\beta = 100$.

Solution: In the common-mode gain, the frequency of the zero is

$$f_z = \frac{1}{2\pi R_o C_o} = \frac{1}{2\pi(10 \times 10^6)(1 \times 10^{-12})} \Rightarrow 15.9 \text{ kHz}$$

Also in the common-mode gain, the frequency of the pole is

$$f_p = 1/(2\pi R_{eq} C_o)$$

where

$$R_{eq} = \frac{R_o \left(1 + \frac{R_B}{r_\pi}\right)}{1 + \frac{R_B}{r_\pi} + \frac{2(1 + \beta)R_o}{r_\pi}} = \frac{(10 \times 10^6) \left(1 + \frac{0.5}{10}\right)}{1 + \frac{0.5}{10} + \frac{2(101)(10 \times 10^6)}{10 \times 10^3}}$$

or

$$R_{eq} = 51.98 \Omega$$

The frequency of the pole is therefore

$$f_p = \frac{1}{2\pi(51.98)(1 \times 10^{-12})} \Rightarrow 3.06 \text{ GHz}$$

Comment: The frequency of the zero in the common-mode gain is fairly low, while the frequency of the pole is extremely large. The relatively low frequency of the zero justifies neglecting the effect of C_π and C_μ . The CMRR frequency response is shown in Figure 11.54, where f_z is the zero frequency of the common-mode gain and f_H is the upper 3 dB frequency of the differential-mode gain.

EXERCISE PROBLEM

Ex 11.17: Repeat Example 11.17 for the case when the output capacitance of the constant current source is $C_o = 0.2 \text{ pF}$. (Ans. $f_z = 79.6 \text{ kHz}$, $f_p = 15.3 \text{ GHz}$)

11.8.3 With Emitter-Degeneration Resistors

Figure 11.55 shows a bipolar diff-amp with two resistances R_E connected in the emitter portion of the circuit. One effect of including an emitter resistor is to reduce the voltage gain, so the presence of these resistors is termed **emitter degeneration**.

In Chapter 7, we found that an emitter-follower circuit, which includes an emitter resistance, is a wide-bandwidth amplifier. Therefore, one effect of resistors R_E is an increase in the bandwidth of the differential amplifier. We rely on a computer simulation to evaluate emitter degeneration effects.

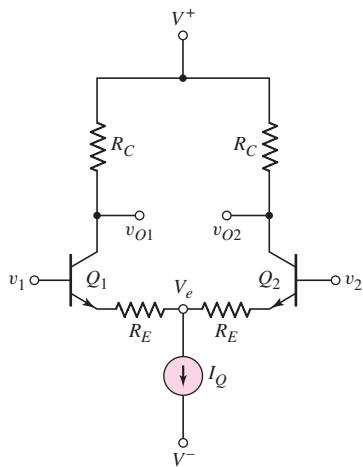


Figure 11.55 BJT differential amplifier with emitter-degeneration resistors

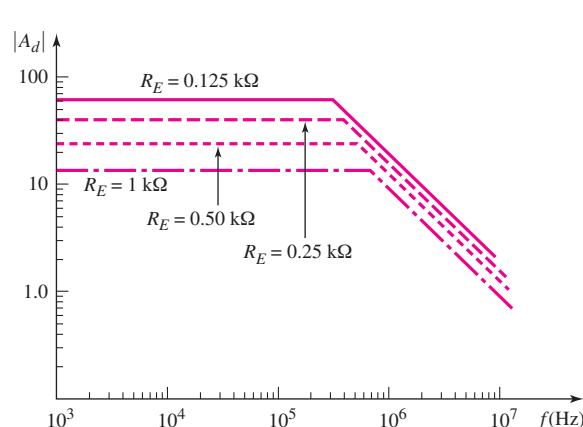


Figure 11.56 PSpice results for frequency response of diff-amp with emitter-degeneration

Figure 11.56 shows the frequency response of a one-sided differential-mode gain, obtained from a PSpice analysis for four R_E resistance values. The diff-amp is biased at $I_Q = 0.5$ mA and the R_C resistors are $R_C = 30$ k Ω . The transistor capacitances are $C_\pi = 34.6$ pF and $C_\mu = 4.3$ pF. As the emitter degeneration increases, the differential-mode voltage gain decreases, but the bandwidth increases, as previously indicated. The figure-of-merit for amplifiers, the gain-bandwidth product, is approximately a constant for the results shown in Figure 11.56.

11.8.4 With Active Load

Figure 11.57 shows a bipolar diff-amp with an active load and a single input at v_1 . The base and collector junctions of Q_3 are connected together, and a one-sided output is taken at v_{O2} .

With the connection of Q_3 , the equivalent load resistance in the collector of Q_1 is on the order of $r_\pi/(1 + \beta)$. This small resistance minimizes the Miller multiplication factor in Q_1 . Also, with the base of Q_2 at ground potential, one side of $C_{\mu 2}$ is grounded, and the Miller multiplication in Q_2 is zero. Therefore, we expect the bandwidth of the diff-amp with an active load to be relatively wide. At high frequencies, however, the effective impedance in the collector of Q_1 also includes the input

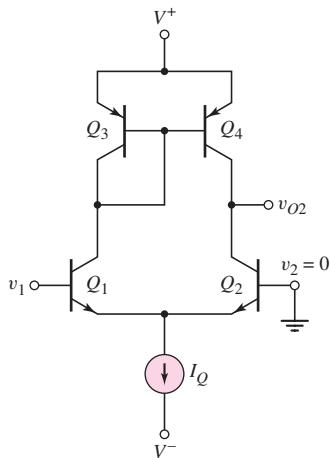


Figure 11.57 BJT diff-amp with active load and single-sided input

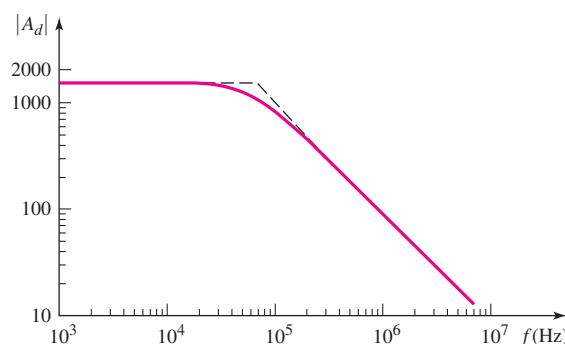


Figure 11.58 PSpice results for frequency response of diff-amp with active load and single-sided input

capacitances of Q_3 and Q_4 . These additional capacitances also affect the frequency response of the diff-amp, potentially narrowing the bandwidth.

Again, we rely on a computer analysis to determine the frequency characteristics of the diff-amp with an active load. Figure 11.58 shows the results of the computer simulation. The diff-amp is biased at $I_Q = 0.5$ mA, and the Early voltage of each transistor is assumed to be 80 V. The transistor capacitances are $C_\pi = 34.6$ pF for each transistor, $C_\mu = 3.8$ pF in Q_1 and Q_2 , and $C_\mu = 7$ pF and 5.5 pF in Q_3 and Q_4 , respectively.

The low-frequency voltage gain is 1560 and the upper 3 dB frequency is 64 kHz. The large gain is as expected for an active load amplifier, but the 3 dB frequency is lower than expected. However, the gain-bandwidth product for the active load diff-amp is approximately four times that of the diff-amp shown in Figure 11.55. The increased gain-bandwidth product implies a reduced Miller multiplication factor in the active load diff-amp, as predicted.



11.9 DESIGN APPLICATION: A CMOS DIFF-AMP

Objective: • Design a CMOS diff-amp with an output gain stage to meet a set of specifications.

Specifications: Design a CMOS diff-amp with an output stage. The magnitude of voltage gain of each stage is to be at least 600. Bias currents are to be $I_Q = I_{\text{REF}} = 100 \mu\text{A}$, and biasing of the circuit is to be $V^+ = 2.5$ V and $V^- = -2.5$ V.

Design Approach: The circuit to be designed has the configuration shown in Figure 11.59. The diff-amp has NMOS amplifying transistors and a PMOS active load. The diff-amp is biased with a cascode current source to provide a large output resistance. The gain stage is a PMOS transistor in a common source configuration that also has an active load.

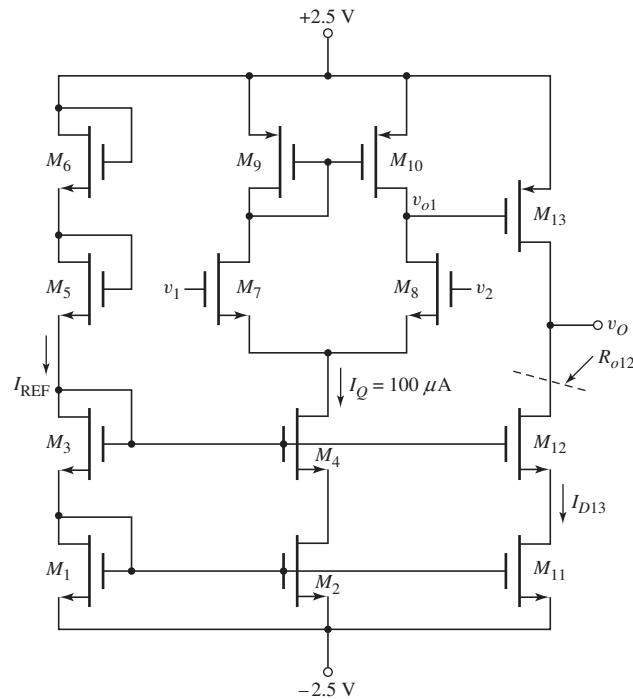


Figure 11.59 A CMOS diff-amp with an output stage for the design application

We will assume that several sets of transistors are matched. In particular, we will assume that M_1 to M_4 , M_{11} , and M_{12} are matched; M_7 and M_8 are matched; M_5 and M_6 are matched; and M_9 and M_{10} are matched.

Choices: Assume NMOS and PMOS transistors are available with parameters $V_{TN} = 0.5$ V, $V_{TP} = -0.5$ V, $k'_n = 80 \mu\text{A}/\text{V}^2$, $k'_p = 40 \mu\text{A}/\text{V}^2$, and $\lambda_n = \lambda_p = 0.01 \text{ V}^{-1}$.

Solution (Differential Pair): The differential gain of the diff-amp is given by

$$A_d = g_m(r_{o8} \| r_{o10})$$

We find

$$r_{o8} = r_{o10} = \frac{1}{\lambda I_D} = \frac{1}{(0.01)(0.05)} = 2000 \text{ k}\Omega$$

Then, for $A_d = 600$, we have

$$600 = g_m(2000 \| 2000)$$

which yields $g_m = 0.6 \text{ mA/V}$. Then

$$g_m = 2\sqrt{\frac{k'_n}{2} \frac{W}{L} I_D} = 0.6 = 2\sqrt{\left(\frac{0.08}{2}\right)\left(\frac{W}{L}\right)_7} \quad (0.05)$$

which yields

$$\left(\frac{W}{L}\right)_7 = \left(\frac{W}{L}\right)_8 = 45$$

We will also, somewhat arbitrarily, make the width-to-length ratios of all other transistors, except M_5 , M_6 , and M_{13} , the same value of 45.

Solution (Current Source): We need to consider the two transistors M_5 and M_6 . The gate-to-source voltages of M_1 and M_3 are found from

$$I_{\text{REF}} = 100 = \frac{80}{2}(45)(V_{GS1} - 0.5)^2$$

which yields $V_{GS1} = V_{GS3} = 0.736$ V. Since M_5 and M_6 are matched, we find

$$V_{GS5} = V_{GS6} = \frac{2.5 - (-2.5) - 2(0.736)}{2} = 1.76 \text{ V}$$

The width-to-length ratios of M_5 and M_6 are found from

$$I_{\text{REF}} = 100 = \frac{80}{2} \left(\frac{W}{L} \right)_{5,6} (1.76 - 0.5)^2$$

which yields

$$\left(\frac{W}{L} \right)_5 = \left(\frac{W}{L} \right)_6 = 1.57$$

Solution (Second Stage): The source-to-gate voltage applied to the common-source transistor M_{13} is equal to the source-to-drain voltage on M_{10} which is the same as the source-to-gate voltage of M_9 since the diff-amp is balanced. We find

$$I_{D9} = 50 = \frac{40}{2}(45)(V_{SG9} - 0.5)^2$$

or

$$V_{SG9} = V_{SG13} = 0.736 \text{ V}$$

The drain current in M_{13} is $I_{D13} = I_Q = 100 \mu\text{A}$ because of the matched transistors in the current source circuit. We now find

$$\frac{I_{D13}}{I_{D9}} = \frac{(W/L)_{13}}{(W/L)_9} \Rightarrow \frac{100}{50} = \frac{(W/L)_{13}}{45}$$

which yields $(W/L)_{13} = 90$. The gain of the second stage is given by

$$A_2 = -g_{m13}(r_{o13} \| R_{o12})$$

We find

$$g_{m13} = 2\sqrt{\frac{k'_p}{2} \left(\frac{W}{L} \right)_{13} I_{D13}} = 2\sqrt{\left(\frac{0.04}{2} \right)(90)(0.1)} = 0.849 \text{ mA/V}$$

and

$$r_{o13} = \frac{1}{\lambda I_{D13}} = \frac{1}{(0.01)(0.1)} = 1000 \text{ k}\Omega$$

The output resistance R_{o12} is given by

$$R_{o12} = r_{o12} + r_{o11}(1 + g_{m12}r_{o12})$$

We find

$$r_{o11} = r_{o12} = \frac{1}{\lambda I_{D12}} = \frac{1}{(0.01)(0.1)} = 1000 \text{ k}\Omega$$

and

$$g_{m12} = 2\sqrt{\frac{k'_n}{2} \left(\frac{W}{L}\right)_{12} I_{D12}} = 2\sqrt{\left(\frac{0.08}{2}\right)(45)(0.1)} = 0.849 \text{ mA/V}$$

Then

$$R_{o12} = 1000 + 1000[1 + (0.849)(1000)] = 851,000 \text{ k}\Omega$$

The second stage voltage gain is then

$$A_2 = -0.849(1000\parallel 851,000) = -849$$

Solution (Overall Voltage Gain): Since there is no loading of the second stage on the diff-amp circuit, the overall voltage gain is

$$A_v = A_d A_2 = (600)(-849) = -5.094 \times 10^5$$

Comment: We may note that the amplifier we have just designed is an all MOSFET circuit. The circuit contains no resistors. An all-transistor circuit is one of the advantages of MOS transistors.

We may also note that a large voltage gain can be obtained from a circuit using active loads.

11.10 SUMMARY

- The ideal transistor differential amplifier amplifies only the difference between two input signals.
- The differential-mode input voltage is defined as the difference between the two input signals and the common-mode input voltage is defined as the average of the two input signals.
- When a differential input voltage is applied, one transistor of the differential pair turns on more than the second transistor of the differential pair so that the currents become unbalanced, producing a signal output voltage.
- A common-mode output signal is generated when the output resistance of the current source is finite rather than ideally infinite.
- The common-mode rejection ratio, CMRR, is defined in terms of decibels as $\text{CMRR}_{\text{dB}} = 20 \log_{10} |A_d/A_{cm}|$, where A_d and A_{cm} are the differential-mode voltage gain and common-mode voltage gain, respectively.
- Differential amplifiers are usually designed with active loads to increase the differential-mode voltage gain.
- BiCMOS circuits may be designed to incorporate the best parameters and characteristics of BJTs and MOSFETs in the same circuit.
- A BJT Darlington pair is typically used as a second stage in a BJT diff-amp. The input impedance is large, which tends to minimize loading effects on the diff-amp, and the effective current gain of the pair is the product of the individual gains.
- As an application, a CMOS diff-amp with an output gain stage was designed to meet a set of specifications.

 **CHECKPOINT**

After studying this chapter, the reader should have the ability to:

- ✓ Describe the mechanism by which a differential-mode signal and common-mode signal are produced in a BJT diff-amp.
- ✓ Describe the dc transfer characteristics of a BJT diff-amp.
- ✓ Define common-mode rejection ratio.
- ✓ Describe the mechanism by which a differential-mode signal and common-mode signal are produced in a MOSFET diff-amp.
- ✓ Describe the dc transfer characteristics of a MOSFET diff-amp.
- ✓ Design a MOSFET diff-amp with an active load to yield a specified differential-mode voltage gain.
- ✓ Analyze a simplified BJT operational amplifier circuit.
- ✓ Design a simplified MOSFET operational amplifier circuit.

 **REVIEW QUESTIONS**

1. Define differential-mode and common-mode input voltages.
2. Sketch the dc transfer characteristics of a BJT differential amplifier.
3. From the dc transfer characteristics, qualitatively define the linear region of operation for a differential amplifier.
4. What is meant by matched transistors and why are matched transistors important in the design of diff-amps?
5. Explain how a differential-mode output signal is generated.
6. Explain how a common-mode output signal is generated.
7. Define the common-mode rejection ratio, CMRR. What is the ideal value?
8. What design criteria will yield a large value of CMRR in an emitter-coupled pair?
9. Sketch the differential-mode and common-mode half-circuit models for an emitter-coupled diff-amp.
10. Define differential-mode and common-mode input resistances.
11. Sketch the dc transfer characteristics of a MOSFET differential amplifier.
12. Sketch and describe the advantages of a MOSFET cascode current source used with a MOSFET differential amplifier.
13. Sketch a simple MOSFET differential amplifier with an active load.
14. Explain the advantages of an active load.
15. Describe the loading effects of connecting a second stage to the output of a BJT diff-amp.
16. Explain the frequency response of the differential-mode voltage gain.
17. Sketch a BJT Darlington pair circuit and explain the advantages.
18. Describe the three stages of a simple BJT operational amplifier.

 **PROBLEMS****Section 11.2 Basic BJT Differential Pair**

- 11.1 (a) A differential-amplifier has a differential-mode gain of $A_d = 250$ and a common-mode rejection ratio of $\text{CMRR}_{\text{dB}} = \infty$. A differential-mode input signal of $v_d = 1.5 \sin \omega t \text{ mV}$ is applied along with a common-mode input signal of $v_{cm} = 3 \sin \omega t \text{ V}$. Assuming the common-mode gain is positive, determine the output voltage. (b) Repeat part (a) if the common-mode

rejection ratio is $\text{CMRR}_{\text{dB}} = 80 \text{ dB}$. (c) Repeat part (a) if the common-mode rejection ratio is $\text{CMRR}_{\text{dB}} = 50 \text{ dB}$.

- 11.2 Consider the circuit shown in Figure P11.2. Assume $g_m = 1.0 \text{ mA/V}$. Assume the input signal voltages are $v_1 = 0.7 + 0.1 \sin \omega t \text{ V}$ and $v_2 = 0.7 - 0.1 \sin \omega t \text{ V}$. (a) Determine the signal voltages (i) v_{o1} , (ii) v_{o2} , and (iii) $v_{o1} - v_{o2}$. (b) Using the results of part (a), determine the small-signal voltage gains (i) $A_{d1} = \Delta v_{o1}/\Delta(v_1 - v_2)$, (ii) $A_{d2} = \Delta v_{o2}/\Delta(v_1 - v_2)$, and (iii) $A_{d3} = \Delta(v_{o1} - v_{o2})/\Delta(v_1 - v_2)$.

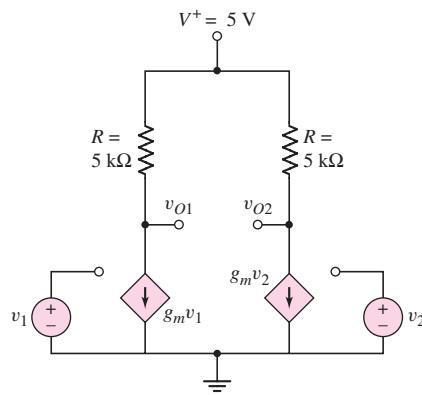


Figure P11.2

- 11.3 Consider the differential amplifier shown in Figure P11.3 with transistor parameters $\beta = 150$, $V_{BE(\text{on})} = 0.7 \text{ V}$, and $V_A = \infty$. (a) Design the circuit such that the Q -point values are $I_{C1} = I_{C2} = 100 \mu\text{A}$ and $v_{o1} = v_{o2} = 1.2 \text{ V}$ for $v_1 = v_2 = 0$. (b) Draw the dc load line and plot the Q -point for transistor Q_2 . (c) What are the maximum and minimum values of the common-mode input voltage?

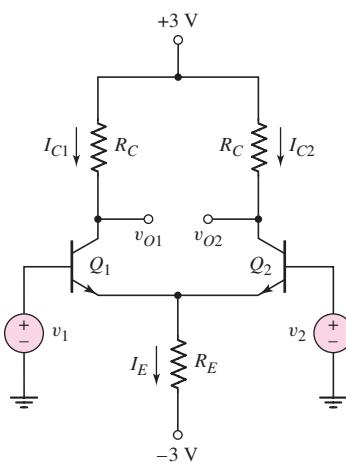


Figure P11.3

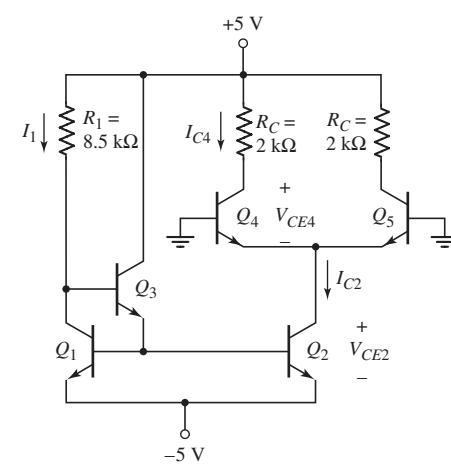


Figure P11.4

- 11.4 The differential amplifier in Figure P11.4 is biased with a three-transistor current source. The transistor parameters are: $\beta = 100$, $V_{BE(\text{on})} = 0.7 \text{ V}$, and

- $V_A = \infty$. (a) Determine I_1 , I_{C2} , I_{C4} , V_{CE2} , and V_{CE4} . (b) Determine a new value of R_1 such that $V_{CE4} = 2.5$ V. What are the values of I_{C4} , I_{C2} , I_1 , and R_1 ?
- *D11.5 For the transistors in the circuit in Figure P11.5, the parameters are $\beta = 100$ and $V_{BE(on)} = 0.7$ V. The Early voltage is $V_A = \infty$ for Q_1 and Q_2 , and is $V_A = 50$ V for Q_3 and Q_4 . (a) Design resistor values such that $I_3 = 400 \mu\text{A}$ and $V_{CE1} = V_{CE2} = 10$ V. (b) Find A_d , A_{cm} , and CMRR_{dB} for a one-sided output at v_{O2} . (c) Determine the differential- and common-mode input resistances.

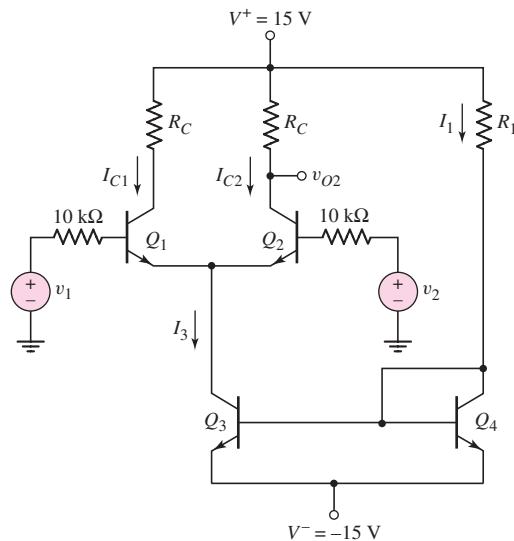


Figure P11.5

- 11.6 The diff-amp in Figure 11.3 of the text has parameters $V^+ = +5$ V, $V^- = -5$ V, $R_C = 8 \text{ k}\Omega$, and $I_Q = 0.5 \text{ mA}$. The transistor parameters are $\beta = 120$, $V_{BE(\text{on})} = 0.7$ V, and $V_A = \infty$. (a) Using Figure 11.3(a), determine the maximum common-mode input voltage v_{cm} that can be applied such that the transistors Q_1 and Q_2 remain biased in the active region. (b) Using Figure 11.3(b), determine the change in v_{C2} from its dc value if $v_d = 18 \text{ mV}$. (c) Repeat part (b) if $v_d = 10 \text{ mV}$.
- D11.7 The diff-amp configuration shown in Figure P11.7 is biased at ± 3 V. The maximum power dissipation in the entire circuit is to be no more than 1.2 mW when $v_1 = v_2 = 0$. The available transistors have parameters: $\beta = 120$, $V_{BE(\text{on})} = 0.7$ V, and $V_A = \infty$. Design the circuit to produce the maximum possible differential-mode voltage gain, but such that the common-mode input voltage can be within the range $-1 \leq v_{CM} \leq 1$ V and the transistors are still biased in the forward-active region. What is the value of A_d ? What are the current and resistor values?
- 11.8 Consider the circuit in Figure P11.8, with transistor parameters: $\beta = 100$, $V_{BE(\text{on})} = 0.7$ V, and $V_A = \infty$. (a) For $v_1 = v_2 = 0$, find I_{C1} , I_{C2} , I_E , V_{CE1} , and V_{CE2} . (b) Determine the maximum and minimum values of the common-mode input voltage. (c) Calculate A_d for a one-sided output at the collector of Q_2 .

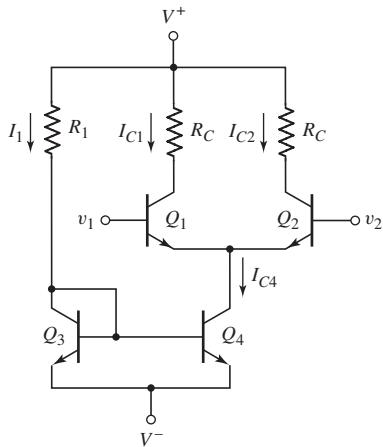


Figure P11.7

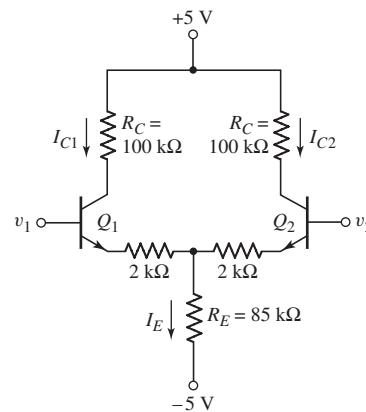


Figure P11.8

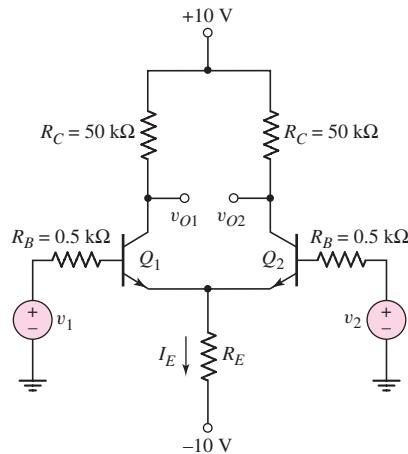


Figure P11.9

- *11.9 The transistor parameters for the circuit in Figure P11.9 are: $\beta = 100$, $V_{BE(on)} = 0.7$ V, and $V_A = \infty$. (a) Determine R_E such that $I_E = 150 \mu\text{A}$. (b) Find A_d , A_{cm} , and CMRR_{dB} for a one-sided output at v_{O2} . (c) Determine the differential- and common-mode input resistances.

- 11.10 The bias voltages for the diff-amp shown in Figure P11.10 are $V^+ = 3$ V and $V^- = -3$ V. The transistor current gains are $\beta = 80$, the nominal value of $V_{BE(on)}$ is 0.6 V, and $V_A = \infty$. (a) Design the circuit such that the quiescent collector currents are $50 \mu\text{A}$ and $v_{C1} = v_{C2} = -1.5$ V for $v_1 = v_2 = 0$. (b) Determine v_{C1} and v_{C2} when (i) $v_1 = v_2 = 1$ V and (ii) $v_1 = 0.994$ V, $v_2 = 1.006$ V.

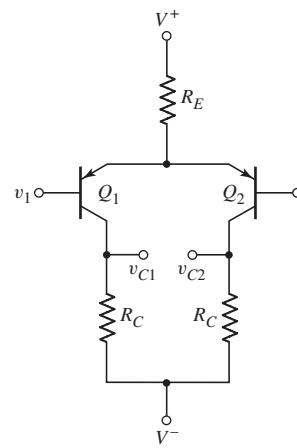


Figure P11.10

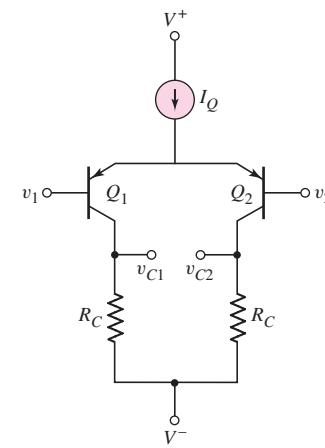


Figure P11.11

- 11.11 Consider the circuit shown in Figure P11.11. The circuit and transistor parameters are $V^+ = +3$ V, $V^- = -3$ V, $R_C = 360 \text{ k}\Omega$, $I_Q = 12 \mu\text{A}$, $\beta = 60$, $V_{EB(\text{on})} = 0.6$ V, and $V_A = \infty$. The output resistance of the current source is $R_o = 4 \text{ M}\Omega$. (a) Determine the Q-points of the transistors for $v_1 = v_2 = 0$. (b) Determine the differential- and common-mode voltage gains for (i) $v_O = v_{C1} - v_{C2}$ and (ii) $v_O = v_{C2}$.

- 11.12 The circuit and transistor parameters for the circuit shown in Figure P11.11 are $V^+ = 5 \text{ V}$, $V^- = -5 \text{ V}$, $I_Q = 0.2 \text{ mA}$, $\beta = 80$, $V_{BE(\text{on})} = 0.6 \text{ V}$, and $V_A = \infty$. (a) Design the circuit such that the minimum common-mode voltage is $v_{cm} = -2.5 \text{ V}$. (b) Using the results of part (a), what is the magnitude of the differential-mode gain, $A_d = |(v_{C1} - v_{C2}) / (v_1 - v_2)|$? (c) Determine v_{C1} and v_{C2} for $v_1 = 0.507 \text{ V}$ and $v_2 = 0.493 \text{ V}$. (d) What is the minimum output resistance of the current source such that $\text{CMRR}_{\text{dB}} = 60 \text{ dB}$ for a one-sided output.

- 11.13 Consider the differential amplifier shown in Figure P11.13 with mismatched collector resistors. The circuit and transistor parameters are $V^+ = 5 \text{ V}$, $V^- = -5 \text{ V}$, $\beta = 120$, $V_{BE(\text{on})} = 0.7 \text{ V}$, and $V_A = \infty$. (a) For $\Delta R = 0$, design the circuit such that $I_{CQ1} = I_{CQ2} = 120 \mu\text{A}$ and $v_{C1} = v_{C2} = 3 \text{ V}$ for $v_1 = v_2 = 0$. (b) Using the results of part (a), determine $|A_d|$ for a two-sided output. (c) For $\Delta R = 500 \Omega$, determine A_d , A_{cm} , and CMRR_{dB} for $v_o = \Delta(v_{C1} - v_{C2})$.

- 11.14 Consider the differential amplifier shown in Figure P11.14 with mismatched transistors. The mismatched transistors result in mismatched transconductances as shown. The circuit and transistor parameters are $V^+ = +10 \text{ V}$, $V^- = -10 \text{ V}$, $R_C = 50 \text{ k}\Omega$, $R_E = 75 \text{ k}\Omega$, $\beta = 120$, $V_{BE(\text{on})} = 0.7 \text{ V}$, and $V_A = \infty$. Determine A_d , A_{cm} , and CMRR_{dB} for $\Delta g_m/g_m = 0.01$ and for $v_o = v_{C1} - v_{C2}$. Assume $v_1 = v_2 = 0$ in the quiescent condition.

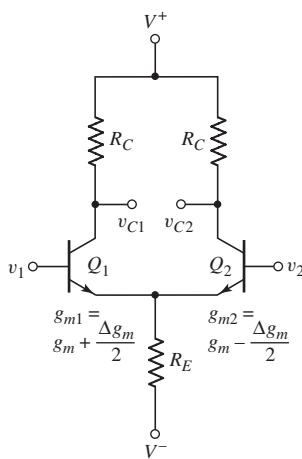


Figure P11.14

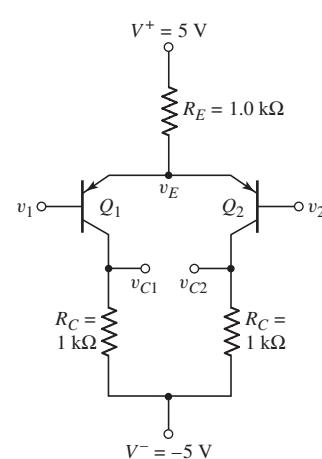


Figure P11.15

- 11.15 Consider the circuit in Figure P11.15. The transistor parameters are $\beta = 120$, $V_{BE(\text{on})} = 0.7 \text{ V}$, and $V_A = \infty$. Determine v_E , v_{C1} , and v_{C2} for (a) $v_1 = v_2 = 0$; (b) $v_1 = 0.5 \text{ V}$, $v_2 = 0$; and (c) $v_1 = 0$, $v_2 = 0.015 \text{ V}$.
- 11.16 (a) Design the circuit shown in Figure P11.16 such that $v_o = v_{C1} - v_{C2} = 1 \text{ V}$ when $v_1 = -5 \text{ mV}$ and $v_2 = +5 \text{ mV}$. The transistor parameters are $\beta = 180$, $V_{BE(\text{on})} = 0.7 \text{ V}$, and $V_A = \infty$. (b) Using the results of part (a), determine the maximum common-mode input voltage.
- 11.17 Consider the differential amplifier in Figure P11.17 with parameters $V^+ = 5 \text{ V}$, $V^- = -5 \text{ V}$, and $I_O = 0.8 \text{ mA}$. Neglect base currents and assume $V_A = \infty$ for all transistors. The emitter currents can be written as

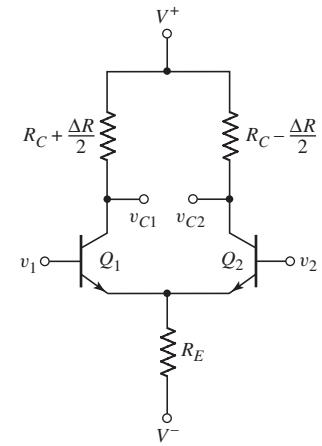


Figure P11.13

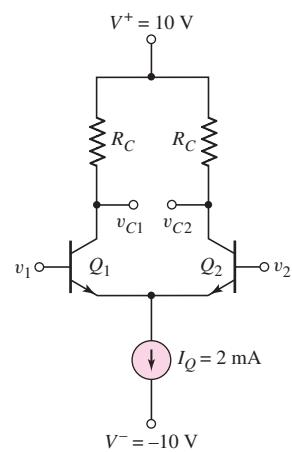


Figure P11.16

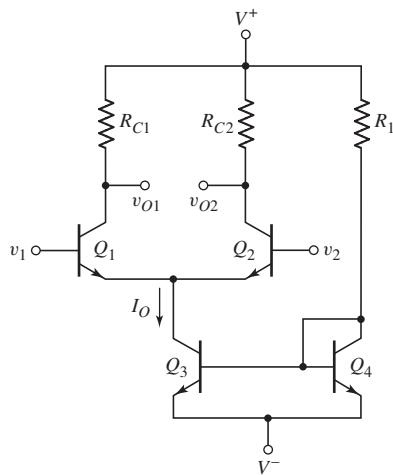


Figure P11.17

$I_{E1} = I_{S1}e^{V_{BE1}/V_T}$ and $I_{E2} = I_{S2}e^{V_{BE2}/V_T}$. (a) If $v_1 = v_2 = 0$ and $I_{S1} = I_{S2} = 3 \times 10^{-15}$ A, determine $(v_{O1} - v_{O2})$ for (i) $R_{C1} = R_{C2} = 7.5$ k Ω and (ii) $R_{C1} = 7.4$ k Ω , $R_{C2} = 7.6$ k Ω . (b) Repeat part (a) for $I_{S1} = 2.9 \times 10^{-15}$ A and $I_{S2} = 3.1 \times 10^{-15}$ A.

- 11.18 For the diff-amp in Figure 11.2, determine the value of $v_d = v_1 - v_2$ that produces (a) $i_{C1} = 0.20I_Q$ and (b) $i_{C2} = 0.90I_Q$.
- 11.19 Consider the expanded dc transfer curves shown in Figure 11.6. Determine the maximum differential input voltage such that the actual curve is within (a) 0.5 percent of the ideal linear extrapolation and (b) 1.5 percent of the ideal extrapolation.
- *D11.20 The diff-amp for the experimental system described in Example 11.4 needs to be redesigned. The range of the output voltage has increased to $-2 \leq V_O \leq 2$ V while the differential-mode voltage gain is still $A_d = 100$. The common-mode input voltage has increased to $v_{CM} = 3.5$ V. The value of CMRR needs to be increased to 80 dB.
- *11.21 The transistor parameters for the circuit in Figure P11.9 are: $\beta = 120$, $V_{BE(\text{on})} = 0.7$ V, and $V_A = \infty$. (a) Determine R_E such that $I_E = 0.25$ mA. (b) Assume the R_B resistance connected to the base of Q_2 is zero while the R_B resistance connected to the base of Q_1 remains at 0.5 k Ω . (i) Determine the differential-mode voltage gain for a one-sided output at v_{O2} . (ii) Determine the common-mode voltage gain for a one-sided output at v_{O2} .
- 11.22 The circuit parameters of the diff-amp shown in Figure 11.2 are $V^+ = 3$ V, $V^- = -3$ V, and $I_Q = 0.25$ mA. Base currents are negligible and $V_A = \infty$ for each transistor. (a) Design the circuit such that a differential-mode output voltage of $v_o = v_{C1} - v_{C2} = 1.2$ V is produced when a differential-mode input voltage of $v_d = v_1 - v_2 = 16$ mV is applied. (b) What is the maximum possible common-mode input voltage that can be applied such that the input transistors remain biased in the forward-active mode? (c) For a one-sided output, what is the value of CMRR_{dB} if the output resistance of the current source is $R_o = 4$ M Ω ?
- *11.23 Consider the circuit in Figure P11.23. Assume the Early voltage of Q_1 and Q_2 is $V_A = \infty$, and assume the current source I_Q is ideal. Derive the

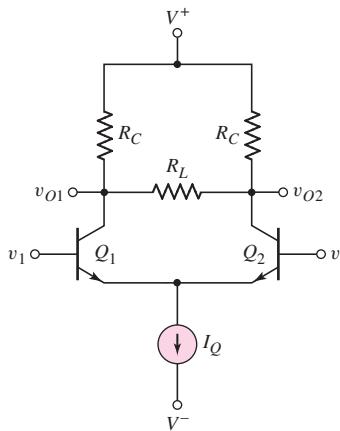


Figure P11.23

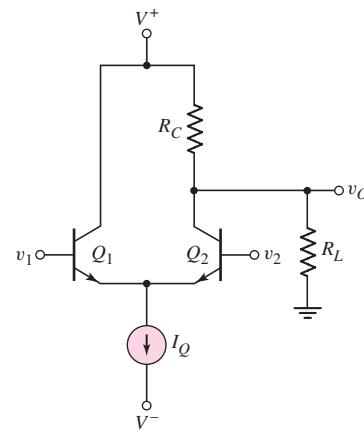


Figure P11.24

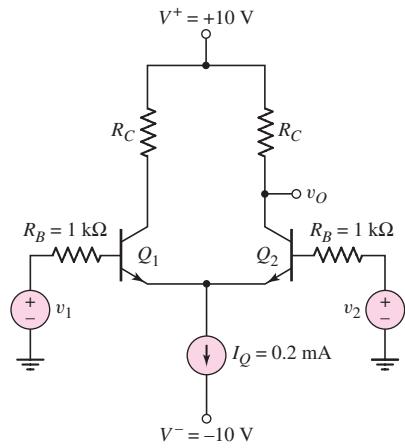


Figure P11.26

expressions for the one-sided differential-mode gain $A_{v1} = v_{o1}/v_d$ and $A_{v2} = v_{o2}/v_d$, and for the two-sided differential-mode gain $A_d = (v_{o2} - v_{o1})/v_d$.

- 11.24 The Early voltage of transistors Q_1 and Q_2 in the circuit in Figure P11.24 is $V_A = \infty$. Assuming an ideal current source I_Q , derive the expression for the differential-mode gain $A_d = v_o/v_d$.
- *11.25 Consider the small-signal equivalent circuit of the differential-pair configuration shown in Figure 11.9. Derive the expressions for the differential- and common-mode voltage gains if the output is a two-sided output defined as $V_o = V_{c2} - V_{c1}$.
- *D11.26 Consider a BJT diff-amp with the configuration in Figure P11.26. The signal sources have nonzero source resistances as shown. The transistor parameters are: $\beta = 150$, $V_{BE}(\text{on}) = 0.7$ V, and $V_A = \infty$. The range of the common-mode input voltage is to be $-3 \leq v_{CM} \leq 3$ V and the CMRR is to be 75 dB. (a) Design the diff-amp to produce the maximum possible differential-mode voltage gain. (b) Design the current source to produce the desired bias current and CMRR.
- 11.27 The bridge circuit in Figure P11.27 is a temperature transducer in which the resistor R_A is a thermistor (a resistor whose resistance varies with temperature). The value of δ varies over the range of $-0.01 \leq \delta \leq 0.01$ as temperature varies over a particular range. Assume the value of $R = 40$ k Ω . The bridge circuit is to be connected to the diff-amp in Figure 11.2. The diff-amp circuit parameters are $V^+ = 5$ V, $V^- = -5$ V, $I_Q = 0.2$ mA, and $R_C = 15$ k Ω . The transistor parameters are $\beta = 120$, $V_{BE}(\text{on}) = 0.7$ V, and $V_A = \infty$. Terminal A of the bridge circuit is connected to the base of Q_1 and terminal B is connected to the base of Q_2 . Determine the range of output voltage v_{O2} as δ changes. [Hint: Make a Thevenin equivalent circuit at terminals A and B of the bridge circuit.]
- 11.28 A diff-amp is biased with a constant-current source $I_Q = 0.25$ mA that has an output resistance of $R_o = 8$ M Ω . The bipolar transistor parameters are $\beta = 120$ and $V_A = \infty$. Determine (a) the differential-mode input resistance and (b) the common-mode input resistance.

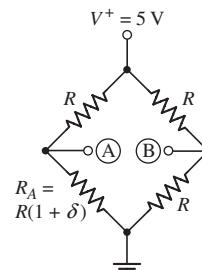


Figure P11.27

- 11.29 The transistor parameters for the circuit shown in Figure P11.29 are $\beta = 180$, $V_{BE(on)} = 0.7$ V (except for Q_4), $V_A = \infty$ for Q_1 and Q_2 , and $V_A = 100$ V for Q_3 and Q_4 . (a) Determine R_1 and R_2 such that $I_1 = 0.5$ mA and $I_Q = 140 \mu\text{A}$. (b) Determine the common-mode input resistance. (c) For $R_C = 40 \text{ k}\Omega$, determine the common-mode voltage gain.

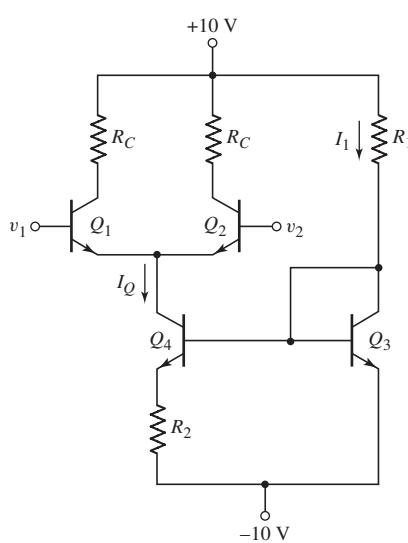


Figure P11.29

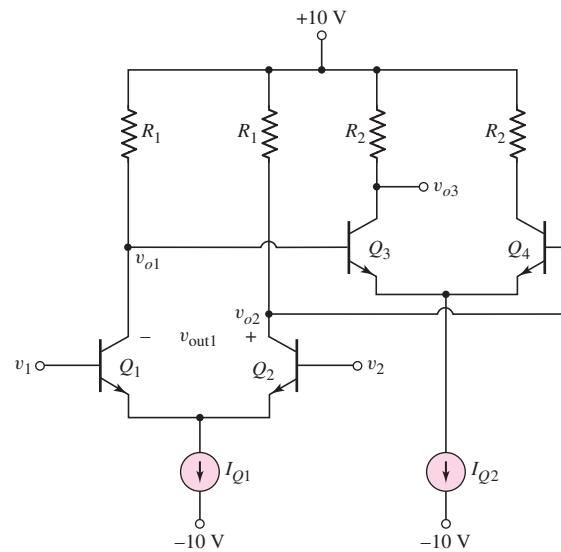


Figure P11.30

- D11.30 Figure P11.30 shows a two-stage cascade diff-amp with resistive loads. Power supply voltages of ± 10 V are available. Assume transistor parameters of: $\beta = 100$, $V_{BE(on)} = 0.7$ V, and $V_A = \infty$. Design the circuit such that the two-sided differential-mode voltage gain is $A_{d1} = (v_{o2} - v_{o1})/(v_1 - v_2) = 20$ for the first stage, and that the one-sided differential-mode voltage gain is $A_{d2} = v_{o3}/(v_{o2} - v_{o1}) = 30$ for the second stage. The circuit is to be designed such that the maximum differential-mode voltage swing is obtained in each stage.

Section 11.3 Basic FET Differential Pair

- 11.31 For the differential amplifier in Figure P11.31 the parameters are $R_1 = 50 \text{ k}\Omega$ and $R_D = 24 \text{ k}\Omega$. The transistor parameters are: $K_n = 0.25 \text{ mA/V}^2$, $\lambda = 0$, and $V_{TN} = 2$ V. (a) Determine I_1 , I_Q , I_{D1} , V_{DS1} , and V_{DS4} when $v_1 = v_2 = 0$. (b) Draw the dc load line and plot the Q-point for transistor M_2 . (c) What are the maximum and minimum values of the common-mode input voltage?

- 11.32 The bias voltages in the diff-amp shown in Figure P11.31 are changed to $V^+ = 3$ V and $V^- = -3$ V. The transistor parameters are $K_{n1} = K_{n2} = 100 \mu\text{A/V}^2$, $K_{n3} = K_{n4} = 200 \mu\text{A/V}^2$, $\lambda_1 = \lambda_2 = 0$, $\lambda_3 = \lambda_4 = 0.01 \text{ V}^{-1}$ and $V_{TN} = 0.3$ V (all transistors). (a) Design the circuit such that $V_{DS1} = V_{DS2} = 4$ V and $I_{D1} = I_{D2} = 60 \mu\text{A}$ when $v_1 = v_2 = -1.15$ V. (i) What are the values of I_Q and I_1 ? (ii) What are the values of R_D and R_1 ? (iii) What are the values of V_{GS1} and V_{GS4} ? (b) Calculate the change in I_Q if $v_1 = v_2 = +1.15$ V.

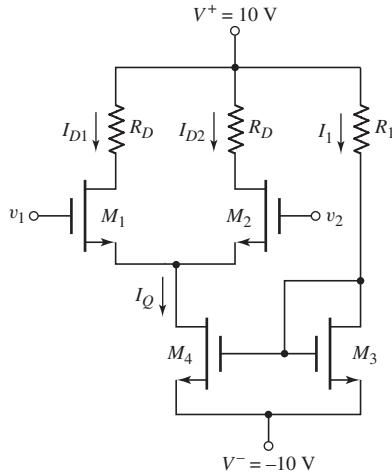


Figure P11.31

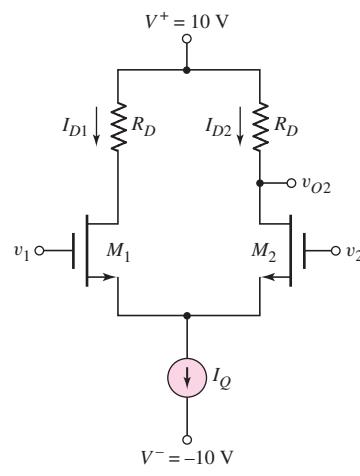


Figure P11.33

- 11.33 The transistor parameters for the differential amplifier shown in Figure P11.33 are $V_{TN} = 0.5 \text{ V}$, $k'_n = 80 \mu\text{A/V}^2$, $W/L = 4$, and $\lambda = 0$.
(a) Find R_D and I_Q such that $I_{D1} = I_{D2} = 80 \mu\text{A}$ and $v_{O2} = 2 \text{ V}$ when $v_1 = v_2 = 0$. (b) Draw the dc load line, and plot the Q -point for M_2 .
(c) What is the maximum common-mode input voltage?
11.34 The diff-amp in Figure P11.34 has parameters $V^+ = 3 \text{ V}$, $V^- = -3 \text{ V}$, and $I_Q = 0.18 \text{ mA}$. The transistor parameters are $V_{TN} = 0.35 \text{ V}$, $k'_n = 100 \mu\text{A/V}^2$, $W/L = 4$, and $\lambda = 0$. (a) Using Figure P11.34(a), determine R_D such that the maximum value of the common-mode input voltage is $v_{cm}(\max) = 2.25 \text{ V}$. The input transistors M_1 and M_2 must remain biased in the saturation region. (b) Using Figure P11.34(b), determine the value of v_{D2} for (i) $v_d = 0$, (ii) $v_d = +120 \text{ mV}$, and (iii) $v_d = -50 \text{ mV}$.
11.35 The bias voltages of the diff-amp shown in Figure P11.35 are $V^+ = 5 \text{ V}$ and $V^- = -5 \text{ V}$. The threshold voltage of each transistor is $V_{TN} = 0.4 \text{ V}$ and

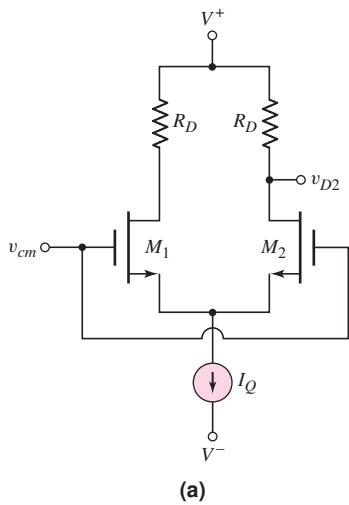
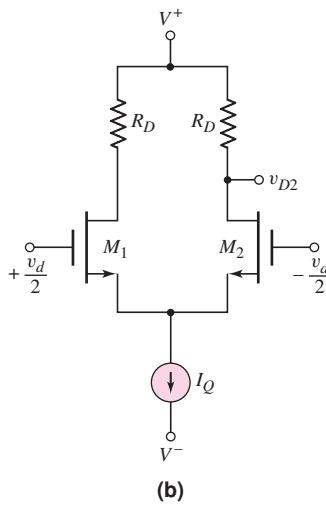


Figure P11.34



(b)

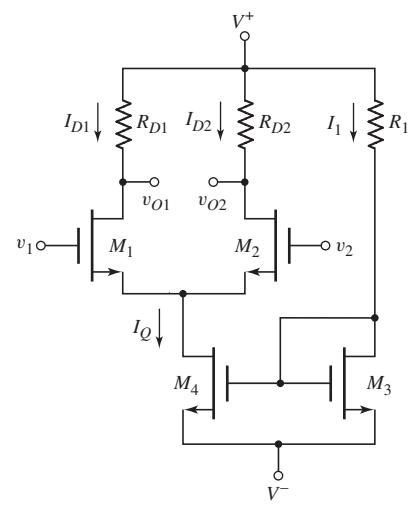


Figure P11.35

assume $\lambda = 0$. Let $K_{n3} = K_{n4} = 0.20 \text{ mA/V}^2$. The drain currents can be written as $I_{D1} = K_{n1}(V_{GS1} - V_{TN})^2$ and $I_{D2} = K_{n2}(V_{GS2} - V_{TN})^2$.

(a) Design the circuit such that $I_Q = 0.25 \text{ mA}$ when $v_1 = v_2 = 0$. (b) If $v_1 = v_2 = 0$ and $K_{n1} = K_{n2} = 0.120 \text{ mA/V}^2$, find $v_{O1} - v_{O2}$ when (i) $R_{D1} = R_{D2} = 15 \text{ k}\Omega$ and (ii) $R_{D1} = 14.5 \text{ k}\Omega$, $R_{D2} = 15.5 \text{ k}\Omega$. (c) Repeat part (b) for $K_{n1} = 0.125 \text{ mA/V}^2$ and $K_{n2} = 0.115 \text{ mA/V}^2$.

- 11.36 The circuit parameters of the diff-amp shown in Figure 11.19 are $V^+ = 3 \text{ V}$, $V^- = -3 \text{ V}$, and $I_Q = 0.15 \text{ mA}$. The transistor parameters are $K_n = 0.2 \text{ mA/V}^2$, $V_{TN} = 0.4 \text{ V}$, and $\lambda = 0$. Determine the value of $v_d = v_{G1} - v_{G2}$ that produces (a) $i_{D1} = 0.2I_Q$, (b) $i_{D2} = 0.8I_Q$, and (c) $i_{D1} = I_Q$.
- 11.37 Consider the normalized dc transfer characteristics of a MOSFET diff-amp shown in Figure 11.21. Assume that $K_n = 0.20 \text{ mA/V}^2$ and $I_Q = 0.10 \text{ mA}$. Determine the maximum differential input voltage such that the actual curve of i_{D1}/I_Q is within (a) 0.5 percent of the ideal linear extrapolation and (b) 1.5 percent of the ideal linear extrapolation.
- 11.38 The parameters of the diff-amp circuit shown in Figure P11.38 are $V^+ = 9 \text{ V}$, $V^- = -9 \text{ V}$, $R_D = 510 \text{ k}\Omega$, and $R_S = 390 \text{ k}\Omega$. The transistor parameters are $V_{TP} = -0.8 \text{ V}$, $K_p = 50 \mu\text{A/V}^2$, and $\lambda = 0$. Determine v_{D1} and v_{D2} for (a) $v_1 = v_2 = 1 \text{ V}$ and (b) $v_1 = 1.050 \text{ V}$, $v_2 = 0.950 \text{ V}$.

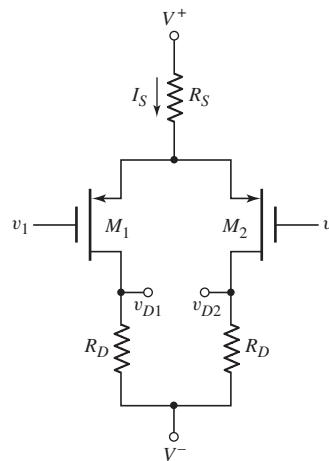


Figure P11.38

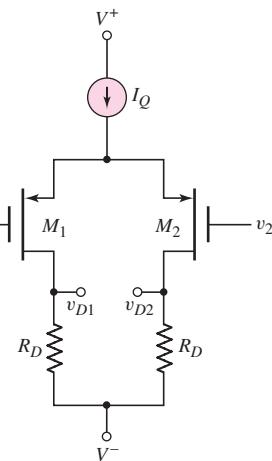


Figure P11.39

- 11.39 Consider the circuit shown in Figure P11.39. The circuit and transistor parameters are $V^+ = +3 \text{ V}$, $V^- = -3 \text{ V}$, $R_D = 360 \text{ k}\Omega$, $I_Q = 12 \mu\text{A}$, $V_{TP} = -0.4 \text{ V}$, $K_p = 30 \mu\text{A/V}^2$, and $\lambda = 0$. The output resistance of the current source is $R_o = 4 \text{ M}\Omega$. (a) Determine the Q -points of the transistors for $v_1 = v_2 = 0$. (b) Determine the differential- and common-mode voltage gains for (i) $v_O = v_{D1} - v_{D2}$ and (ii) $v_O = v_{D2}$.
- 11.40 The circuit and transistor parameters for the circuit shown in Figure P11.39 are $V^+ = 5 \text{ V}$, $V^- = -5 \text{ V}$, $I_Q = 0.15 \text{ mA}$, $R_D = 30 \text{ k}\Omega$, $V_{TP} = -0.5 \text{ V}$, $K_p = 0.12 \text{ mA/V}^2$, and $\lambda = 0$. The output resistance of the current source is $R_o = 4 \text{ M}\Omega$. (a) Determine $v_{D1} - v_{D2}$ for (i) $v_1 = +0.05 \text{ V}$, $v_2 = -0.05 \text{ V}$ and (ii) $v_1 = +0.10 \text{ V}$, $v_2 = -0.10 \text{ V}$. (b) Determine the change in v_{D2} as the inputs change (i) from $v_1 = v_2 = 0$ to $v_1 = 0.10 \text{ V}$ and $v_2 = -0.10 \text{ V}$, and (ii) from $v_1 = v_2 = 0$ to $v_1 = 1.10 \text{ V}$ and $v_2 = 0.90 \text{ V}$.

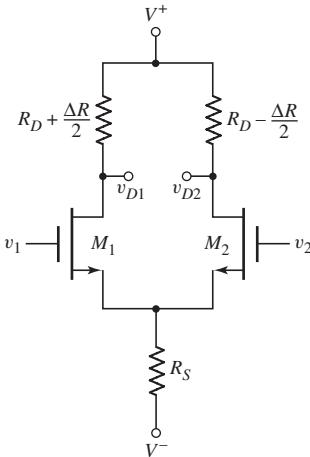


Figure P11.41

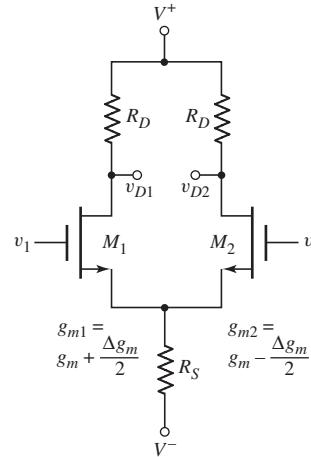


Figure P11.42

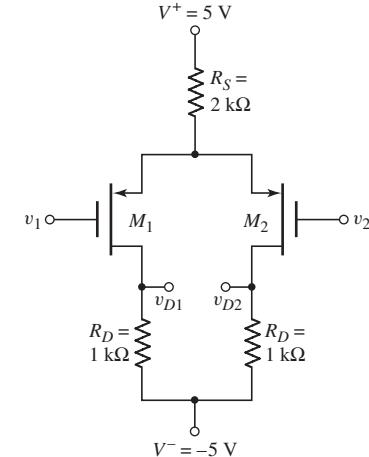


Figure P11.43

- 11.41 Consider the differential amplifier shown in Figure P11.41 with mismatched drain resistors. The circuit and transistor parameters are $V^+ = +10$ V, $V^- = -10$ V, $R_D = 50$ k Ω , $R_S = 75$ k Ω , $V_{TN} = 1$ V, $K_n = 0.15$ mA/V 2 , and $\lambda = 0$. Determine A_d , A_{cm} , and $CMRR|_{dB}$ for $\Delta R = 500$ Ω and for $v_O = v_{D1} - v_{D2}$. Assume $v_1 = v_2 = 0$ in the quiescent condition.
- 11.42 Consider the differential amplifier shown in Figure P11.42 with mismatched transistors. The mismatched transistors result in mismatched transconductances as shown. The circuit and transistor parameters are $V^+ = +10$ V, $V^- = -10$ V, $R_D = 50$ k Ω , $R_S = 75$ k Ω , $V_{TN} = 1$ V, $K_n = 0.15$ mA/V 2 , and $\lambda = 0$. Determine A_d , A_{cm} , and $CMRR|_{dB}$ for $\Delta g_m/g_m = 0.01$ and for $v_O = v_{D1} - v_{D2}$. Assume $v_1 = v_2 = 0$ in the quiescent condition.
- 11.43 Consider the circuit in Figure P11.43. The transistor parameters are $K_p = 1.2$ mA/V 2 , $V_{TP} = -0.6$ V, and $\lambda = 0$. Determine v_S , v_{D1} , and v_{D2} for (a) $v_1 = v_2 = 0$; (b) $v_1 = v_2 = 1$ V; (c) $v_1 = -0.1$ V, $v_2 = 0.1$ V; and (d) $v_1 = 0.9$ V, $v_2 = 1.1$ V.
- D11.44 (a) Design the circuit shown in Figure P11.44 such that $v_O = v_{D1} - v_{D2} = 1$ V when $v_1 = -50$ mV and $v_2 = +50$ mV. The transistor parameters are $V_{TN} = 0.8$ V, $K_n = 0.4$ mA/V 2 , and $\lambda = 0$. (b) Using the results of part (a), determine the maximum common-mode input voltage.
- *D11.45 The Hall effect experimental arrangement was described in Example 11.4. The required diff-amp is to be designed in the circuit configuration in Figure P11.35. The transistor parameters are $V_{TN} = 0.8$ V, $k'_n = 80$ μ A/V 2 , $\lambda_1 = \lambda_2 = 0$, and $\lambda_3 = \lambda_4 = 0.01$ V $^{-1}$. If the CMRR requirement cannot be met, a more sophisticated current source may have to be designed.
- *11.46 Consider the diff-amp in Figure P11.46. The transistor parameters are: $K_{n1} = K_{n2} = 50$ μ A/V 2 , $\lambda_1 = \lambda_2 = 0.02$ V $^{-1}$, and $V_{TN1} = V_{TN2} = 1$ V. (a) Determine I_S , I_{D1} , I_{D2} , and v_{O2} for $v_1 = v_2 = 0$. (b) Using the small-signal equivalent circuit, determine the differential-mode voltage gain $A_d = v_{o2}/v_d$, the common-mode voltage gain $A_{cm} = v_{o2}/v_{cm}$, and the $CMRR|_{dB}$.

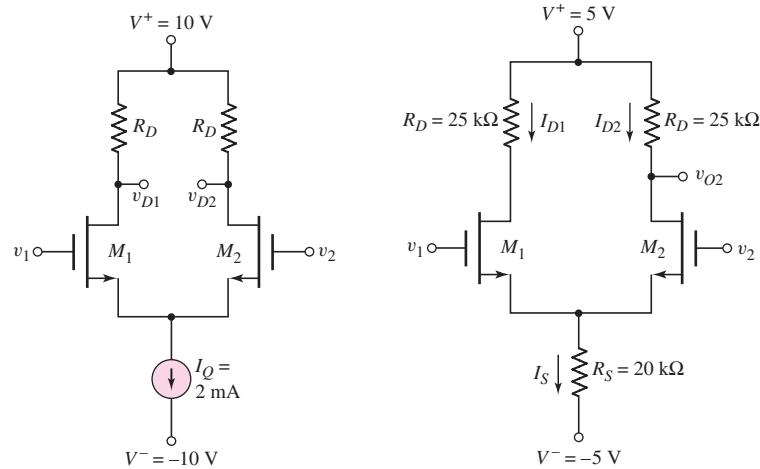


Figure P11.44

Figure P11.46

- 11.47 Consider the circuit shown in Figure P11.47. Assume that $\lambda = 0$ for M_1 and M_2 . Also assume an ideal current source I_Q . Derive the expression for the one-sided differential mode gains $A_{d1} = v_{o1}/v_d$ and $A_{d2} = v_{o2}/v_d$, and the two-sided differential-mode gain $A_d = (v_{o2} - v_{o1})/v_d$.

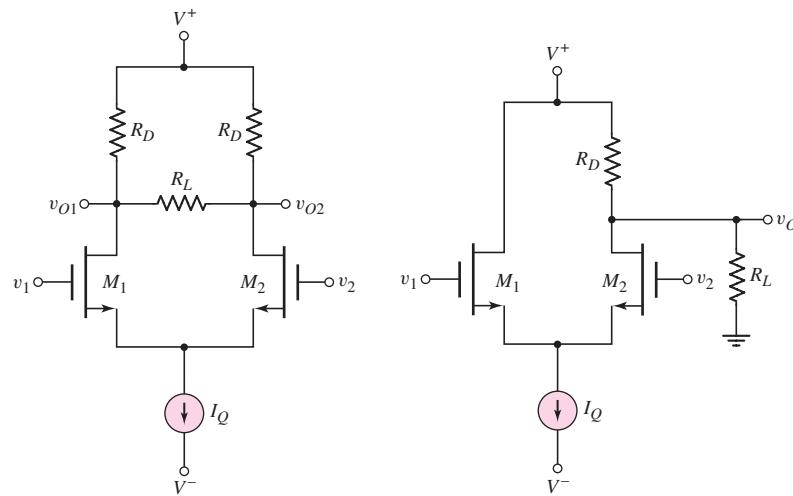


Figure P11.47

Figure P11.48

- 11.48 Consider the diff-amp shown in Figure P11.48. Assume $\lambda_1 = \lambda_2 = 0$ and assume the current source has an output resistance of R_o . (a) Derive the expression for the differential-mode voltage gain $A_d = \Delta v_O/v_d$ where $v_d = v_1 - v_2$. (b) Derive the expression for the common-mode voltage gain $A_{cm} = \Delta v_O/v_{cm}$ where $v_{cm} = (v_1 + v_2)/2$.
- 11.49 The bias voltages of the diff-amp circuit shown in Figure 11.19 are $V^+ = 5\text{ V}$ and $V^- = -5\text{ V}$, and the bias current is $I_Q = 0.2\text{ mA}$. The transistor parameters are $V_{TN} = 0.4\text{ V}$, $K_n = 0.15\text{ mA/V}^2$, and $\lambda = 0$. (a) Design the circuit such that a differential-mode output voltage of $\Delta v_O = 0.5\text{ V}$ is

produced when a differential-mode input voltage of $v_d = v_1 - v_2 = 100 \text{ mV}$ is applied. (b) Using the results of part (a), determine the maximum possible common-mode input voltage that can be applied such that the transistors remain biased in the saturation region.

- 11.50 Consider the small-signal equivalent circuit in Figure 11.23. Assume the output is a two-sided output defined as $V_o = V_{d2} - V_{d1}$, where V_{d2} and V_{d1} are the signal voltages at the drains of M_2 and M_1 , respectively. Derive expressions for the differential- and common-mode voltage gains.
- 11.51 Consider the MOSFET diff-amp with the configuration in Figure P11.33. The circuit parameters are $V^+ = 3 \text{ V}$, $V^- = -3 \text{ V}$, and $I_Q = 0.2 \text{ mA}$. The transistor parameters are $V_{TN} = 0.4 \text{ V}$, $k'_n = 100 \mu\text{A/V}^2$, $W/L = 10$, and $\lambda = 0$. The range of the common-mode input voltage is to be $-1.5 \leq v_{cm} \leq +1.5 \text{ V}$, and the common-mode rejection ratio is to be $\text{CMRR}_{\text{dB}} = 50 \text{ dB}$. (a) Design the diff-amp to produce the maximum possible differential-mode voltage gain. (b) Design an all MOSFET current source to produce the desired bias current and CMRR. The minimum W/L ratio of any transistor is to be 0.8, and assume $\lambda = 0.02 \text{ V}^{-1}$ for all transistors in the current source circuit.
- 11.52 Consider the bridge circuit and diff-amp described in Problem 11.27. The BJT diff-amp is to be replaced with a MOSFET diff-amp as shown in Figure 11.19. The transistor parameters are $V_{TN} = 0.4 \text{ V}$, $K_n = 1 \text{ mA/V}^2$, and $\lambda = 0$. The bias voltages of the MOSFET diff-amp are $V^+ = 5 \text{ V}$ and $V^- = -5 \text{ V}$, and the reference current is $I_Q = 0.2 \text{ mA}$. Let $R_D = 20 \text{ k}\Omega$. Terminal A of the bridge circuit is to be connected to the gate of M_1 and terminal B is to be connected to the gate of M_2 . (a) Determine the range of output voltage v_O as δ changes. (b) Explain the advantages and disadvantages of this circuit configuration compared to that in Problem 11.27.
- *D11.53 Figure P11.53 shows a two-stage cascade diff-amp with resistive loads. Power supply voltages of $\pm 10 \text{ V}$ are available. Assume transistor parameters of $V_{TN} = 1 \text{ V}$, $k'_n = 60 \mu\text{A/V}^2$, and $\lambda = 0$. Design the circuit such

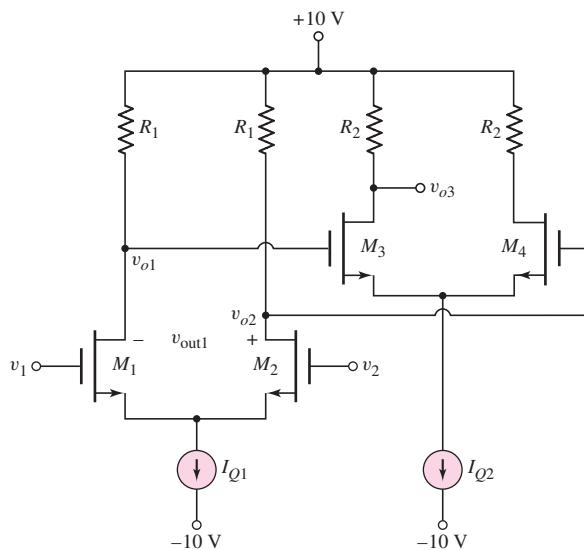


Figure P11.53

that the two-sided differential-mode voltage gain is $A_{d1} = (v_{o2} - v_{o1})/(v_1 - v_2) = 20$ for the first stage, and that the one-sided differential-mode voltage gain is $A_{d2} = v_{o3}/(v_{o2} - v_{o1}) = 30$ for the second stage. The circuit is to be designed such that the maximum differential-mode voltage swing is obtained in each stage.

- *11.54 Figure P11.54 shows a matched JFET differential pair biased with a current source I_Q . (a) Starting with

$$i_D = I_{DSS} \left(1 - \frac{v_{GS}}{V_P}\right)^2$$

show that

$$\frac{i_{D1}}{I_Q} = \frac{1}{2} + \left(\frac{1}{-2V_P}\right) v_d \sqrt{2 \left(\frac{I_{DSS}}{I_Q}\right) - \left(\frac{I_{DSS}}{I_Q}\right)^2 \left(\frac{v_d}{V_P}\right)^2}$$

and

$$\frac{i_{D2}}{I_Q} = \frac{1}{2} - \left(\frac{1}{-2V_P}\right) v_d \sqrt{2 \left(\frac{I_{DSS}}{I_Q}\right) - \left(\frac{I_{DSS}}{I_Q}\right)^2 \left(\frac{v_d}{V_P}\right)^2}$$

- (b) Show that the I_Q bias current is switched entirely to one transistor or the other when

$$|v_d| = |V_P| \sqrt{\frac{I_Q}{I_{DSS}}}$$

- (c) Show that the maximum forward transconductance is given by

$$g_f(\max) = \frac{di_{D1}}{dv_d} \Big|_{v_d=0} = \left(\frac{1}{-V_P}\right) \sqrt{\frac{I_Q \cdot I_{DSS}}{2}}$$

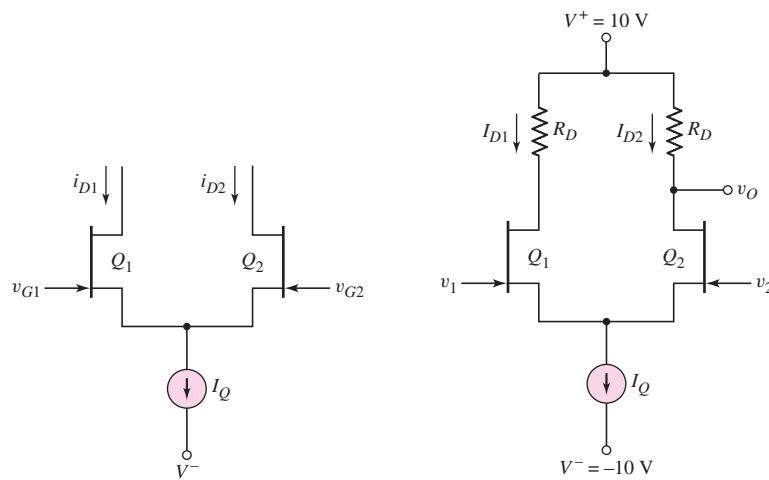


Figure P11.54

Figure P11.55

- 11.55 A JFET differential amplifier is shown in Figure P11.55. The transistor parameters are: $V_P = -4$ V, $I_{DSS} = 2$ mA, and $\lambda = 0$. (a) Find R_D and I_Q such that $I_{D1} = I_{D2} = 0.5$ mA and $v_{o2} = 7$ V when $v_1 = v_2 = 0$.

- (b) Calculate the maximum forward transconductance. (c) Determine the one-sided differential-mode voltage gain $A_d = v_o/v_d$.
- *11.56 Consider the JFET diff-amp shown in Figure P11.56. The transistor parameters are: $I_{DSS} = 0.8 \text{ mA}$, $\lambda = 0$, and $V_P = -2 \text{ V}$. (a) Determine I_S , I_{D1} , I_{D2} , and v_{o2} for $v_1 = v_2 = 0$. (b) Using the small-signal equivalent circuit, determine the differential-mode voltage gain $A_d = v_{o2}/v_d$, the common-mode voltage gain $A_{cm} = v_o/v_{cm}$, and the CMRR_{dB}.

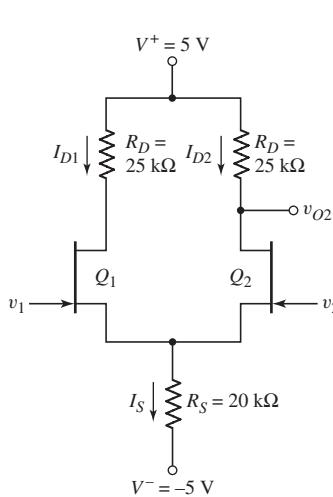


Figure P11.56

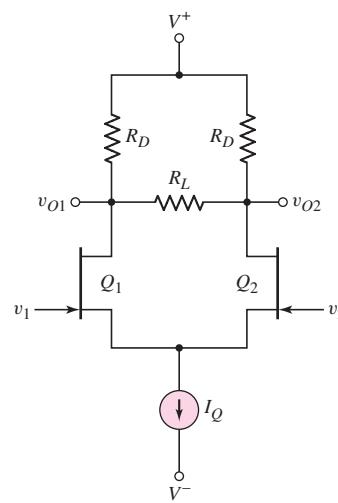


Figure P11.57

- *11.57 Consider the circuit in Figure P11.57. Assume that $\lambda = 0$ for the transistors, and assume an ideal current source I_Q . Derive the expressions for the one-sided differential-mode gains $A_{d1} = v_{o1}/v_d$ and $A_{d2} = v_{o2}/v_d$, and for the two-sided differential-mode gain $A_d = (v_{o2} - v_{o1})/v_d$.

Section 11.4 Differential Amplifier with Active Load

- 11.58 The circuit parameters for the diff-amp shown in Figure 11.30 are $V^+ = 3.3 \text{ V}$, $V^- = -3.3 \text{ V}$, and $I_Q = 0.4 \text{ mA}$. The transistor parameters are $\beta = 120$, $V_{A1} = V_{A2} = 120 \text{ V}$, $V_{A3} = V_{A4} = 80 \text{ V}$, and $V_{A5} = \infty$. (a) Determine the open-circuit differential-mode voltage gain. (b) What is the output resistance of the diff-amp? (c) Find the value of load resistance R_L that reduces the differential-mode gain to 75 percent of the open-circuit value.
- D11.59 Design a differential amplifier with the configuration shown in Figure 11.28 incorporating a basic two-transistor current source to establish I_Q . The bias voltages are to be $V^+ = +5 \text{ V}$ and $V^- = -5 \text{ V}$, the bias current is to be $I_Q = 250 \mu\text{A}$, and the available transistors have parameters $\beta = 180$, $V_{BE(\text{on})} = V_{EB(\text{on})} = 0.7 \text{ V}$, $V_{AN} = 150 \text{ V}$, and $V_{AP} = 100 \text{ V}$. (a) Show the complete circuit. (b) What is the open-circuit differential-mode voltage gain. (c) Determine the differential-mode input resistance and the output resistance. (d) Determine the common-mode input voltage range.
- 11.60 The differential amplifier shown in Figure P11.60 has a pair of pnp bipolars as input devices and a pair of npn bipolars connected as an active load.

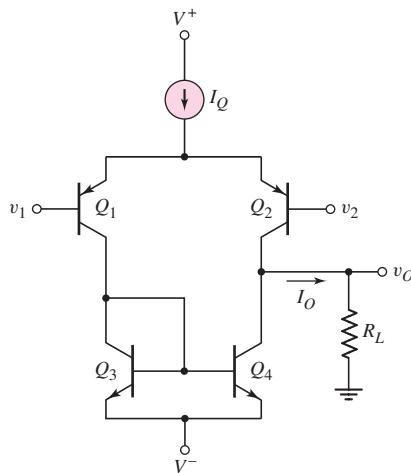


Figure P11.60

The circuit bias is $I_Q = 0.2$ mA, and the transistor parameters are $\beta = 100$ and $V_A = 100$ V. (a) Determine I_Q such that the dc currents in the diff-amp are balanced. (b) Find the open-circuit differential-mode voltage gain. (c) Determine the differential-mode voltage gain if a load resistance $R_L = 250$ k Ω is connected to the output.

- 11.61 The bias voltages for the diff-amp shown in Figure 11.30 are $V^+ = 5$ V and $V^- = -5$ V. A load resistance of $R_L = 250$ k Ω is capacitively coupled to the output. The transistor parameters are $\beta = 120$, $V_{A1} = V_{A2} = 90$ V, and $V_{A3} = V_{A4} = 60$ V. (a) Determine the bias current I_Q that will produce a differential-mode voltage gain of $A_d = 1000$. (b) If $V_{EB(on)} = 0.6$ V, what is the maximum common-mode voltage that can be applied such that all transistors are biased in the forward-active mode?
- 11.62 Consider the diff-amp shown in Figure P11.62. The circuit parameters are $V^+ = 3$ V, $V^- = -3$ V, and $I_Q = 0.4$ mA. The npn transistor parameters are $\beta_{npn} = 180$, $V_{BE(on)} = 0.7$ V, and $V_{AN} = 120$ V, and the pnp transistor parameters are $\beta_{pnp} = 120$, $V_{EB(on)} = 0.7$ V, and $V_{AP} = 80$ V. (a) Sketch the small-signal equivalent circuit for the diff-amp assuming an ideal differential-mode input signal. (b) Determine the one-sided differential-mode gain $A_{d1} = \Delta v_{O1}/v_d$. (c) Determine the one-sided differential-mode gain $A_{d2} = \Delta v_{O2}/v_d$. (d) Find the two-sided differential-mode gain $A_{d3} = \Delta(v_{O2} - v_{O1})/v_d$.
- 11.63 Consider the MOSFET diff-amp shown in Figure P11.63. The bias voltages are $V^+ = 3$ V and $V^- = -3$ V. The current source is $I_Q = 200$ μ A and has an output resistance of $R_o = 2$ M Ω . The transistor parameters are $V_{TN} = 0.4$ V, $V_{TP} = -0.4$ V, $K_n = K_p = 0.5$ mA/V 2 , $\lambda_2 = 0.02$ V, V $^{-1}$, $\lambda_4 = 0.03$ V $^{-1}$ and $\lambda_1 = \lambda_3 = 0$. (a) Determine the voltage gain $A = v_o/v_d$ for $v_1 = v_d$ and $v_2 = 0$. (b) Determine the voltage gain $A = v_o/v_d$ for $v_1 = 0$ and $v_2 = -v_d$. (c) Determine the voltage gain $A = v_o/v_d$ for $v_1 = v_d/2$ and $v_2 = -v_d/2$.
- 11.64 The differential amplifier in Figure P11.64 has a pair of PMOS transistors as input devices and a pair of NMOS transistors connected as an active load. The circuit is biased with $I_Q = 0.2$ mA, and the transistor parameters are: $K_n = K_p = 0.1$ mA/V 2 , $\lambda_n = 0.01$ V $^{-1}$, $\lambda_p = 0.015$ V $^{-1}$,

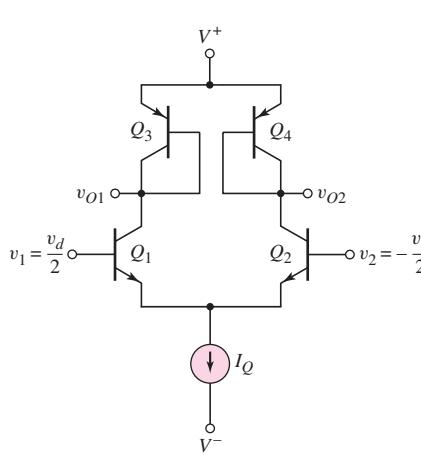


Figure P11.62

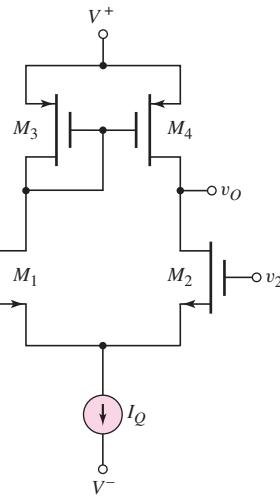


Figure P11.63

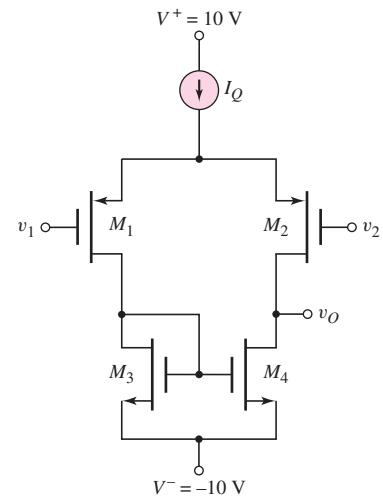


Figure P11.64

$V_{TN} = 1$ V, and $V_{TP} = -1$ V. (a) Determine the quiescent drain-to-source voltage in each transistor. (b) Find the open-circuit differential-mode voltage gain. (c) What is the output resistance?

- 11.65 The circuit parameters for the diff-amp shown in Figure 11.32 are $V^+ = 1.8$ V, $V^- = -1.8$ V, and $I_Q = 120 \mu\text{A}$. The NMOS transistor parameters are $V_{TN} = 0.3$ V, $k'_n = 100 \mu\text{A/V}^2$, $(W/L)_n = 8$, and $\lambda_n = 0.025 \text{ V}^{-1}$. The parameters of the PMOS transistors are $V_{TP} = -0.3$ V, $k'_p = 40 \mu\text{A/V}^2$, $(W/L)_p = 10$, and $\lambda_p = 0.04 \text{ V}^{-1}$. (a) Determine the small-signal differential-mode voltage gain $A_d = v_o/v_d$. (b) What is the maximum common-mode voltage gain that can be applied such that all transistors are still biased in the saturation region?

- *11.66 Consider the diff-amp with active load in Figure P11.66. The Early voltages are $V_{AN} = 120$ V for Q_1 and Q_2 and $V_{AP} = 80$ V for Q_3 and Q_4 .

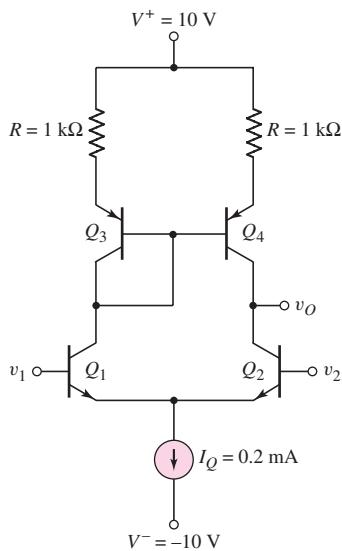


Figure P11.66

- (a) Determine the open-circuit differential-mode voltage gain. (b) Compare this value to the gain obtained when $R = 0$. (c) Determine the output resistance R_o for parts (a) and (b). Assume $\beta = 100$. [Hint: As a good approximation, use the output resistance results from a Widlar current source.]

11.67 The diff-amp in Figure P11.67 has a three-transistor active load circuit and a Darlington pair configuration connected to the output. Determine the bias current I_{Q1} in terms of I_Q such that the diff-amp dc currents are balanced.

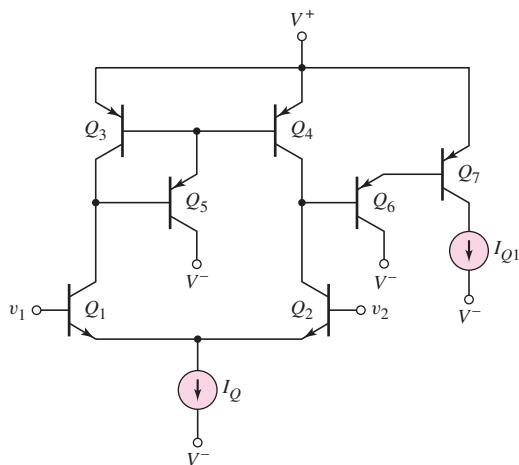


Figure P11.67

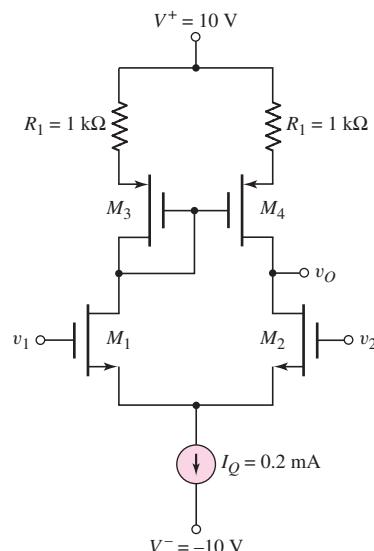


Figure P11.68

- *11.68 Consider the diff-amp in Figure P11.68. The PMOS parameters are: $K_p = 80 \mu\text{A}/\text{V}^2$, $\lambda_p = 0.02 \text{ V}^{-1}$, $V_{TP} = -2 \text{ V}$. The NMOS parameters are: $K_n = 80 \mu\text{A}/\text{V}^2$, $\lambda_n = 0.015 \text{ V}^{-1}$, $V_{TN} = +2 \text{ V}$. (a) Determine the open-circuit differential-mode voltage gain. (b) Compare this value to the gain obtained when $R_1 = 0$. (c) What is the output resistance of the diff-amp for parts (a) and (b)?

*11.69 Reconsider the circuit in Figure P11.60 except that $1 \text{ k}\Omega$ resistors are inserted at the emitters of the active load transistors Q_3 and Q_4 as in the circuit in Figure P11.66. Assume the same transistor parameters as in Problem 11.60. (a) Determine the output resistance looking into the output of the diff-amp circuit. (b) Find the open-circuit differential-mode voltage gain.

*11.70 Consider the circuit in Figure P11.70, in which the input transistors to the diff-amp are Darlington pairs. Assume transistor parameters of $\beta(\text{npn}) = 120$, $\beta(\text{pnp}) = 80$, $V_A(\text{npn}) = 100 \text{ V}$, and $V_A(\text{pnp}) = 80 \text{ V}$. Let the power supply voltages be $\pm 10 \text{ V}$ and let $I_Q = 1 \text{ mA}$. (a) Determine the output resistance R_o . (b) Calculate the differential-mode voltage gain. (c) Find the differential-mode input resistance R_{id} .

*D11.71 Design a BJT diff-amp with an active load similar to the configuration in Figure P11.70 except that the input devices are to be pnp transistors and

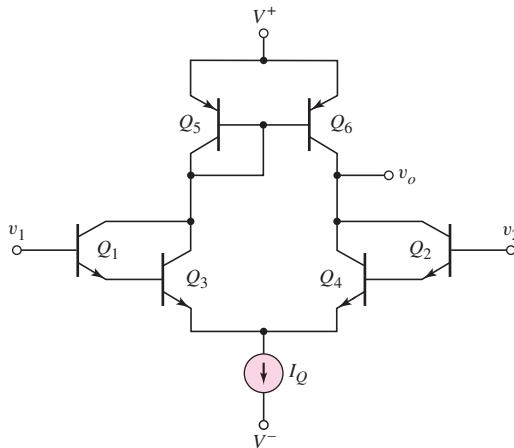


Figure P11.70

the active load will have npn transistors. Using the same parameters as in Problem 11.70, determine the small-signal differential-mode voltage gain.

- D11.72 Reconsider the diff-amp specifications listed in Problem 11.45. Design an all-CMOS diff-amp with the configuration in Figure 11.32 to meet the specifications. The NMOS transistor parameters are $V_{TN} = 0.4$ V, $k'_n = 100 \mu\text{A/V}^2$, and $\lambda_n = 0.025 \text{ V}^{-1}$. The parameters of the PMOS transistors are $V_{TP} = -0.4$ V, $k'_p = 40 \mu\text{A/V}^2$, and $\lambda_p = 0.04 \text{ V}^{-1}$.
- D11.73 An all-CMOS diff-amp, including the current source circuit, with the configuration in Figure 11.32 is to be designed to have a differential-mode gain of $A_d = 240$. The bias voltages are $V^+ = 3$ V and $V^- = -3$ V. The total power dissipation in the circuit is to be limited to 0.8 mW. Assume the NMOS transistor parameters are $V_{TN} = 0.4$ V, $k'_n = 100 \mu\text{A/V}^2$, and $\lambda_n = 0.02 \text{ V}^{-1}$. Assume PMOS transistor parameters of $V_{TP} = -0.4$ V, $k'_p = 40 \mu\text{A/V}^2$, and $\lambda_p = 0.03 \text{ V}^{-1}$.
- D11.74 The differential amplifier with the configuration shown in Figure 11.36 is to be designed to achieve a differential-mode voltage gain of $A_d = 400$. The circuit parameters are to be $V^+ = +5$ V, $V^- = -5$ V, and $I_Q = 200 \mu\text{A}$. The available transistors have parameters for the PMOS of $V_{TP} = -0.5$ V, $k'_p = 40 \mu\text{A/V}^2$, and $\lambda_p = 0.02 \text{ V}^{-1}$, and for the NMOS of $V_{TN} = +0.5$ V, $k'_n = 80 \mu\text{A/V}^2$, and $\lambda_n = 0.015 \text{ V}^{-1}$.
- *11.75 Consider the fully cascoded diff-amp in Figure 11.37. Assume $I_Q = 80 \mu\text{A}$ and transistor parameters of: $V_{TN} = 0.8$ V, $k'_n = 60 \mu\text{A/V}^2$, $\lambda_n = 0.015 \text{ V}^{-1}$, $V_{TP} = -0.8$ V, $k'_p = 25 \mu\text{A/V}^2$, and $\lambda_p = 0.02 \text{ V}^{-1}$. The transistor width-to-length ratios are $W/L = 60/4$ for transistors M_1-M_4 , $W/L = 40/4$ for transistors M_5-M_6 , and $W/L = 4/4$ for transistors M_7-M_8 . (a) Determine the output resistance of the diff-amp. (b) Calculate the differential-mode voltage gain of the diff-amp. (c) Find the common-mode voltage gain of the diff-amp using a computer simulation.
- 11.76 Consider the diff-amp that was shown in Figure P11.63. The circuit and transistor parameters are $V^+ = 2.8$ V, $V^- = -2.8$ V, $I_Q = 120 \mu\text{A}$, $K_n = K_p = 0.2 \text{ mA/V}^2$, $V_{TN} = +0.3$ V, $V_{TP} = -0.3$ V, and $\lambda_n = \lambda_p = 0.025 \text{ V}^{-1}$. (a) Determine the differential-mode voltage gain. (b) What is

the output resistance of the diff-amp? (c) What is the maximum common-mode voltage that may be applied?

- 11.77 The diff-amp in Figure P11.63 is biased at $I_Q = 0.5$ mA. The transistor parameters are $K_n = K_p = 0.25$ mA/V², $V_{TN} = 0.4$ V, $V_{TP} = -0.4$ V, and $\lambda_n = \lambda_p = 0.02$ V⁻¹. (a) What are the minimum power supply voltages if the common-mode input voltage is to be in the range ± 3 V? Assume symmetrical supply voltages. (b) Determine the differential-mode voltage gain.
- 11.78 The circuit and transistor parameters of the bipolar diff-amp shown in Figure P11.78 are $I_Q = 200$ μ A, $\beta_{npn} = 125$, $\beta_{pnp} = 80$, $V_{BE(on)} = V_{EB(on)} = 0.7$ V, $V_{AN} = 100$ V, and $V_{AP} = 60$ V. (a) What are the minimum power supply voltages (assume symmetrical supply voltages) if the common-mode input voltage is to be in the range of ± 2 V. (b) What is the differential-mode voltage gain?
- 11.79 Repeat Problem 11.78 if $I_Q = 120$ μ A, $V_{AN} = 75$ V, and $V_{AP} = 40$ V. All other parameters remain the same.

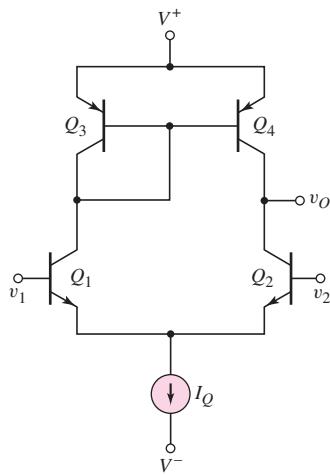


Figure P11.78

Section 11.5 BiCMOS Circuits

- 11.80 (a) The Darlington pair circuit in Figure 11.45 has new bias current levels of $I_{BIAS1} = 0.25$ mA and $I_{BIAS2} = 0.50$ mA. The transistor parameters are $K_n = 0.2$ mA/V², $V_{TN} = 0.4$ V, and $\lambda = 0$ for M_1 ; and $\beta = 150$, $V_{BE(on)} = 0.7$ V, and $V_A = \infty$ for M_2 . Determine the small-signal parameters for each transistor and find the composite transconductance. (b) Repeat part (a) for bias currents of $I_{BIAS1} = 0.05$ mA and $I_{BIAS2} = 0.50$ mA.
- 11.81 Consider the BiCMOS diff-amp in Figure 11.44, biased at $I_Q = 0.4$ mA. The transistor parameters for M_1 and M_2 are: $K_n = 0.2$ mA/V², $V_{TN} = 1$ V, and $\lambda = 0.01$ V⁻¹. The parameters for Q_1 and Q_2 are: $\beta = 120$, $V_{EB(on)} = 0.7$ V, and $V_A = 80$ V. (a) Determine the differential-mode voltage gain. (b) If the output resistance of the current source is $R_o = 500$ k Ω , determine the common-mode voltage gain using a computer simulation analysis.

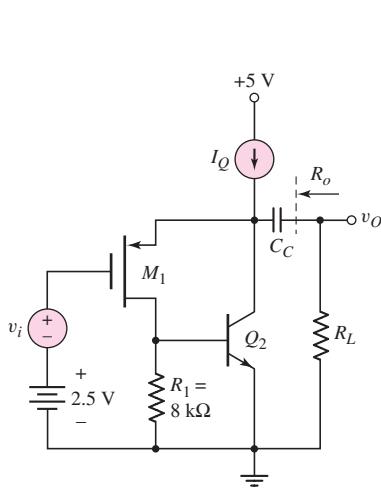


Figure P11.82

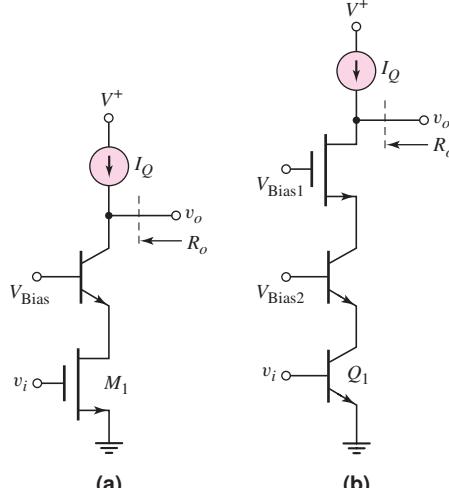


Figure P11.84

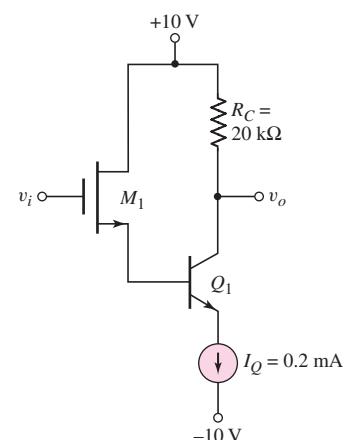


Figure P11.85

- 11.82 The BiCMOS circuit shown in Figure P11.82 is equivalent to a pnp bipolar transistor with an infinite input impedance. The bias current is $I_Q = 0.5 \text{ mA}$. The MOS transistor parameters are $V_{TP} = -0.5 \text{ V}$, $K_p = 0.7 \text{ mA/V}^2$, and $\lambda = 0$, and the BJT parameters are $\beta = 180$, $V_{BE(\text{on})} = 0.7 \text{ V}$, and $V_A = \infty$. (a) Sketch the small-signal equivalent circuit. (b) Calculate the small-signal parameters for each transistor. (c) Determine the small-signal voltage gain $A_v = v_o/v_i$ for (i) $R_L = 10 \text{ k}\Omega$ and (ii) $R_L = 100 \text{ k}\Omega$.
- 11.83 The bias current in the BiCMOS circuit shown in Figure P11.82 is $I_Q = 0.8 \text{ mA}$. The transistor parameters are the same as described in Problem 11.82. (a) Sketch the small-signal equivalent circuit and calculate the small-signal parameters for each transistor. (b) Determine the output resistance as defined in the figure.
- *11.84 The bias current I_Q is $25 \mu\text{A}$ in each circuit in Figure P11.84. The BJT parameters are $\beta = 100$ and $V_A = 50 \text{ V}$, and the MOSFET parameters are $V_{TN} = 0.8 \text{ V}$, $K_n = 0.25 \text{ mA/V}^2$, and $\lambda = 0.02 \text{ V}^{-1}$. Assume the two amplifying transistors M_1 and Q_1 are biased in the saturation region and forward-active region, respectively. Determine the small-signal voltage gain $A_v = v_o/v_i$ and the output resistance R_o for each circuit.
- 11.85 For the circuit shown in Figure P11.85, determine the small-signal voltage gain, $A_v = v_o/v_i$. Assume transistor parameters of $V_{TN} = 1 \text{ V}$, $K_n = 0.2 \text{ mA/V}^2$, and $\lambda = 0$ for M_1 and $\beta = 80$ and $V_A = \infty$ for Q_1 .

Section 11.6 Gain Stage and Simple Output Stage

- 11.86 The output stage in the circuit shown in Figure P11.86 is a Darlington pair emitter-follower configuration. Assume $\beta = 120$ for all npn transistors and $\beta = 90$ for all pnp transistors. Let $V_{A7} = 60 \text{ V}$ for Q_7 , $V_{A11} = 120 \text{ V}$ for Q_{11} , and $V_A = \infty$ for all other transistors. Determine the output resistance R_o .
- *11.87 For the circuit in Figure P11.87, the transistor parameters are $\beta = 100$ and $V_A = \infty$. The bias currents in the transistors are indicated on the figure.

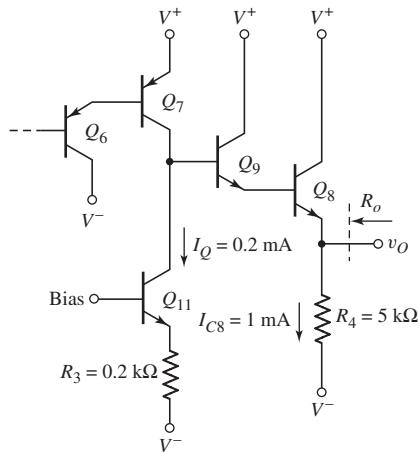


Figure P11.86

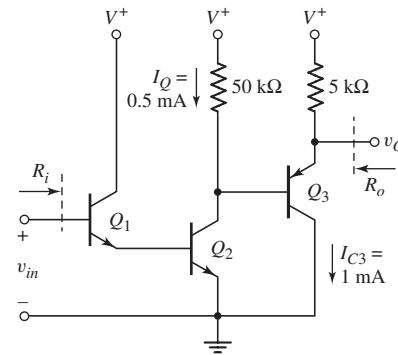


Figure P11.87

Determine the input resistance R_i , the output resistance R_o , and the small-signal voltage gain $A_v = v_o/v_{in}$.

- 11.88 Consider the circuit in Figure P11.88. The bias currents I_1 and I_2 are such that a zero dc output voltage is established. The transistor parameters are: $K_p = 0.2\text{ mA/V}^2$, $K_n = 0.5\text{ mA/V}^2$, $V_{TP} = -0.8\text{ V}$, $V_{TN} = +0.8\text{ V}$, and $\lambda_n = \lambda_p = 0.01\text{ V}^{-1}$. Determine the small-signal voltage gain $A_v = v_o/v_{in}$ and the output resistance R_o .

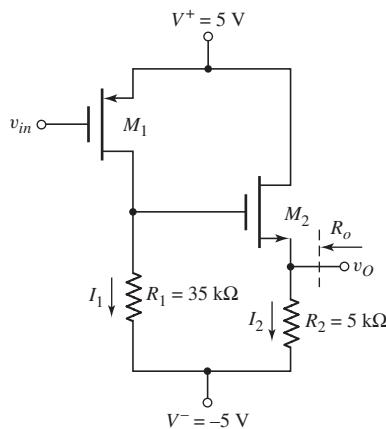


Figure P11.88

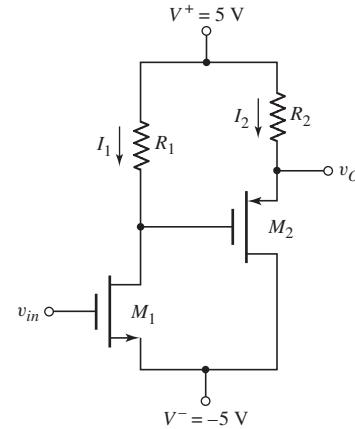


Figure P11.89

- 11.89 The bias currents in the circuit shown in Figure P11.89 are $I_1 = 0.25\text{ mA}$ and $I_2 = 1.0\text{ mA}$. The transistor parameters are $K_n = 0.5\text{ mA/V}^2$, $K_p = 1.0\text{ mA/V}^2$, $V_{TN} = 0.8\text{ V}$, $V_{TP} = -0.8\text{ V}$, and $\lambda_n = \lambda_p = 0.02\text{ V}^{-1}$. (a) Determine the resistor values R_1 and R_2 such that the dc value of the output voltage is zero. (b) Sketch the small-signal equivalent circuit and find the small-signal transistor parameters. (c) Determine the small-signal voltage gain $A_v = v_o/v_{in}$. (d) Determine the output resistance R_o .

Section 11.7 Simplified Op-Amp Circuits

- *11.90 Consider the multistage bipolar circuit in Figure P11.90, in which dc base currents are negligible. Assume the transistor parameters are $\beta = 120$, $V_{BE(on)} = 0.7$ V, and $V_A = \infty$. The output resistance of the constant current source is $R_o = 200$ k Ω . (a) For $v_1 = v_2 = -1.5$ V, design the circuit such that $v_{o2} = v_o = 0$, $I_{CQ3} = 0.25$ mA, and $I_{CQ4} = 2$ mA. (b) Assuming C_E acts as a short circuit, determine the differential-mode voltage gains $A_{d1} = v_{o2}/v_d$ and $A_d = v_o/v_d$. (c) Determine the common-mode gains $A_{cm1} = v_{o2}/v_d$ and $A_{cm} = v_o/v_d$, and the overall CMRR_{dB}.

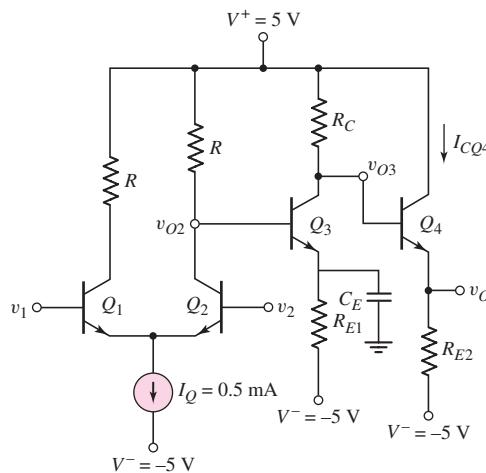


Figure P11.90

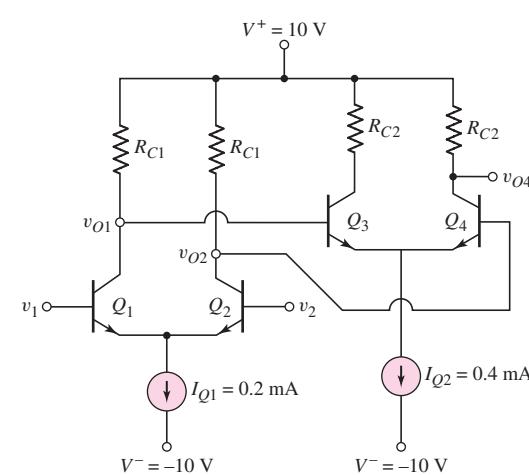


Figure P11.91

- *D11.91 The circuit in Figure P11.91 has two bipolar differential amplifiers in cascade, biased with ideal current sources I_{Q1} and I_{Q2} . Assume the transistor parameters are $\beta = 180$ and $V_A = \infty$. (a) Design the circuit such that $v_{o1} = v_{o2} = 2$ V and $v_{o4} = 6$ V when $v_1 = v_2 = 0$. (b) Determine the differential-mode voltage gains $A_{d1} = (v_{o1} - v_{o2})/v_d$ and $A_d = v_{o4}/v_d$.
- *11.92 The transistor parameters for the circuit in Figure P11.92 are: $\beta = 200$, $V_{BE(on)} = 0.7$ V, and $V_A = 80$ V. (a) Determine the differential-mode voltage gain $A_d = v_{o3}/v_d$ and the common-mode voltage gain $A_{cm} = v_{o3}/v_{cm}$. (b) Determine the output voltage v_{o3} if $v_1 = 2.015 \sin \omega t$ V and $v_2 = 1.985 \sin \omega t$ V. Compare this output to the ideal output that would be obtained if $A_{cm} = 0$. (c) Find the differential-mode and common-mode input resistances.
- *11.93 For the transistors in the circuit in Figure P11.93, the parameters are: $K_n = 0.2$ mA/V², $V_{TN} = 2$ V, and $\lambda = 0.02$ V⁻¹. (a) Determine the differential-mode voltage gain $A_d = v_{o3}/v_d$ and the common-mode voltage gain $A_{cm} = v_{o3}/v_{cm}$. (b) Determine the output voltage v_{o3} if $v_1 = 2.15 \sin \omega t$ V and $v_2 = 1.85 \sin \omega t$ V. Compare this output to the ideal output that would be obtained if $A_{cm} = 0$.

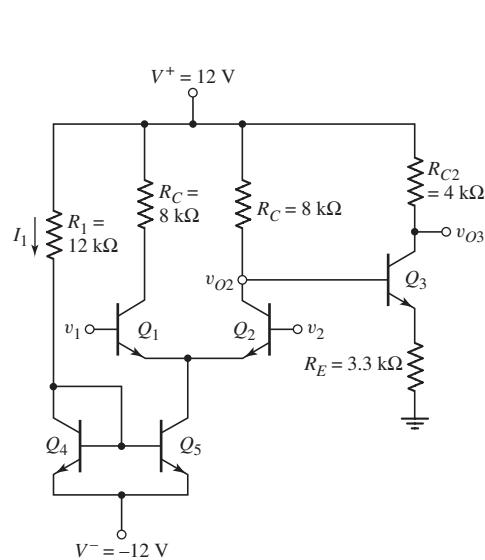


Figure P11.92

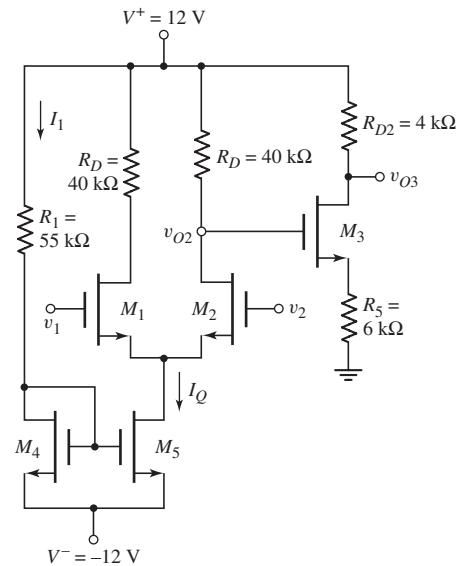


Figure P11.93

Section 11.8 Diff-Amp Frequency Response

- 11.94 Consider the differential amplifier in Figure 11.50(a) with parameters $I_Q = 0.8 \text{ mA}$, $R_C = 10 \text{ k}\Omega$, and $R_B = 0.5 \text{ k}\Omega$. The transistor parameters are $\beta = 150$, $V_{BE}(\text{on}) = 0.7 \text{ V}$, $V_A = \infty$, $C_\pi = 1.2 \text{ pF}$, and $C_\mu = 0.2 \text{ pF}$.
 (a) Determine the low-frequency differential-mode gain $A_d = v_{o2}/v_d$.
 (b) Find the equivalent Miller capacitance of each transistor.
 (c) Determine the upper 3 dB frequency.
- 11.95 The differential amplifier in Figure 11.51(a) has the same circuit and transistor parameters as described in Problem 11.94. The equivalent impedance of the current source is $R_o = 10 \text{ M}\Omega$ and $C_o = 0.4 \text{ pF}$.
 (a) Determine the frequency of the zero in the common-mode gain.
 (b) Find the frequency of the pole in the common-mode gain.
- 11.96 A BJT diff-amp is biased with a current source $I_Q = 2 \text{ mA}$, and the circuit parameters are $R_C = 10 \text{ k}\Omega$ and $R_B = 1 \text{ k}\Omega$. The transistor parameters are: $\beta = 120$, $f_T = 800 \text{ MHz}$, and $C_\mu = 1 \text{ pF}$.
 (a) Determine the upper 3 dB frequency of the differential-mode gain.
 (b) If the current source impedance parameters are $R_o = 10 \text{ M}\Omega$ and $C_o = 1 \text{ pF}$, find the frequency of the zero in the common-mode gain.
- 11.97 Consider the diff-amp in Figure 11.55. The circuit and transistor parameters are the same as in Problem 11.6. For a one-sided output at v_{o2} , determine the differential-mode gain for:
 (a) $R_E = 100 \Omega$, and
 (b) $R_E = 250 \Omega$.



COMPUTER SIMULATION PROBLEMS

- 11.98 Using a computer simulation, verify the results of Example 11.12.
- 11.99 Using a computer simulation, verify the results of Example 11.13 for the simple op-amp circuit.

- 11.100 Consider the circuit in Figure P11.100. Use standard transistors. Using a computer simulation, determine the small-signal differential-mode voltage gain and common-mode voltage gain for (a) $R_L = 10 \text{ M}\Omega$ and (b) $R_L = 200 \text{ k}\Omega$.

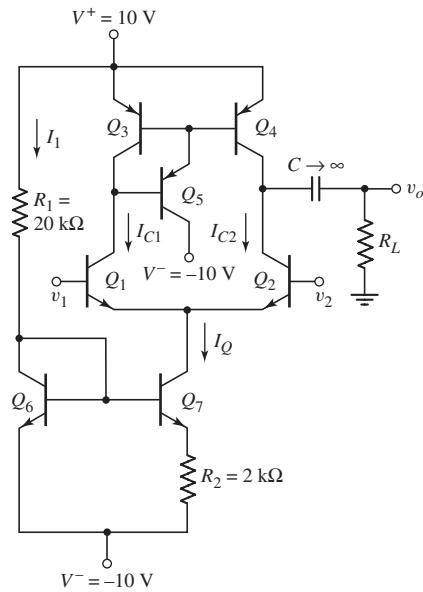


Figure P11.100

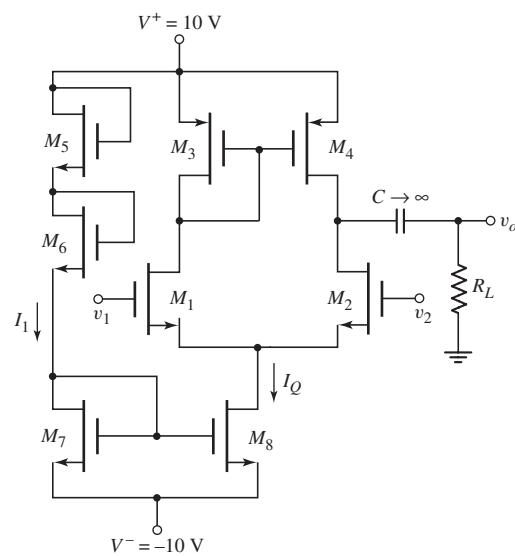


Figure P11.101

- 11.101 Consider the circuit in Figure P11.101. Use standard transistors. Using a computer simulation, determine the small-signal differential-mode voltage gain and common-mode voltage gain for (a) $R_L = 10 \text{ M}\Omega$ and (b) $R_L = 200 \text{ k}\Omega$.

DESIGN PROBLEMS

[Note: Each design is to be correlated with a computer simulation analysis.]

- *D11.102 Design a basic BJT diff-amp with an active load and constant current-source biasing. The bias voltages are to be $\pm 3 \text{ V}$ and the maximum power dissipation is to be limited to 2 mW . The open-circuit differential-mode voltage gain should be $|A_d| = 1800$ and the common-mode rejection ratio should be $\text{CMRR}_{\text{dB}} = 75 \text{ dB}$. Specify bias currents, minimum Early voltages, and the minimum output impedance of the current source. Design the current source to achieve the required output impedance.

- *D11.103 Design a basic MOSFET diff-amp with an active load and constant current-source biasing. The bias voltages are to be $\pm 3 \text{ V}$ and the maximum power dissipation is to be limited to 2 mW . The open-circuit differential-mode voltage gain should be $|A_d| = 180$ and the common-mode rejection ratio should be $\text{CMRR}_{\text{dB}} = 80 \text{ dB}$. Use appropriate transistor parameters. Specify bias currents, minimum λ values, and the minimum output impedance of the current source. Design the current source to achieve the required output impedance.

*D11.104 Consider the bipolar op-amp configuration in Figure 11.49. The bias voltages are ± 10 V, as shown, the current I_{R7} is to be $I_{R7} = 3$ mA, and the maximum dc power dissipation in the circuit is to be 120 mW. The output voltage is to be $v_o = 0$ for $v_1 = v_2 = 0$. Design the circuit, using reasonable resistance and current values. What is the overall differential-mode voltage gain?

*D11.105 The transistor parameters for the circuit in Figure P11.105 are: $K_n = 0.2$ mA/V², $V_{TN} = 0.8$ V, and $\lambda = 0$. The output resistance of the constant-current source is $R_o = 100$ k Ω . (a) For $v_1 = v_2 = 0$, design the circuit such that: $v_{o2} = 2$ V, $v_{o3} = 3$ V, $v_o = 0$, $I_{DQ3} = 0.25$ mA, and $I_{DQ4} = 2$ mA. (b) Determine the differential-mode gains $A_{d1} = v_{o2}/v_d$ and $A_d = v_o/v_d$. (c) Determine the common-mode voltage gains $A_{cm1} = v_{o2}/v_{cm}$ and $A_{cm} = v_o/v_{cm}$, and the overall CMRR_{dB}.

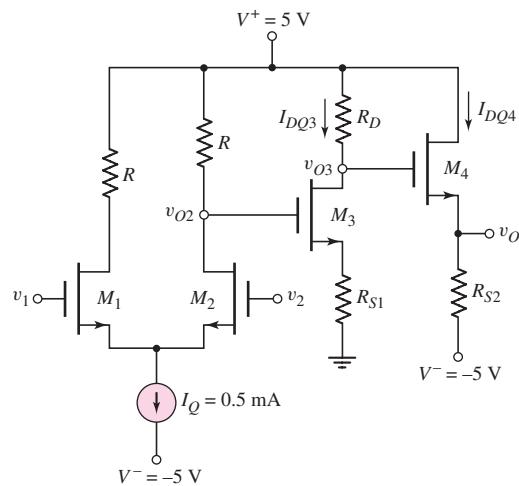


Figure P11.105

Digital Electronics

Part 2 of the text dealt with analog electronic circuits. Part 3 now deals with digital electronics, another important category of electronics.

Chapter 16 examines field-effect transistor digital circuits. MOSFET digital circuits have revolutionized digital electronics, with CMOS technology producing high-density, low-power digital circuits. Initially, we briefly consider the NMOS inverter and NMOS logic gates. We then analyze the basic CMOS inverter and then develop CMOS logic gates. Finally, we analyze FET shift registers and flip-flops and then discuss some basic A/D and D/A converters.

Bipolar digital circuits are considered in Chapter 17. We initially examine emitter-coupled logic, which is primarily used in specialized high-speed applications. We then briefly consider the basic aspects of transistor-transistor logic (TTL), which was the mainstay of logic design for many years. Low-power Schottky TTL circuits are analyzed in order to obtain a good comparison between FET and bipolar digital technologies.

MOSFET Digital Circuits

This chapter presents the basic concepts of MOSFET digital integrated circuits, which is the most widely used technology for the fabrication of digital systems. The small transistor size and low power dissipation of CMOS circuits allows for a high level of integration for logic and memory circuits. JFET logic circuits are very specialized and are therefore not considered here.

A discussion of NMOS logic circuits will serve as an introduction to the analysis and design of digital systems. This technology, although old, deals with only one type of transistor (n-channel) and therefore makes the analysis more straightforward than dealing with two types of transistors in the same circuit. This discussion will also serve as a baseline to point out advantages of CMOS technology.

Initially, we consider basic digital logic circuits such as NOR and NAND gates, and then discuss additional logic circuits such as flip flops, shift registers, and adders. Finally, we consider memories, and then A/D and D/A converters.

PREVIEW

In this chapter, we will:

- Analyze and design NMOS inverters
- Analyze and design NMOS logic gates
- Analyze and design CMOS inverters
- Analyze and design static CMOS logic gates
- Analyze and design clocked CMOS logic gates
- Analyze and understand the characteristics of NMOS and CMOS transmission gates
- Analyze and understand the characteristics of shift registers and various flip-flop designs
- Discuss semiconductor memories
- Analyze and design random-access memory (RAM) cells
- Analyze read-only memories (ROM)
- Discuss the basic concepts in A/D and D/A converters
- As an application, design a static CMOS logic gate to implement a specific logic function.

16.1 NMOS INVERTERS

Objective: • Analyze and design NMOS inverters.

The inverter is the basic circuit of most MOS logic circuits. The design techniques used in NMOS logic circuits are developed from the dc analysis results for the NMOS inverter. Extending the concepts developed from the inverter to NOR and NAND gates is then direct. Alternative inverter load elements are compared in terms of power consumption, packing density, and transfer characteristics.

16.1.1 n-Channel MOSFET Revisited

We studied the structure, operation, and characteristics of MOS transistors in Chapter 3. In this section, we will quickly review the n-channel MOSFET characteristics, emphasizing specific properties important in digital circuit design.

The simplified n-channel MOSFET that we have considered is shown in Figure 16.1(a). A more detailed view of the n-channel MOSFET is shown in Figure 16.1(b). The active transistor region is the surface of the semiconductor and comprises heavily doped n^+ source and drain regions and the p-type channel region. The channel length is L and the channel width is W . The body, or substrate, is a single-crystal silicon wafer, which is the starting material for circuit fabrication and provides physical support for the integrated circuit.

In an integrated circuit, all n-channel transistors are fabricated in the same p-type substrate material. The substrate is connected to the most negative potential in the circuit, which for digital circuits is normally at ground potential or zero volts. However, the source terminal of many transistors will not be at zero volts, which means that a reverse-biased pn junction will exist between the source and substrate. When the source and body terminals are not at the same potential, the threshold voltage of the transistor becomes a function of the source-to-body voltage. This **body effect** must then be taken into account in determining logic levels in digital circuits.

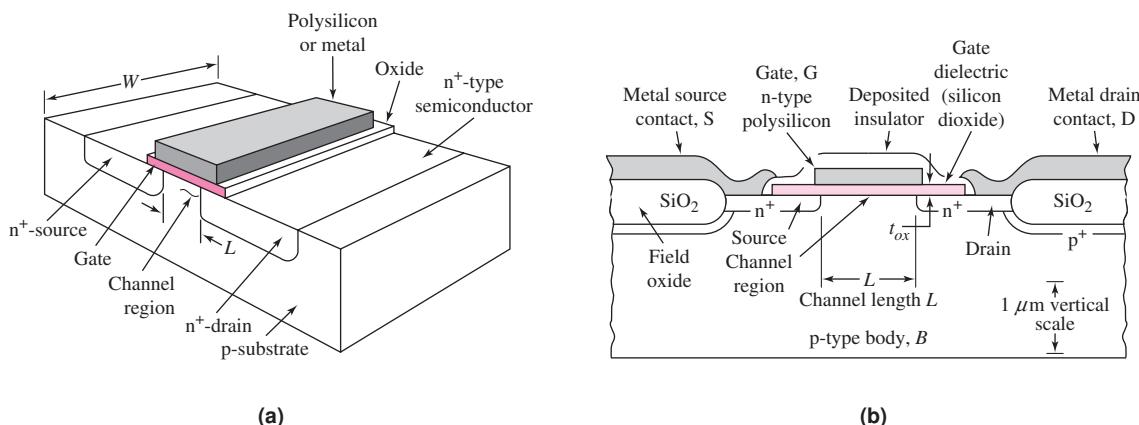


Figure 16.1 (a) n-channel MOSFET simplified view and (b) n-channel MOSFET detailed cross section

Current-Voltage Relation

The current-voltage characteristics of the n-channel MOSFET are functions of both the electrical and geometric properties of the device. When the transistor is biased in the nonsaturation region, for $v_{GS} \geq V_{TN}$ and $v_{DS} \leq (v_{GS} - V_{TN})$, we can write

$$i_D = K_n [2(v_{GS} - V_{TN})v_{DS} - v_{DS}^2] \quad (16.1(a))$$

In the saturation region, for $v_{GS} \geq V_{TN}$ and $v_{DS} \geq (v_{GS} - V_{TN})$, we have

$$i_D = K_n (v_{GS} - V_{TN})^2 \quad (16.1(b))$$

The transition point separates the nonsaturation and saturation regions and is the drain-to-source saturation voltage, which is given by

$$v_{DS} = v_{DS}(\text{sat}) = v_{GS} - V_{TN} \quad (16.2)$$

The term $(1 + \lambda v_{DS})$ is sometimes included in Equation (16.1(b)) to account for channel length modulation and the finite output resistance. In most cases, it has little effect on the operating characteristics of MOS digital circuits. In our analysis, the term λ is assumed to be zero unless otherwise stated.

The parameter K_n is the NMOS transistor conduction parameter and is given by

$$K_n = \left(\frac{1}{2} \mu_n C_{\text{ox}} \right) \left(\frac{W}{L} \right) = \frac{k'_n}{2} \frac{W}{L} \quad (16.3)$$

The electron mobility μ_n and oxide capacitance C_{ox} are assumed to be constant for all devices in a particular IC.

The current-voltage characteristics are directly related to the channel width-to-length ratio, or the size of the transistor. In general, in a given IC, the length L is fixed, but the designer can control the channel width W .

Since the MOS transistor is a majority carrier device, the switching speed of MOS digital circuits is limited by the time required to charge and discharge the capacitances between device electrodes and between interconnect lines and ground. Figure 16.2 shows the significant capacitances in a MOSFET. The capacitances C_{sb} and C_{db} are the source-to-body and drain-to-body n^+p junction capacitances. The total input gate capacitance, to a first approximation, is a constant equal to

$$C_g = WLC_{\text{ox}} = WL \left(\frac{\epsilon_{\text{ox}}}{t_{\text{ox}}} \right) \quad (16.4)$$

where C_{ox} is the oxide capacitance per unit area, and is a function of the oxide thickness. The parameter C_{ox} also appears in the expression for the conduction parameter.

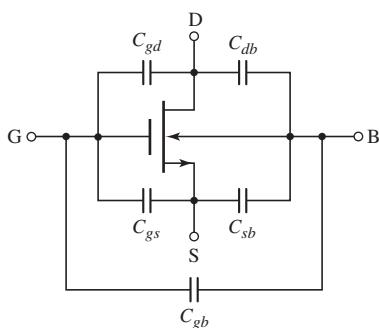


Figure 16.2 n-channel MOSFET and device capacitances

Small Geometry Effects

The current–voltage relationships given by Equations (16.1(a)), (16.1(b)), and (16.2) are first-order approximations that apply to “long” channel devices. The tendency in device design is to make the devices as small as possible, which means the channel length is being reduced to values substantially smaller than 1 μm . The corresponding channel widths are also being reduced. As the channel length is reduced, several effects alter the current–voltage characteristics. First, the threshold voltage becomes a function of the geometry of the device and is dependent on the channel length. This effect must be taken into account in the design of the transistor. Second, carrier velocity saturation reduces the saturation-mode current below the current value predicted by Equation (16.1(b)). The current is no longer a quadratic function of gate-to-source voltage, and tends to become a linear function of voltage. Channel length modulation means that the current tends to be larger than that predicted by the ideal equation. Third, the electron mobility is a function of the gate voltage so that the current tends to be smaller than the predicted value as the gate-to-source voltage increases. All of these effects complicate the analysis considerably.

We can, however, determine the basic operation and behavior of MOSFET logic circuits by using the first-order equations. We will use these first-order equations in our design of logic circuits. To determine the effect of small device size, a computer simulation may be performed in which the appropriate device models are incorporated in the simulation.

16.1.2 NMOS Inverter Transfer Characteristics

Since the inverter is the basis for most logic circuits, we will describe the NMOS inverter and will develop the dc transfer characteristics for three types of inverters with different load devices. This discussion will introduce voltage transfer functions and will define the maximum and minimum logic levels.

NMOS Inverter with Resistor Load

Figure 16.3(a) shows a single NMOS transistor connected to a resistor to form an inverter. The transistor characteristics and load line are shown in Figure 16.3(b), along with the parametric curve separating the saturation and nonsaturation regions. We determine the voltage transfer characteristics of the inverter by examining the various regions in which the transistor can be biased.

When the input voltage is less than or equal to the threshold voltage, or $v_I \leq V_{TN}$, the transistor is cut off, $i_D = 0$, and the output voltage is $v_O = V_{DD}$. The maximum output voltage is defined as the **logic 1** level. As the input voltage becomes

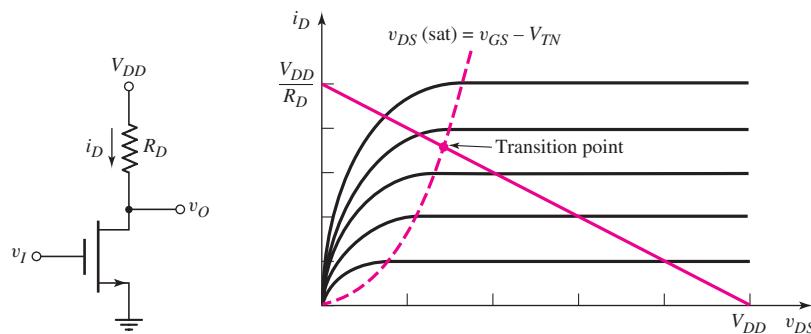


Figure 16.3 (a) NMOS inverter with resistor load and (b) transistor characteristics and load line

just greater than V_{TN} , the transistor turns on and is biased in the saturation region. The output voltage is then

$$v_o = V_{DD} - i_D R_D \quad (16.5)$$

where the drain current is given by

$$i_D = K_n(v_{GS} - V_{TN})^2 = K_n(v_I - V_{TN})^2 \quad (16.6)$$

Combining Equations (16.5) and (16.6) yields

$$v_o = V_{DD} - K_n R_D (v_I - V_{TN})^2 \quad (16.7)$$

which relates the output and input voltages as long as the transistor is biased in the saturation region.

As the input voltage increases, the Q -point of the transistor moves up the load line. At the transition point, we have

$$V_{Ot} = V_{It} - V_{TN} \quad (16.8)$$

where V_{Ot} and V_{It} are the drain-to-source and gate-to-source voltages, respectively, at the transition point. Substituting Equation (16.8) into (16.7), we determine the input voltage at the transition point from

$$K_n R_D (V_{It} - V_{TN})^2 + (V_{It} - V_{TN}) - V_{DD} = 0 \quad (16.9)$$

As the input voltage becomes greater than V_{It} , the Q -point continues to move up the load line, and the transistor becomes biased in the nonsaturation region. The drain current is then

$$i_D = K_n [2(v_{GS} - V_{TN})v_{DS} - v_{DS}^2] = K_n [2(v_I - V_{TN})v_O - v_O^2] \quad (16.10)$$

Combining Equations (16.5) and (16.10) yields

$$v_o = V_{DD} - K_n R_D [2(v_I - V_{TN})v_O - v_O^2] \quad (16.11)$$

which relates the input and output voltages as long as the transistor is biased in the nonsaturation region.

Figure 16.4 shows the voltage transfer characteristics of this inverter for three resistor values. Also shown is the line, given by Equation (16.8), which separates the

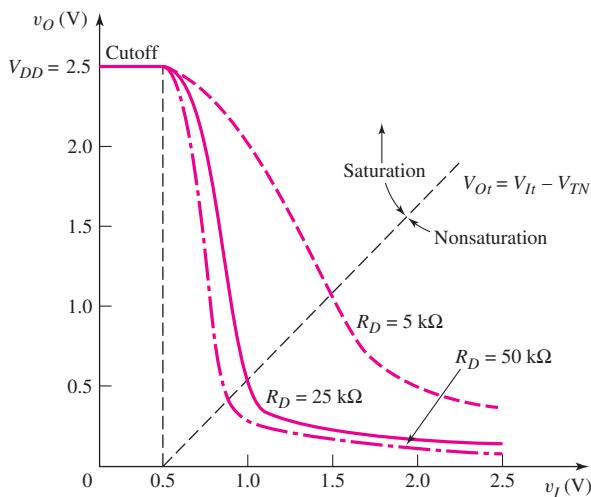


Figure 16.4 Voltage transfer characteristics of NMOS inverter with resistive load, using parameters from Example 16.1 and for the three resistor values.

saturation and nonsaturation bias regions of the transistor. The figure shows that the minimum output voltage, or the logic 0 level, for a high input decreases with increasing load resistance, and the sharpness of the transition region between a low input and a high input increases with increasing load resistance.

It should be noted that a large resistance is difficult to fabricate in an IC. A large resistor value in the inverter will limit current and power consumption as well as provide a small V_{OL} value. But it would also require a large chip area if fabricated in a standard MOS process. To avoid this problem MOS transistors can be used as load devices, replacing the resistor, as discussed in subsequent paragraphs.

EXAMPLE 16.1

Objective: Determine the transition point, minimum output voltage, maximum drain current, and maximum power dissipation of an NMOS inverter with resistor load.

Specifications: Consider the circuit in Figure 16.3(a) with parameters $V_{DD} = 2.5$ V and $R_D = 20 \text{ k}\Omega$. The transistor parameters are $V_{TN} = 0.5$ V and $K_n = 0.3 \text{ mA/V}^2$.

Solution: The input voltage at the transition point is found from Equation (16.9). We have

$$(0.3)(25)(V_{I_t} - 0.5)^2 + (V_{I_t} - 0.5) - 2.5 = 0$$

which yields

$$V_{I_t} - 0.5 = 0.515 \text{ V} \quad \text{or} \quad V_{I_t} = 1.015 \text{ V}$$

The output voltage at the transition point is

$$V_{O_t} = V_{I_t} - V_{TN} = 1.015 - 0.5 = 0.515 \text{ V}$$

When v_I is high at $v_I = 2.5$ V, the output voltage is found from Equation (16.11). We find

$$v_O = 2.5 - (0.3)(25)[2(2.5 - 0.5)v_O - v_O^2]$$

which yields the output low level as

$$v_O = V_{OL} = 82.3 \text{ mV}$$

The maximum drain current in the inverter occurs when $v_O = V_{OL}$ and is found to be

$$i_{D,\max} = \frac{2.5 - 0.0823}{25} \Rightarrow 96.7 \mu\text{A}$$

The maximum power dissipated in the inverter is

$$P_{D,\max} = i_{D,\max} \cdot V_{DD} = (0.0967)(2.5) = 0.242 \text{ mW}$$

Comment: The level of V_{OL} is less than the threshold voltage V_{TN} ; therefore, if the output of this inverter is used to drive a similar inverter, the driver transistor of the load inverter would be cut off and its output would be high, which is the desired condition. We will compare the maximum drain currents and maximum power dissipations of the three basic NMOS inverters.

EXERCISE PROBLEM

Ex 16.1: Consider the NMOS inverter with resistor load in Figure 16.3(a) biased at $V_{DD} = 3$ V. Assume transistor parameters of $k'_n = 100 \mu\text{A}/\text{V}^2$, $W/L = 4$, and $V_{TN} = 0.5$ V. (a) Find the value of R_D such that $v_O = 0.1$ V when $v_I = 3$ V. (b) Using the results of part (a), determine the maximum current and maximum power dissipation in the inverter. (c) Using the results of part (a), determine the transition point for the driver transistor. (Ans. (a) $R_D = 29.6 \text{ k}\Omega$; (b) $i_{D,\text{max}} = 0.098 \text{ mA}$, $P_{D,\text{max}} = 0.294 \text{ mW}$; (c) $V_{I_t} = 1.132 \text{ V}$, $V_{O_t} = 0.632 \text{ V}$)

An n-channel enhancement-mode MOSFET with the gate connected to the drain can be used as a load device in an NMOS inverter. This device configuration was analyzed in Chapter 3. We found that, when $v_{GS} = v_{DS} \geq V_{TN}$, the transistor always operates in the saturation region. The drain current is given by

$$i_D = K_n(v_{GS} - V_{TN})^2 = K_n(v_{DS} - V_{TN})^2 \quad (16.12)$$

We continue to neglect the effect of the output resistance and the λ parameter.

Figure 16.5(a) shows an NMOS inverter with the enhancement load device. The driver transistor parameters are denoted by V_{TND} and K_D , and the load transistor parameters are denoted by V_{TNL} and K_L . The substrate connections are not shown. In the following analysis, we neglect the body effect and we assume all threshold voltages are constant. These assumptions do not seriously affect the basic analysis, nor the inverter characteristics.

The driver transistor characteristics and the load curve are shown in Figure 16.5(b). When the inverter input voltage is less than the driver threshold voltage, the driver is cut off and the drain currents are zero. From Equation (16.12), we have

$$i_{DL} = 0 = K_L(v_{DSL} - V_{TNL})^2 \quad (16.13)$$

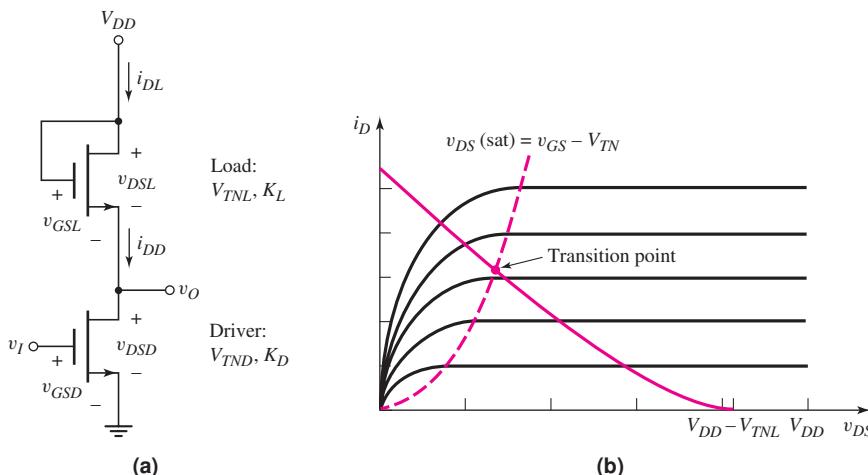


Figure 16.5 (a) NMOS inverter with saturated load and (b) driver transistor characteristics and load curve

From Figure 16.5(a), we see that $v_{DSL} = V_{DD} - v_O$, which means that

$$v_{DSL} - V_{TNL} = V_{DD} - v_O - V_{TNL} = 0 \quad (16.14(a))$$

The **maximum output voltage** is then

$$v_{O,\max} \equiv V_{OH} = V_{DD} - V_{TNL} \quad (16.14(b))$$

For the enhancement-load NMOS inverter, the maximum output voltage, which is the **logic 1 level**, does not reach the full V_{DD} value. This cutoff point is shown in the load curve in Figure 16.5(b).

As the input voltage becomes just greater than the **driver threshold voltage** V_{TND} , the driver transistor turns on and is biased in the saturation region. In steady-state, the two drain currents are equal since the output will be connected to the gates of other MOS transistors. We have $i_{DD} = i_{DL}$, which can be written as

$$K_D(v_{GSD} - V_{TND})^2 = K_L(v_{GSL} - V_{TNL})^2 \quad (16.15)$$

Equation (16.15) is expressed in terms of the individual transistor parameters. In terms of the input and output voltages, the expression becomes

$$K_D(v_I - V_{TND})^2 = K_L(V_{DD} - v_O - V_{TNL})^2 \quad (16.16)$$

Solving for the output voltage yields

$$v_O = V_{DD} - V_{TNL} - \sqrt{\frac{K_D}{K_L}(v_I - V_{TND})} \quad (16.17)$$

As the input voltage increases, the driver Q -point moves up the load curve and the output voltage decreases linearly with v_I .

At the driver transition point, we have

$$v_{DSD}(\text{sat}) = v_{GSD} - V_{TND}$$

or

$$V_{Ot} = V_{It} - V_{TND} \quad (16.18)$$

Substituting Equation (16.18) into (16.17), we find the input voltage at the transition point, which is

$$V_{It} = \frac{V_{DD} - V_{TNL} + V_{TND} \left(1 + \sqrt{\frac{K_D}{K_L}} \right)}{1 + \sqrt{\frac{K_D}{K_L}}} \quad (16.19)$$

As the input voltage becomes greater than V_{It} , the driver transistor Q -point continues to move up the load curve and the driver becomes biased in the nonsaturation region. Since the driver and load drain currents are still equal, or $i_{DD} = i_{DL}$, we now have

$$K_D[2(v_{GSD} - V_{TND})v_{DSD} - v_{DSD}^2] = K_L(v_{DSL} - V_{TNL})^2 \quad (16.20)$$

Writing Equation (16.20) in terms of the input and output voltages produces

$$K_D[2(v_I - V_{TND})v_O - v_O^2] = K_L(V_{DD} - v_O - V_{TNL})^2 \quad (16.21)$$

Obviously, the relationship between v_I and v_O in this region is not linear.

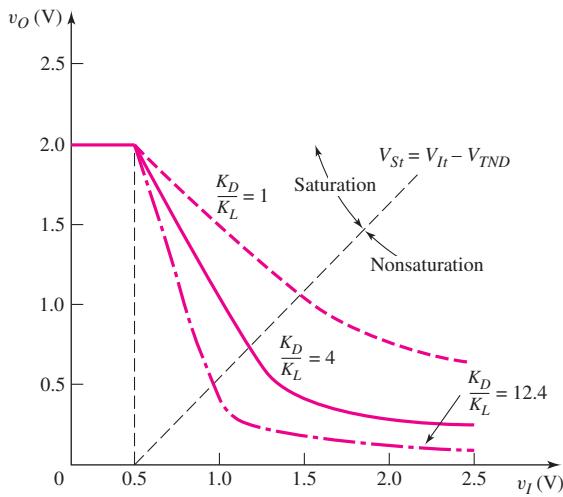


Figure 16.6 Voltage transfer characteristics of NMOS inverter with saturated load, using parameters from Example 16.2 and for three aspect ratios.

Figure 16.6 shows the voltage transfer characteristics of this inverter for three K_D -to- K_L ratios. The ratio K_D/K_L is the aspect ratio and is related to the width-to-length parameters of the driver and load transistors.

The line, given by Equation (16.18), separating the driver saturation and nonsaturation regions is also shown in the figure. We see that the minimum output voltage, or the logic 0 level, for a high input decreases with an increasing K_D/K_L ratio. As the width-to-length ratio of the load transistor decreases, the effective resistance increases, which means that the general behavior of the transfer characteristics is the same as for the resistor load. However, the high output voltage is

$$V_{OH} = V_{DD} - V_{TNL}$$

When the driver is biased in the saturation region, we find the slope of the transfer curve, which is the **inverter gain**, by taking the derivative of Equation (16.17) with respect to v_I . We see that

$$dv_O/dv_I = -\sqrt{K_D/K_L}$$

When the aspect ratio is greater than unity, the inverter gain magnitude is greater than unity. A logic circuit family with an inverter transfer curve that exhibits a gain greater than unity for some region is called a **restoring logic family**. Restoring logic is so named because logic signals that are degraded for some reason in one circuit can be restored by the gain of subsequent logic circuits.

DESIGN EXAMPLE 16.2

Objective: Design an NMOS inverter to meet a set of specifications and determine the power dissipation in the inverter.

Specifications: The NMOS inverter with saturated load shown in Figure 16.5(a) is to be designed such that $v_O = 0.1$ V when $v_I = 2.0$ V. The circuit is biased at $V_{DD} = 2.5$ V. (Neglect the body effect.)

Choices: Transistors are available with parameters $V_{TN} = 0.5 \text{ V}$ and $k'_n = 100 \mu\text{A}/\text{V}^2$.

Solution: The maximum output voltage (defined as a logic 1), neglecting the body effect, is

$$V_{OH} = V_{DD} - V_{TNL} = 2.5 - 0.5 = 2.0 \text{ V}$$

For $v_I = 2.0 \text{ V}$, the driver is biased in the **nonsaturation** region and the load is always biased in the saturation region. Setting the two drain currents equal to each other, we find, using Equation (16.21),

$$K_D[2(2.0 - 0.5)(0.1) - (0.1)^2] = K_L(2.5 - 0.1 - 0.5)^2$$

which yields

$$\frac{K_D}{K_L} = 12.4$$

If we choose $(W/L)_L = 1$, and since

$$\frac{K_D}{K_L} = \frac{(W/L)_D}{(W/L)_L}$$

then we have

$$\left(\frac{W}{L}\right)_D = 12.4$$

The maximum inverter current occurs for $v_O = V_{OL} = 0.1 \text{ V}$ and is found from

$$\begin{aligned} i_{D,\max} &= \frac{k'_n}{2} \cdot \left(\frac{W}{L}\right)_D [2(v_I - V_{TND})v_O - v_O^2] \\ &= \left(\frac{0.1}{2}\right) (12.4) [2(2.0 - 0.5)(0.1) - (0.1)^2] = 0.180 \text{ mA} \end{aligned}$$

The maximum **power dissipated** in the inverter is

$$P_{D,\max} = i_{D,\max} \cdot V_{DD} = (0.18)(2.5) = 0.45 \text{ mW}$$

Comment: In the NMOS inverter with enhancement-mode load, a relatively large difference in sizes of the driver and load transistors is required to produce a relatively low output voltage V_{OL} . The load transistor width-to-length ratio cannot be reduced substantially, so the maximum power dissipation cannot be substantially reduced from the 0.45 mW.

EXERCISE PROBLEM

Ex 16.2: The enhancement-load NMOS inverter in Figure 16.5(a) is biased at $V_{DD} = 3 \text{ V}$. The transistor parameters are $k'_n = 100 \mu\text{A}/\text{V}^2$, $V_{TND} = V_{TNL} = 0.4 \text{ V}$, $(W/L)_D = 16$, and $(W/L)_L = 2$. (a) Determine v_O when (i) $v_I = 0.1 \text{ V}$ and (ii) $v_I = 2.6 \text{ V}$. Neglect the body effect. (b) Determine the maximum current and maximum power dissipation in the inverter. (c) Determine the transition point for the driver transistor. (Ans. (a) (i) $v_O = 2.6 \text{ V}$, (ii) $v_O = 0.174 \text{ V}$; (b) $i_{D,\max} = 0.589 \text{ mA}$, $P_{D,\max} = 1.766 \text{ mW}$; (c) $V_{It} = 1.08 \text{ V}$, $V_{Ot} = 0.68 \text{ V}$)

NMOS Inverter with Depletion Load

Depletion-mode MOSFETs can also be used as load elements in NMOS inverters. Figure 16.7(a) shows the NMOS inverter with depletion load. The gate and source of the depletion-mode transistor are connected together. The driver transistor is still an enhancement-mode device. As before, the driver transistor parameters are V_{TND} ($V_{TND} > 0$) and K_D , and the load transistor parameters are V_{TNL} ($V_{TNL} < 0$) and K_L . Again, the substrate connections are not shown. The fabrication process for this inverter is slightly more complicated than for the enhancement-load inverter, since the threshold voltages of the two devices are not equal. However, as we will see, the advantages of this inverter make the extra processing steps worthwhile. This inverter has been the basis of many microprocessor and static memory designs.

The current–voltage characteristic curve for the depletion load, neglecting the body effect, is shown in Figure 16.7(b). Since the gate is connected to the source, $v_{GSL} = 0$, and the Q -point of the load is on this particular curve.

The driver transistor characteristics and the ideal load curve are shown in Figure 16.7(c). When the inverter input is less than the driver threshold voltage, the driver is cut off and the drain currents are zero. From Figure 16.7(b), we see that for

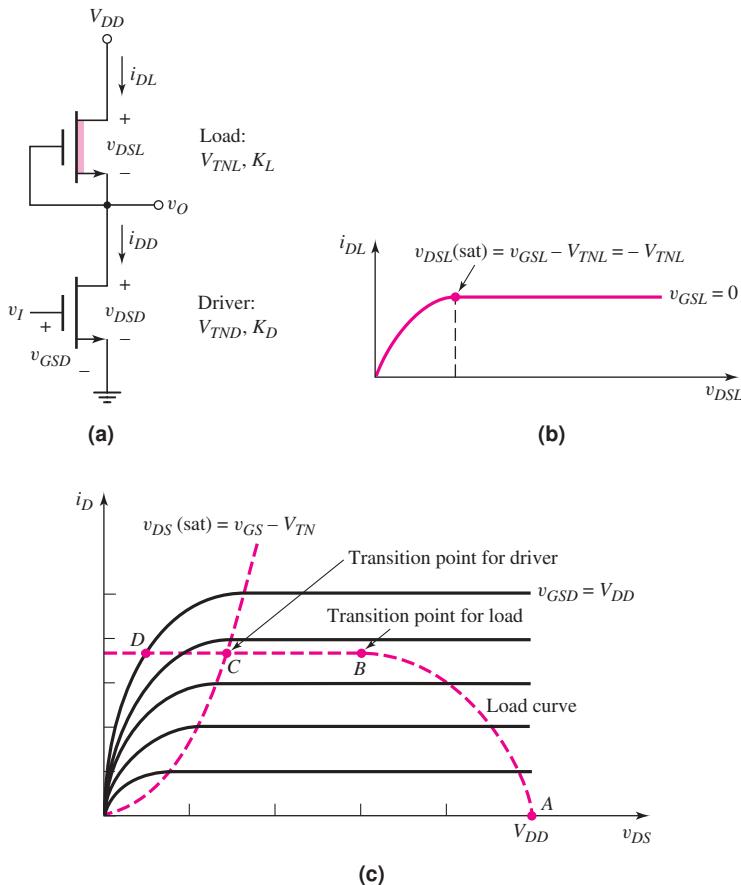


Figure 16.7 (a) NMOS inverter with depletion load, (b) current–voltage characteristic of depletion load, and (c) driver transistor characteristics and load curve

$i_D = 0$, the drain-to-source voltage of the load transistor must be zero; therefore, $v_O = V_{DD}$ for $v_I \leq V_{TND}$. An advantage of the depletion-load inverter over the enhancement-load inverter is that the high output voltage, or the logic 1 level, is at the full V_{DD} value.

As the input voltage becomes just greater than the driver threshold voltage V_{TND} , the driver turns on and is biased in the saturation region; however, the load is biased in the nonsaturation region. The Q -point lies between points A and B on the load curve shown in Figure 16.7(c). We again set the two drain currents equal, or $i_{DD} = i_{DL}$, which means that

$$K_D[v_{GSD} - V_{TND}]^2 = K_L[2(v_{GSL} - V_{TNL})v_{DSL} - v_{DSL}^2] \quad (16.22)$$

Writing Equation (16.22) in terms of the input and output voltages yields

$$K_D[v_I - V_{TND}]^2 = K_L[2(-V_{TNL})(V_{DD} - v_O) - (V_{DD} - v_O)^2] \quad (16.23)$$

This equation relates the input and output voltages as long as the driver is biased in the saturation region and the load is biased in the nonsaturation region.

There are two transition points for the NMOS inverter with a depletion load: one for the load and one for the driver. These are points B and C , respectively, in Figure 16.7(c). The transition point for the load is given by

$$v_{DSL} = V_{DD} - V_{Ot} = v_{GSL} - V_{TNL} = -V_{TNL} \quad (16.24(a))$$

or

$$V_{Ot} = V_{DD} + V_{TNL} \quad (16.24(b))$$

Since V_{TNL} is negative, the output voltage at the transition point is less than V_{DD} . The transition point for the driver is given by

$$v_{DSD} = v_{GSD} - V_{TND}$$

or

$$V_{Ot} = V_{It} - V_{TND} \quad (16.25)$$

When the Q -point lies between points B and C on the load curve, both devices are biased in the saturation region, and

$$K_D(v_{GSD} - V_{TND})^2 = K_L(v_{GSL} - V_{TNL})^2 \quad (16.26(a))$$

or

$$\sqrt{\frac{K_D}{K_L}}(v_I - V_{TND}) = -V_{TNL} \quad (16.26(b))$$

Equation (16.26(b)) demonstrates that the input voltage is a constant as the Q -point passes through this region. This effect is also shown in Figure 16.7(c); the load curve between points B and C lies on a constant v_{GSD} curve. (This characteristic will change when the body effect is taken into account.)

For an input voltage greater than the value given by Equation (16.26(b)), the driver is biased in the nonsaturation region while the load is biased in the saturation region. The Q -point is now between points C and D on the load curve in Figure 16.7(c). Equating the two drain currents, we obtain

$$K_D[2(v_{GSD} - V_{TND})v_{DSD} - v_{DSD}^2] = K_L(v_{GSL} - V_{TNL})^2 \quad (16.27(a))$$

which becomes

$$\frac{K_D}{K_L}[2(v_I - V_{TND})v_O - v_O^2] = (-V_{TNL})^2 \quad (16.27(b))$$

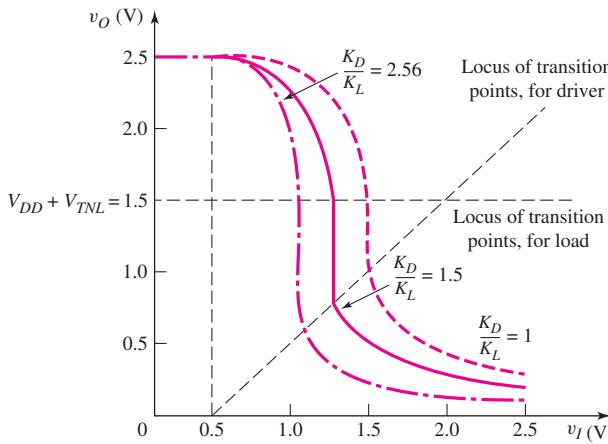


Figure 16.8 Voltage transfer characteristics of an NMOS inverter with depletion load, using parameters from Example 16.3 and for three aspect ratios

This equation implies that the relationship between the input and output voltages are not linear in this region.

Figure 16.8 shows the voltage transfer characteristics of this inverter for three values of K_D/K_L . Also shown are the loci of transition points for the load and driver transistors as given by Equations (16.24(b)) and (16.25), respectively.

DESIGN EXAMPLE 16.3

Objective: Design an NMOS inverter to meet a set of specifications and determine the power dissipation in the inverter.

Specifications: The NMOS inverter with depletion load shown in Figure 16.7(a) is to be designed such that $v_O = V_{OL} = 0.10$ V when $v_I = 2.5$ V. The circuit is biased at $V_{DD} = 2.5$ V. (Neglect the body effect.)

Choices: Transistors are available with process conduction parameters of $k'_n = 100 \mu\text{A}/\text{V}^2$. The driver transistor threshold voltage is $V_{TND} = 0.5$ V and the load transistor threshold voltage is $V_{TNL} = -1$ V.

Solution: For $v_I = 2.5$ V, the driver transistor is biased in the nonsaturation region and the load transistor is biased in the saturation region. Using Equation (16.27(b)), we find

$$K_D[2(2.5 - 0.5)(0.1) - (0.1)^2] = K_L[0 - (-1)]^2$$

which yields

$$\frac{K_D}{K_L} = 2.56$$

If we choose $(W/L)_L = 1$, then

$$\frac{K_D}{K_L} = \frac{(W/L)_D}{(W/L)_L} \Rightarrow 2.56 = \frac{(W/L)_D}{1} \Rightarrow \left(\frac{W}{L}\right)_D = 2.56$$

The maximum current in the inverter occurs when the output is in its low state, so, from the load transistor, we find

$$i_{D,\max} = \frac{k'_n}{2} \cdot \left(\frac{W}{L} \right)_L (0 - V_{TNL})^2 = \left(\frac{100}{2} \right) (1)[0 - (-1)]^2 = 50 \mu\text{A}$$

The maximum power dissipation in the inverter is

$$P_{D,\max} = i_{D,\max} \cdot V_{DD} = (50)(2.5) = 125 \mu\text{W}$$

Comment: A relatively low output voltage V_{OL} can be produced in the NMOS inverter with depletion load, even when the load and driver transistors are not vastly different in size. The power dissipation in this inverter is also substantially less than in the enhancement-load inverter since the aspect ratio is smaller.

Design Consideration: The static analysis of the three types of NMOS inverters clearly demonstrates the advantage of the depletion load inverter. The size of the driver transistor is smaller for a given load device size to produce a given low output state. This allows a greater number of inverters to be fabricated in a given chip area. In addition, since the power dissipation is less, more inverters can be fabricated on a chip for a given total power dissipation.

EXERCISE PROBLEM

Ex 16.3: The depletion-load NMOS inverter shown in Figure 16.7(a) is biased at $V_{DD} = 3$ V. The transistor parameters are $k'_n = 100 \mu\text{A/V}^2$, $V_{TND} = 0.4$ V, $V_{TNL} = -0.8$ V, $(W/L)_D = 6$, and $(W/L)_L = 2$. (a) Determine v_O for $v_I = 3$ V. Neglect the body effect. (b) Determine the maximum current and maximum power dissipation in the inverter. (c) Find the transition points for the driver and load transistors. (Ans. (a) $v_O = 0.0414$ V; (b) $i_{D,\max} = 0.064$ mA, $P_{D,\max} = 0.192$ mW; (c) driver: $V_{I_t} = 0.862$ V, $V_{O_t} = 0.462$ V; load: $V_{I_t} = 0.862$ V, $V_{O_t} = 2.2$ V)

16.1.3 Body Effect

Up to this point, we have neglected the body effect and assumed that all threshold voltages are constant. Figure 16.9 shows enhancement-load and depletion-load

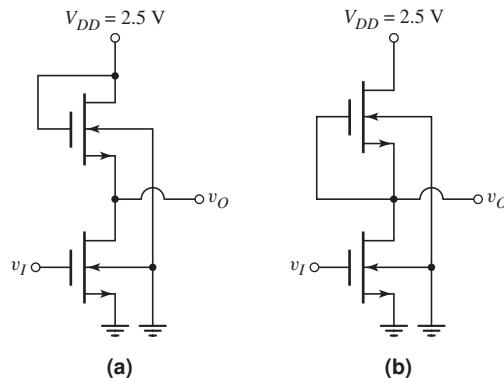


Figure 16.9 NMOS inverters, showing substrate connections to ground potential: (a) enhancement-load inverter and (b) depletion-load inverter

NMOS inverters with the substrates of all transistors tied to ground. A nonzero source-to-body voltage will then exist in the load devices. In fact, the source terminal of the depletion load can increase to V_{DD} . The threshold voltage equation, including the body effect, must be used in the circuit calculations for the load transistor. This significantly complicates the equations for the voltage transfer calculations, making them very cumbersome for hand analyses.

EXAMPLE 16.4

Objective: Determine the change in the high output voltage of an NMOS inverter with enhancement load, taking the body effect into account.

Consider the NMOS inverter with enhancement load in Figure 16.9(a). The transistor parameters are $V_{TNDO} = V_{TNLO} = 0.5$ V and $K_D/K_L = 16$. Assume the inverter is biased at $V_{DD} = 2.5$ V, assume the body effect coefficient is $\gamma = 0.5 \text{ V}^{1/2}$, and let $\phi_{fp} = 0.365$ V.

Solution: When $v_I < V_{TNDO}$, the driver is cut off and the output goes high. From Equation (16.14(b)), the maximum output voltage is

$$v_{O,\max} = V_{OH} = V_{DD} - V_{TNL}$$

where V_{TNL} is given by

$$V_{TNL} = V_{TNLO} + \gamma \left[\sqrt{2\phi_{fp} + V_{SB}} - \sqrt{2\phi_{fp}} \right]$$

From Figure 16.9(a), we see that $V_{SB} = v_O$. Therefore, Equation (16.14(b)) can be written as

$$v_{O,\max} = V_{DD} - \left\{ V_{TNLO} + \gamma \left[\sqrt{2\phi_{fp} + v_{O,\max}} - \sqrt{2\phi_{fp}} \right] \right\}$$

Defining $v_{O,\max} \equiv V_{OH}$, we have

$$V_{OH} - 2.427 = -0.5 \sqrt{0.73 + V_{OH}}$$

Squaring both sides and rearranging terms yields

$$V_{OH}^2 - 5.1044V_{OH} + 5.7088 = 0$$

Consequently, the maximum output voltage, or the logic 1 level, is

$$V_{OH} = 1.655 \text{ V}$$

Comment: Neglecting the body effect, the logic 1 output level is

$$V_{OH} = V_{DD} - V_{TNLO} = 2.5 - 0.5 = 2.0 \text{ V}$$

The body effect, then, can significantly influence the logic high state of the NMOS inverter with enhancement load. These results also impact the inverter noise margins.

The source and body terminals of the depletion load device in the NMOS inverter shown in Figure 16.9(b) are not at the same potential when the output goes high. However, when the driver is cut off, the drain-to-source voltage of the depletion device must be zero in order that $v_{O,\max} = V_{OH} = V_{DD}$.

Computer Simulation: A computer analysis of the inverters in Figure 16.9 was performed, neglecting the body effect and taking the body effect into account. The parameters are $V_{DD} = 5$ V, $V_{TNDO} = 0.8$ V for the driver transistors, $V_{TNLO} = 0.8$ V

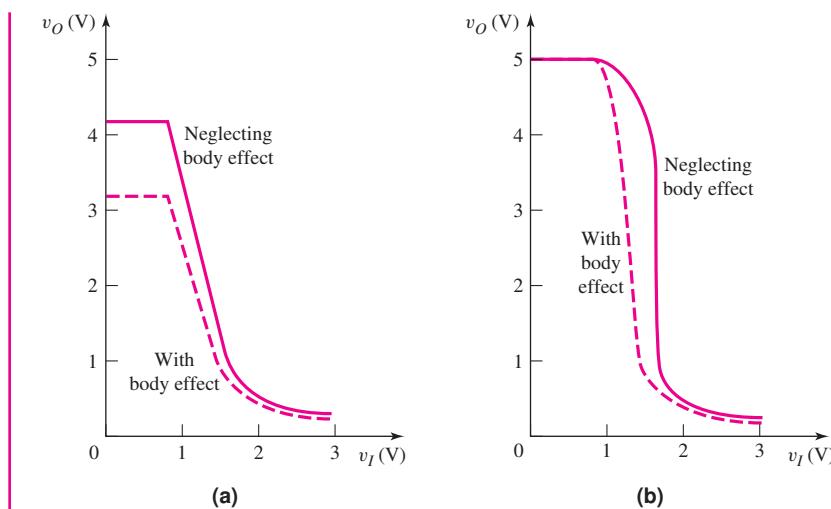


Figure 16.10 Voltage transfer characteristics of NMOS inverters with and without the body effect (a) enhancement load and (b) depletion load

for the saturated load transistor, and $V_{TNLO} = -2$ V for the depletion load transistor. The body effect coefficient was assumed to be $\gamma = 0.9 \text{ V}^{1/2}$.

The body effect changes the voltage transfer characteristics of both the enhancement load and depletion load inverters. Figure 16.10(a) shows the voltage transfer characteristics for the enhancement load inverter. For $v_I = 0$, the output voltage is 3.15 V when the body effect is taken into account. This compares to 4.2 V when the body effect is neglected.

Figure 16.10(b) shows the voltage transfer characteristics for the depletion load inverter. The output voltage is 5 V in the high state, which is independent of the body effect. However, the characteristics during the transition region are a function of the body effect.

EXERCISE PROBLEM

Ex 16.4: Repeat Example 16.4 for the case when the body effect coefficient is $\gamma = 0.3 \text{ V}^{1/2}$. (Ans. $V_{OH} = 1.781$ V)

Test Your Understanding

TYU 16.1 Consider the NMOS inverter with enhancement load, as shown in Figure 16.5(a), biased at $V_{DD} = 1.8$ V. The threshold voltages are $V_{TND} = V_{TNL} = 0.4$ V. Assume $k'_n = 100 \mu\text{A/V}^2$. Design the width-to-length ratios such that the output voltage is 0.12 V and the maximum inverter power dissipation is 0.50 mW when $v_I = 1.4$ V. Neglect the body effect. (Ans. $(W/L)_L = 3.39$, $(W/L)_D = 24.6$)

TYU 16.2 Consider the depletion load inverter in Figure 16.7(a) biased at $V_{DD} = 1.8$ V. The threshold voltages are $V_{TND} = 0.4$ V and $V_{TNL} = -0.6$ V. Assume $k'_n = 100 \mu\text{A/V}^2$. Design the inverter such that the maximum inverter power dissipation is 0.2 mW and the output voltage is 0.08 V when $v_I = 1.8$ V. Neglect the body effect. (Ans. $(W/L)_L = 6.17$, $(W/L)_D = 10.2$)

TYU 16.3 (a) Consider the results of Exercise Ex 16.1. Assume 100,000 resistor-load inverters are fabricated on a single chip and the input voltage of each inverter is high. Determine the current that must be supplied to each chip and the maximum power that will be dissipated on each chip. (b) Repeat part (a) for Exercise Ex 16.2 and the enhancement-load inverter. (c) Repeat part (a) for Exercise Ex 16.3 and the depletion-load inverter. (Ans. (a) $I = 9.8 \text{ A}$, $P = 29.4 \text{ W}$; (b) $I = 58.9 \text{ A}$, $P = 176.6 \text{ W}$; (c) $I = 6.4 \text{ A}$, $P = 19.2 \text{ W}$)

16.2 NMOS LOGIC CIRCUITS

Objective: • Analyze and design NMOS logic gates.

NMOS logic circuits are formed by combining driver transistors in parallel, series, or series-parallel combinations to produce a desired output logic function.

16.2.1 NMOS NOR and NAND Gates

The NMOS NOR logic gate contains additional driver transistors connected in parallel. Figure 16.11 shows a two-input NMOS NOR logic gate with a depletion load. If $A = B = \text{logic 0}$, then both M_{DA} and M_{DB} are cut off and $v_O = V_{DD}$. If $A = \text{logic 1}$ and $B = \text{logic 0}$, then M_{DB} is cut off and the NMOS inverter configuration with M_L and M_{DA} is the same as previously considered, and the output voltage goes low. Similarly, if $A = \text{logic 0}$ and $B = \text{logic 1}$, we again have the same inverter configuration.

If $A = B = \text{logic 1}$, then both M_{DA} and M_{DB} turn on and the two driver transistors are effectively in parallel. The value of the output voltage now changes slightly. Figure 16.12 shows the NOR gate when both input voltages are a logic 1. From our previous analysis, we can assume that the two driver transistors are biased

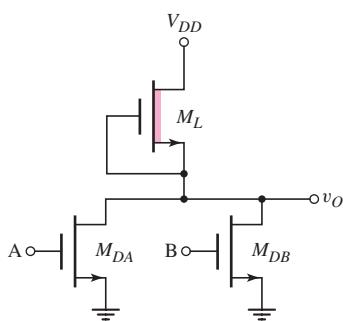


Figure 16.11 Two-input NMOS NOR logic gate with depletion load

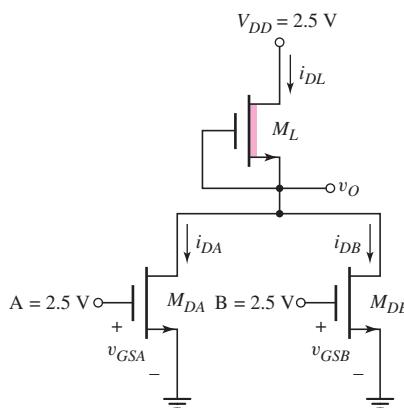


Figure 16.12 Two-input NMOS NOR logic gate for Example 16.5

in the nonsaturation region and the load device is biased in the saturation region. We then have

$$i_{DL} = i_{DA} + i_{DB}$$

which in general terms can be written

$$\begin{aligned} K_L[v_{GSL} - V_{TNL}]^2 &= K_{DA}[2(v_{GSA} - V_{TNA})v_{DSA} - v_{DSA}^2] \\ &\quad + K_{DB}[2(v_{GSB} - V_{TNB})v_{DSB} - v_{DSB}^2] \end{aligned} \quad (16.28)$$

If we assume the two driver transistors are identical, then the driver conduction parameters and threshold voltages are also identical, or $K_{DA} = K_{DB} \equiv K_D$ and $V_{TNA} = V_{TNB} \equiv V_{TND}$. Noting that $v_{GSL} = 0$, $v_{GSA} = v_{GSB} = V_{DD}$, and $v_{DSA} = v_{DSB} = v_O$, we can write Equation (16.28) as

$$[-V_{TNL}]^2 = 2\left(\frac{K_D}{K_L}\right)[2(V_{DD} - V_{TND})v_O - v_O^2] \quad (16.29)$$

Equation (16.29) shows that when both drivers are conducting, the effective width-to-length ratio of the composite driver transistor doubles. This means that the output voltage becomes slightly smaller when both inputs are high.

EXAMPLE 16.5

Objective: Determine the low output voltage of an NMOS NOR circuit.

Consider the NOR logic circuit in Figure 16.12 biased at $V_{DD} = 2.5$ V. Assume transistor parameters of $k'_n = 100 \mu A/V^2$, $V_{TND} = 0.4$ V, $V_{TNL} = -0.6$ V, $(W/L)_D = 4$, and $(W/L)_L = 1$. Neglect the body effect.

Solution: If, for example, $A = \text{logic 1} = 2.5$ V and $B = \text{logic 0}$, then M_{DA} is biased in the nonsaturation region and M_{DB} is cut off. The output voltage is determined from Equation (16.27(b)), which is

$$\frac{K_D}{K_L}[2(v_I - V_{TND})v_O - v_O^2] = (-V_{TNL})^2$$

or

$$\frac{4}{1}[2(2.5 - 0.4)v_O - v_O^2] = [-(-0.6)]^2$$

The output voltage is found to be $v_O = 21.5$ mV.

If both inputs go high, then $A = B = \text{logic 1} = V_{DD} = 2.5$ V and the output voltage can be found using Equation (16.29), which is

$$(-V_{TNL})^2 = 2\left(\frac{K_D}{K_L}\right)[2(V_{DD} - V_{TND})v_O - v_O^2]$$

or

$$[-(-0.6)]^2 = 2\left(\frac{4}{1}\right)[2(2.5 - 0.4)v_O - v_O^2]$$

The output voltage is found to be $v_O = 10.7$ mV.

Comment: An NMOS NOR gate must be designed to achieve a specified V_{OL} output voltage when only one input is high. This will give the largest logic 0 value. When more than one input is high, the output voltage is smaller than the specified V_{OL} value, since the effective width-to-length ratio of the composite driver transistor increases.

EXERCISE PROBLEM

Ex 16.5: Consider the two-input NMOS NOR logic gate shown in Figure 16.11. Let $V_{DD} = 1.8$ V. Assume transistor parameters of $k'_n = 100 \mu\text{A/V}^2$, $V_{TND} = 0.4$ V, $V_{TNL} = -0.6$ V, $(W/L)_D = 5$, and $(W/L)_L = 1$. Neglect the body effect.
 (a) Determine V_{OL} when: (i) $A = \text{logic 1}$, $B = \text{logic 0}$, and (ii) $A = B = \text{logic 1}$.
 (b) Calculate the power dissipation in the circuit for the input condition given in part (a). (Ans. (a) (i) $v_O = 26$ mV, (ii) $v_O = 12.9$ mV; (b) For (i) and (ii), $P = 32.4 \mu\text{W}$)

The NMOS NAND logic gate contains additional driver transistors connected in series. Figure 16.13 shows a two-input NMOS NAND logic gate with a depletion load. If both $A = B = \text{logic 0}$, or if either A or B is a logic 0, at least one driver is cut off, and the output is high. If both $A = B = \text{logic 1}$, then the composite driver of the NMOS inverter conducts and the output goes low.

Since the gate-to-source voltages of M_{DA} and M_{DB} are not equal, determining the actual voltage V_{OL} of a NAND gate is difficult. The drain-to-source voltages of M_{DA} and M_{DB} must adjust themselves to produce the same current. In addition, if the body effect is also included, the analysis becomes even more difficult. Since the two driver transistors are in series, a good approximation assumes that the width-to-length ratio of the drivers must be twice that of a single driver in an NMOS inverter to achieve a given V_{OL} value.

The composite width-to-length ratios of the driver transistors in the two-input NMOS NOR and NAND gates are shown schematically in Figure 16.14. For the NOR gate, the effective *width* doubles; for the NAND gates, the effective *length* doubles.

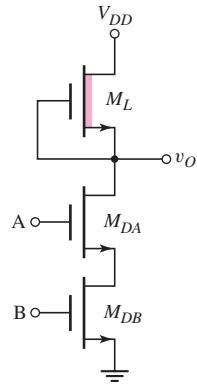


Figure 16.13 Two-input NMOS NAND logic gate with depletion load

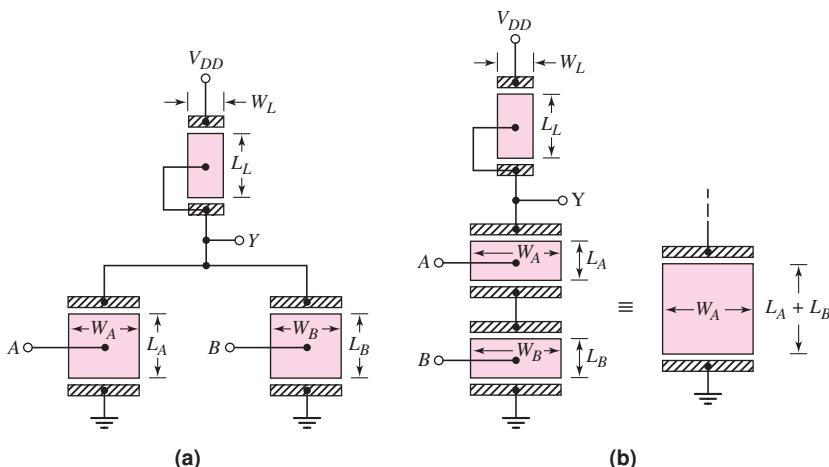


Figure 16.14 Composite width-to-length ratios of driver transistors in two-input NMOS logic configurations (a) NOR and (b) NAND

EXAMPLE 16.6

Objective: Determine the low output voltage of an NMOS NAND circuit.

Consider the NAND logic circuit shown in Figure 16.13 biased at $V_{DD} = 2.5$ V. Assume transistor parameters of $k'_n = 100 \mu\text{A/V}^2$, $V_{TND} = 0.4$ V, $V_{TNL} = -0.6$ V, $(W/L)_D = 8$, and $(W/L)_L = 1$. Neglect the body effect.

Solution: If either A or B is a logic 0, then $v_O = \text{logic 1} = 2.5$ V.

If $A = B = \text{logic 1} = 2.5$ V, then both driver transistors are driven in the non-saturation region and the output goes low. As a good approximation, we will assume the effective length of the driver transistor doubles. Then, using Equation (16.27(b)), we have

$$\frac{\frac{1}{2} \cdot \left(\frac{W}{L}\right)_D}{\left(\frac{W}{L}\right)_L} [2(v_I - V_{TND})v_O - v_O^2] = (-V_{TNL})^2$$

or

$$\frac{8}{(2)(1)} [2(2.5 - 0.4)v_O - v_O^2] = [-(-0.6)]^2$$

The output voltage is found to be $v_O = 21.5$ mV.

This output voltage is the same value that would be obtained for a simple inverter with $(W/L)_D = 4$ and $(W/L)_L = 1$.

Comment: If an N -input NMOS NAND logic gate were to be fabricated then the width-to-length ratio of the drivers would need to be N times that of a single driver in an NMOS inverter to achieve a given value of V_{OL} . The increase in the required area of the driver transistors in a NAND logic gate means that logic gates with more than three or four inputs are not attractive.

EXERCISE PROBLEM

Ex 16.6: Repeat Example 16.6 for a three-input NMOS NAND logic gate with depletion load with $(W/L)_L = 1$ and (a) $(W/L)_D = 12$ and (b) $(W/L)_D = 4$. (Ans. (a) $v_O = 21.5$ mV, (b) $v_O = 65.3$ mV)

16.2.2 NMOS Logic Circuits

The series-parallel combination of drivers can be expanded to synthesize more complex logic functions. Consider the circuit in Figure 16.15. We can show that the Boolean output function is

$$f = \overline{(A \cdot B + C)}$$

Also, the individual transistor width-to-length ratios shown produce an effective K_D/K_L ratio of 4 for an effective single inverter when only M_{DA} and M_{DB} are conducting, or only M_{DC} is conducting. The actual complexity of the Boolean function is limited since the required width-to-length ratios of individual transistors may become unreasonably large.

Two additional logic functions are the exclusive-OR and exclusive-NOR. Figure 16.16 shows a circuit configuration that produces the exclusive-OR function.

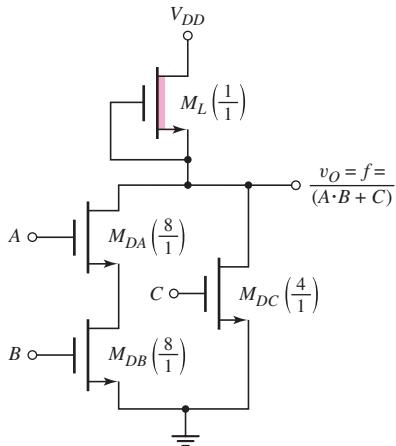


Figure 16.15 NMOS logic circuit example

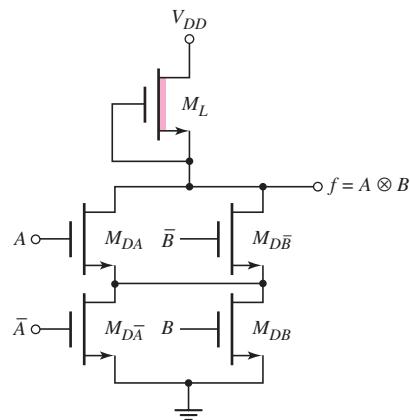
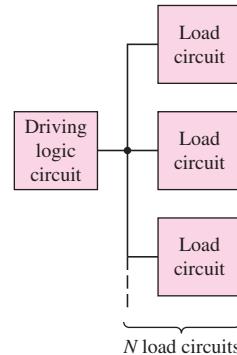


Figure 16.16 NMOS exclusive-OR logic gate

If $A = B = \text{logic 1}$, a path exists from the output to ground through drivers M_{DA} and M_{DB} , and the output goes low. Similarly, if $A = B = \text{logic 0}$, which means that $\bar{A} = \bar{B} = \text{logic 1}$, a path exists from the output to ground through the drivers $M_{D\bar{B}}$ and $M_{D\bar{A}}$, and the output again goes low. For all other input logic signal combinations, the output is isolated from ground so the output goes high.

16.2.3 Fanout

An NMOS inverter or NMOS logic gate must be capable of driving more than one load, as shown in Figure 16.17. It is assumed that each load is identical to the driver logic circuit. The number of identical-load circuits connected to the output of a driver logic circuit is defined as the **fanout**. For MOS logic circuits, the inputs to the load circuits are the oxide-insulated gates of the MOS transistors; therefore, the static loading caused by multiple driver loads is so small that the dc transfer curve is essentially identical to a no-load condition. The dc characteristics of MOS logic circuits are unaffected by the fanout to other MOS logic inputs. However, the load capacitance due to a large fanout seriously degrades the switching speed and propagation delay times. Consequently, maintaining the propagation delay time below a specified maximum value determines the fanout of MOS digital circuits.

Figure 16.17 Logic circuit driving N load circuits

Test Your Understanding

TYU 16.4 (a) Design a three-input NMOS NOR logic gate with depletion load such that $V_{OL(\max)} = 50 \text{ mV}$ and such that the maximum power dissipation is $50 \mu\text{W}$. Let $V_{DD} = 2.5 \text{ V}$. The transistor parameters are $k'_n = 100 \mu\text{A}/\text{V}^2$, $V_{TND} = 0.4 \text{ V}$, and $V_{TNL} = -0.6 \text{ V}$. (b) Using the results of part (a), determine V_{OL} when all inputs are a logic 1. (Ans. (a) $(W/L)_L = 1.11$, $(W/L)_D = 1.93$; (b) $V_{OL} = 16.5 \text{ mV}$)

TYU 16.5 Consider the NMOS logic circuit in Figure 16.18. Assume transistor parameters of $k'_n = 100 \mu\text{A}/\text{V}^2$ and $V_{TN} = 0.4 \text{ V}$. Assume all driver transistors are identical. Neglect the body effect. (a) If $(W/L)_L = 0.5$, determine (W/L) for the drivers such that $V_{OL(\max)} = 80 \mu\text{V}$. Assume logic 1 input voltages are 2.1 V .

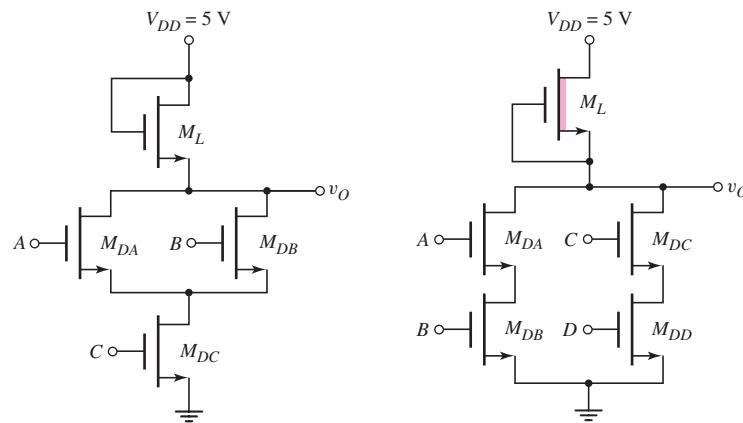


Figure 16.18 Figure for Exercise
TYU 16.5

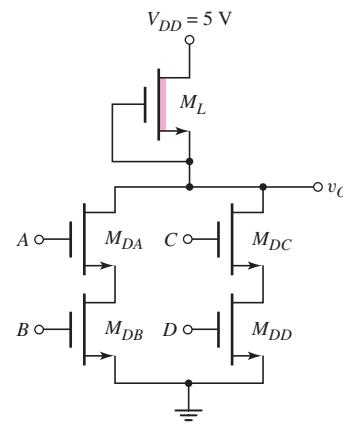


Figure 16.19 Figure for Exercise
TYU 16.6

(b) Determine the maximum power dissipation in the logic circuit. (Ans. (a) $(W/L)_D = 15.4$, (b) $P = 255 \mu\text{W}$)

TYU 16.6 Repeat Exercise TYU 16.5 for the NMOS logic circuit in Figure 16.19, except assume the threshold voltage of the load device is $V_{TNL} = -0.6 \text{ V}$. (Ans. (a) $(W/L)_D = 1.09$, (b) $P = 22.5 \mu\text{W}$)

16.3 CMOS INVERTER

Objective: • Analyze and design CMOS inverters.

Complementary MOS, or CMOS, circuits contain both n-channel and p-channel MOSFETs. As we will see, the power dissipation in CMOS logic circuits is much smaller than in NMOS circuits, which makes CMOS very attractive. We briefly review the characteristics of p-channel transistors, and will then analyze the CMOS inverter, which is the basis of most CMOS logic circuits. We will examine the CMOS NOR and NAND gates and other basic CMOS logic circuits, covering power dissipation, noise margin, fanout, and switching characteristics.

16.3.1 p-Channel MOSFET Revisited

Figure 16.20 shows a simplified view of a p-channel MOSFET. The p- and n-regions are reversed from those in an n-channel device. Again, the channel length is L and the channel width is W . Usually in any given fabrication process, the channel length is the same for all devices, so the channel width W is the variable in logic circuit design.

Normally, in an integrated circuit, more than one p-channel device will be fabricated in the same n-substrate so the p-channel transistors will exhibit a body effect. The n-substrate is connected to the most positive potential. The source terminal may

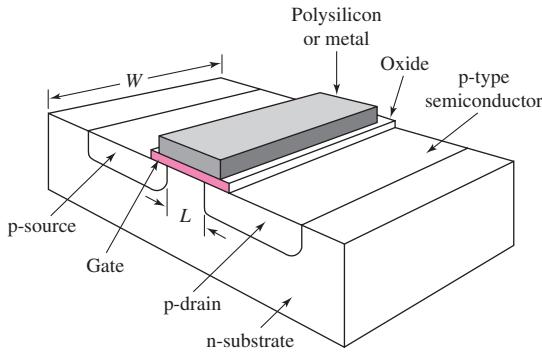


Figure 16.20 Simplified cross section of p-channel MOSFET

be negative with respect to the substrate; therefore, voltage V_{BS} may exist between the body and the source. The threshold voltage is

$$\begin{aligned} V_{TP} &= V_{TPO} - \frac{\sqrt{2e\varepsilon_s N_d}}{C_{ox}} [\sqrt{2\phi_{fn} + V_{BS}} - \sqrt{2\phi_{fn}}] \\ &= V_{TPO} - \gamma [\sqrt{2\phi_{fn} + V_{BS}} - \sqrt{2\phi_{fn}}] \end{aligned} \quad (16.30)$$

where V_{TPO} is the threshold voltage for zero body-to-source voltage, or $V_{BS} = 0$. The parameter N_d is the n-substrate doping concentration and ϕ_{fn} is a potential related to the substrate doping. The parameter γ is the body effect coefficient.

Current-Voltage Relation

The current-voltage characteristics of the p-channel MOSFET are functions of both the electrical and geometric properties of the device. When the transistor is biased in the nonsaturation region, we have $v_{SD} \leq v_{SG} + V_{TP}$. Therefore,

$$i_D = K_p [2(v_{SG} + V_{TP})v_{SD} - v_{SD}^2] \quad (16.31(a))$$

In the saturation region, we have $v_{SD} \geq v_{SG} + V_{TP}$, which means that

$$i_D = K_p (v_{SG} + V_{TP})^2 \quad (16.31(b))$$

The gate potential is negative with respect to the source. For the p-channel transistor to conduct, we must have $v_{GS} < V_{TP}$, where V_{TP} is negative for an enhancement-mode device. We also see that $v_{SG} > |V_{TP}|$ when the p-channel device is conducting.

In most cases, the channel length modulation factor λ has very little effect on the operating characteristics of MOS digital circuits. Therefore, the term λ is assumed to be zero unless otherwise stated.

The transition point, which separates the nonsaturation and saturation bias regions, is given by

$$v_{SD} = v_{SD}(\text{sat}) = v_{SG} + V_{TP} \quad (16.32)$$

The parameter K_p is the conduction parameter and is given by

$$K_p = \left(\frac{1}{2} \mu_p C_{ox} \right) \left(\frac{W}{L} \right) = \frac{k'_p}{2} \frac{W}{L} \quad (16.33)$$

As before, the hole mobility μ_p and oxide capacitance C_{ox} are assumed to be constant for all devices. The hole mobility in p-channel silicon MOSFETs is approximately one-half the electron mobility μ_n in n-channel silicon MOSFETs. This means that a p-channel device width must be approximately twice as large as that of an n-channel device in order that the two devices be electrically equivalent (that is, that they have the same conduction parameter values).

Small Geometry Effects

The same small geometry effects apply to the p-channel devices as we discussed for the n-channel devices in Section 16.1.1. As with the NMOS inverters and logic circuits, we can use Equations (16.31(a)), (16.31(b)), and (16.32) as first-order equations in the design of CMOS logic circuits. The basic operation and behavior of CMOS logic circuits can be predicted using these first-order equations.

16.3.2 DC Analysis of the CMOS Inverter

The **CMOS inverter**, shown in Figure 16.21, is a series combination of a p-channel and an n-channel MOSFET. The gates of the two MOSFETs are connected together to form the input and the **two drains are connected** together to form the output. Both transistors are enhancement-mode devices. The parameters of the NMOS are denoted by K_n and V_{TN} , where $V_{TN} > 0$, and the parameters of the PMOS are denoted by K_p and V_{TP} , where $V_{TP} < 0$.

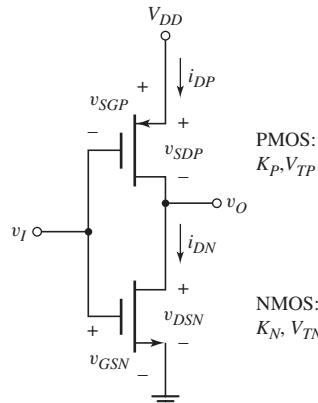


Figure 16.21 CMOS inverter

Figure 16.22 shows a simplified cross section of a CMOS inverter. In this process, a separate p-well region is formed within the starting n-substrate. The n-channel device is fabricated in the p-well region and the p-channel device is fabricated in the n-substrate. Although other approaches, such as an n-well in a p-substrate, are also used to fabricate CMOS circuits, the important point is that the processing is more complicated for CMOS circuits than for NMOS circuits.

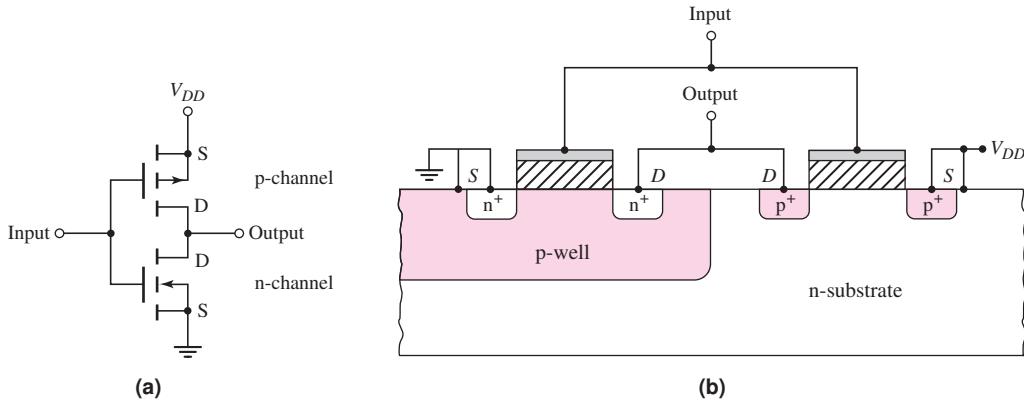


Figure 16.22 Simplified cross section, CMOS inverter

However, the advantages of CMOS digital logic circuits over NMOS circuits justify their use.

Voltage Transfer Curve

Figure 16.23 shows the transistor characteristics for both the n- and p-channel devices. We can determine the voltage transfer characteristics of the inverter by evaluating the various transistor bias regions. For $v_I = 0$, the NMOS device is cut off, $i_{DN} = 0$, and $i_{DP} = 0$. The PMOS source-to-gate voltage is V_{DD} , which means that the PMOS is biased on the curve marked B in Figure 16.23(b). Since the only point on the curve corresponding to $i_{DP} = 0$ occurs at $v_{SDP} = 0 = V_{DD} - v_O$, the output voltage is $v_O = V_{DD}$. This condition exists as long as the **NMOS transistor is cut off**, or $v_I \leq V_{TN}$.

For $v_I = V_{DD}$, the PMOS device is cut off, $i_{DP} = 0$, and $i_{DN} = 0$. The NMOS gate-to-source voltage is V_{DD} and the NMOS is biased on the curve marked A in Figure 16.23(a). The only point on the curve corresponding to $i_{DN} = 0$ occurs at $v_{DSN} = v_O = 0$. The output voltage is zero as long as the PMOS transistor is cut off, or $v_{SGP} = V_{DD} - v_I \leq |V_{TP}|$. This means that the **(input voltage)** is in the **range** $V_{DD} - |V_{TP}| \leq v_I \leq V_{DD}$.

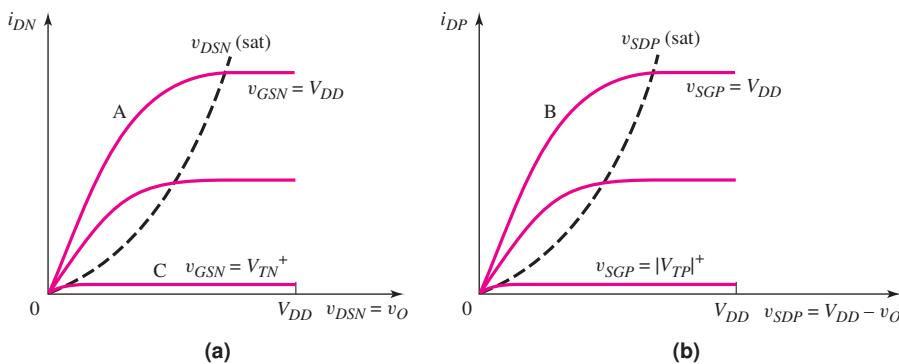


Figure 16.23 Current–voltage characteristics, (a) NMOS transistor and (b) PMOS transistor

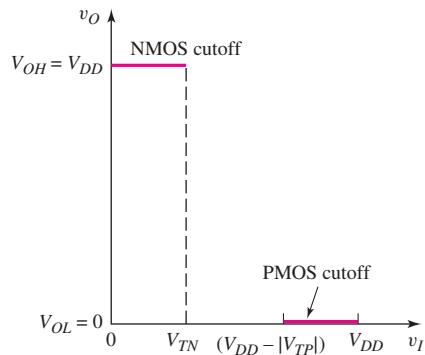


Figure 16.24 CMOS inverter output voltage for input voltage in either high state or low state

Figure 16.24 shows the voltage transfer characteristics generated thus far for the CMOS inverter. The more positive output voltage corresponds to a logic 1, or $V_{OH} = V_{DD}$, and the more negative output voltage corresponds to a logic 0, or $V_{OL} = 0$. When the output is in the logic 1 state, the NMOS transistor is cut off; when the output is in the logic 0 state, the PMOS transistor is cut off.

Ideally, the current in the CMOS inverter in either steady-state condition is zero, which means that, ideally, the quiescent power dissipation is zero. This result is the attractive feature of CMOS digital circuits. In actuality, CMOS inverter circuits exhibit a small leakage current in both steady-state conditions, due to the reverse-biased pn junctions. However, the power dissipation may be in the nanowatt range rather than in the milliwatt range of NMOS inverters. Without this feature, VLSI would not be possible.

When the input voltage is just greater than V_{TN} , or

$$v_I = v_{GSN} = V_{TN}^+$$

the NMOS begins to conduct and the Q -point falls on the curve marked C in Figure 16.23(a). The current is small and $v_{DSN} \approx V_{DD}$, which means that the NMOS is biased in the saturation region. The PMOS source-to-drain voltage is small, so the PMOS is biased in the nonsaturation region. Setting $i_{DN} = i_{DP}$, we can write

$$K_n[v_{GSN} - V_{TN}]^2 = K_p[2(v_{SGP} + V_{TP})v_{SDP} - v_{SDP}^2] \quad (16.34)$$

Relating the gate-to-source and drain-to-source voltages in each transistor to the inverter input and output voltages, respectively, we can rewrite Equation (16.34) as follows:

$$K_n[v_I - V_{TN}]^2 = K_p[2(V_{DD} - v_I + V_{TP})(V_{DD} - v_O) - (V_{DD} - v_O)^2] \quad (16.35)$$

Equation (16.35) relates the input and output voltages as long as the NMOS is biased in the saturation region and the PMOS is biased in the nonsaturation region.

The transition point for the PMOS is defined from

$$v_{SDP}(\text{sat}) = v_{SGP} + V_{TP} \quad (16.36)$$

Using Figure 16.25, Equation (16.36) can be written

$$V_{DD} - V_{OPt} = V_{DD} - V_{IPt} + V_{TP} \quad (16.37\text{(a)})$$

or

$$V_{OPt} = V_{IPt} - V_{TP} \quad (16.37\text{(b)})$$

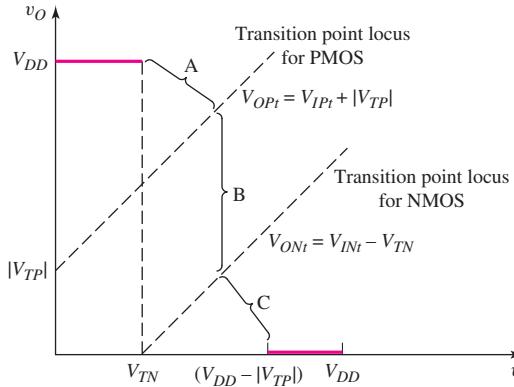


Figure 16.25 Regions of the CMOS transfer characteristics indicating NMOS and PMOS transistor bias conditions. The NMOS device is biased in the saturation region in areas A and B and in the nonsaturation region in area C. The PMOS device is biased in the saturation region in areas B and C and in the nonsaturation region in area A.

where V_{OPt} and V_{IPt} are the PMOS output and input voltages, respectively, at the transition point.

The transition point for the NMOS is defined from

$$v_{DSN}(\text{sat}) = v_{GSN} - V_{TN} \quad (16.38\text{(a)})$$

or

$$V_{ONt} = V_{INt} - V_{TN} \quad (16.38\text{(b)})$$

where V_{ONt} and V_{INt} are the NMOS output and input voltages, respectively, at the transition point.

On the basis that V_{TP} is negative for an enhancement-mode PMOS, Equations (16.37(b)) and (16.38(b)) are plotted in Figure 16.25. We determine the input voltage at the transition points by setting the two drain currents equal to each other when both transistors are biased in the saturation region. The result is

$$K_n(v_{GSN} - V_{TN})^2 = K_p(v_{SGP} + V_{TP})^2 \quad (16.39)$$

With the gate-to-source voltages related to the input voltage, Equation (16.39) becomes

$$K_n(v_I - V_{TN})^2 = K_p(V_{DD} - v_I + V_{TP})^2 \quad (16.40)$$

For this ideal case, the output voltage does not appear in Equation (16.40), and the input voltage is a constant, as long as the two transistors are biased in the saturation region.

Voltage v_I from Equation (16.40) is the input voltage at the PMOS and NMOS transition points. Solving for v_I , we find that

$$v_I = v_{It} = \frac{V_{DD} + V_{TP} + \sqrt{\frac{K_n}{K_p}}V_{TN}}{1 + \sqrt{\frac{K_n}{K_p}}} \quad (16.41)$$

For $v_I > V_{It}$, the NMOS is biased in the nonsaturation region and the PMOS is biased in the saturation region. Again equating the two drain currents, we have

$$K_n[2(v_{GSN} - V_{TN})v_{DSN} - v_{DSN}^2] = K_p(v_{SGP} + V_{TP})^2 \quad (16.42)$$

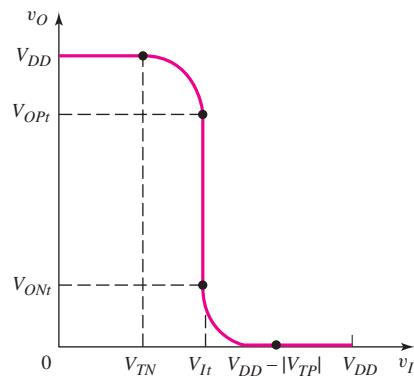


Figure 16.26 Complete voltage transfer characteristics, CMOS inverter

Also, relating the gate-to-source and drain-to-source voltages to the input and output voltages, respectively, modifies Equation (16.42) as follows:

$$K_n[2(v_I - V_{TN})v_O - v_O^2] = K_p(V_{DD} - v_I + V_{TP})^2 \quad (16.43)$$

Equation (16.43) relates the input and output voltages as long as the NMOS is biased in the nonsaturation region and the PMOS in the saturation region. Figure 16.26 shows the complete voltage transfer curve.

EXAMPLE 16.7

Objective: Determine the critical voltages on the voltage transfer curve of a CMOS inverter.

Consider a CMOS inverter biased at $V_{DD} = 5$ V with transistor parameters $K_n = K_p$ and $V_{TN} = -V_{TP} = 0.8$ V. Then consider another CMOS inverter biased at $V_{DD} = 3$ V with transistor parameters $K_n = K_p$ and $V_{TN} = -V_{TP} = 0.6$ V.

Solution ($V_{DD} = 5$ V): The input voltage at the transition points is, from Equation (16.41),

$$V_{It} = \frac{5 + (-0.8) + \sqrt{1}(0.8)}{1 + \sqrt{1}} = 2.5 \text{ V}$$

The output voltage at the transition point for the PMOS is, from Equation (16.37(b)),

$$V_{OPt} = V_{It} - V_{TP} = 2.5 - (-0.8) = 3.2 \text{ V}$$

and the output voltage at the transition point for the NMOS is, from Equation (16.38(b)),

$$V_{ONt} = V_{It} - V_{TN} = 2.5 - 0.8 = 1.7 \text{ V}$$

Solution ($V_{DD} = 3$ V): The critical voltages are

$$V_{It} = 1.5 \text{ V} \quad V_{OPt} = 2.1 \text{ V} \quad V_{ONt} = 0.9 \text{ V}$$

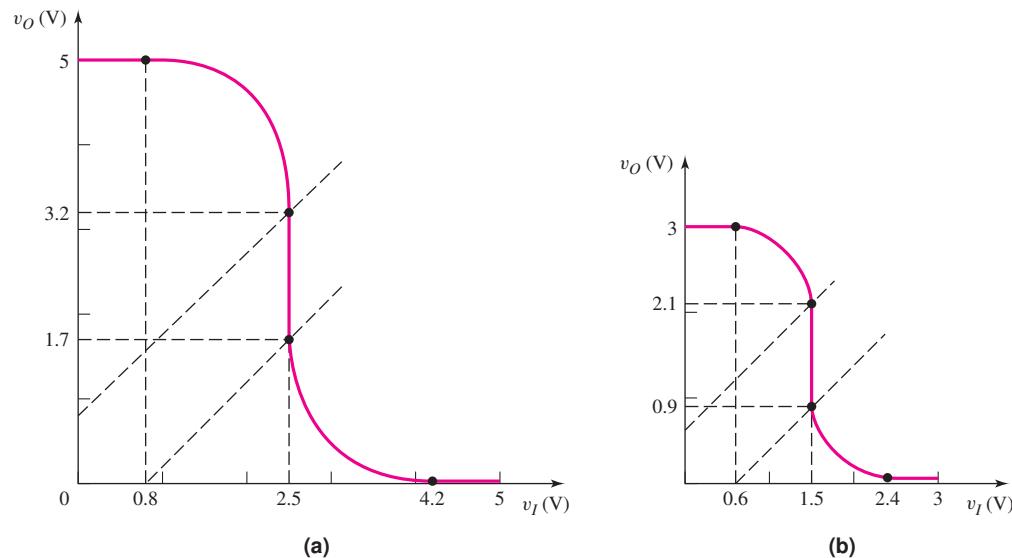


Figure 16.27 Voltage transfer characteristics of CMOS inverter in Example 16.7 biased at (a) $V_{DD} = 5$ V and (b) $V_{DD} = 3$ V

Comment: The two voltage transfer curves are shown in Figure 16.27. These figures depict another advantage of CMOS technology, that is CMOS circuits can be biased over a relatively wide range of voltages.

EXERCISE PROBLEM

Ex 16.7: The CMOS inverter in Figure 16.21 is biased at $V_{DD} = 2.1$ V, and the transistor threshold voltages are $V_{TN} = -V_{TP} = 0.4$ V. Sketch the voltage transfer curve and show the critical voltages as in Figure 16.26 for (a) $K_n/K_p = 1$, (b) $K_n/K_p = 0.5$, and (c) $K_n/K_p = 2$. (Ans. (a) $V_{It} = 1.05$ V, $V_{OPt} = 1.45$ V, $V_{ONT} = 0.65$ V; (b) $V_{It} = 1.16$ V, $V_{OPt} = 1.56$ V, $V_{ONT} = 0.76$ V; (c) $V_{It} = 0.938$ V, $V_{OPt} = 1.338$ V, $V_{ONT} = 0.538$ V)

Transistor Sizing

We may note that both voltage transfer curves shown in Figure 16.27 are symmetrical about the switching point $V_{DD}/2$. This effect is a direct consequence of the fact that the NMOS and PMOS transistors are matched: that is, $K_n = K_p$ and $V_{TN} = |V_{TP}|$. In general, the process conduction parameters, k'_n and k'_p , are not equal. Therefore, in order for the two transistors to be matched, we must adjust the width-to-length ratios. In order for $K_n = K_p$, we have $k'_n(W/L)_n = k'_p(W/L)_p$. In general, $k'_p < k'_n$, so we must have $(W/L)_p > (WL)_n$. The PMOS device must be larger than the NMOS device to make the two devices electrically equivalent.

CMOS Inverter Currents

When the CMOS inverter input voltage is either a logic 0 or a logic 1, the current in the circuit is zero, since one of the transistors is cut off. When the input voltage is in

the range $V_{TN} < v_I < V_{DD} - |V_{TP}|$, both transistors are conducting and a current exists in the inverter.

When the NMOS transistor is biased in the saturation region, the current in the inverter is controlled by v_{GSN} and the PMOS source-to-drain voltage adjusts such that $i_{DP} = i_{DN}$. This condition is demonstrated in Equation (16.34). We can write

$$i_{DN} = i_{DP} = K_n(v_{GSN} - V_{TN})^2 = K_n(v_I - V_{TN})^2 \quad (16.44(a))$$

Taking the square root yields

$$\sqrt{i_{DN}} = \sqrt{i_{DP}} = \sqrt{K_n}(v_I - V_{TN}) \quad (16.44(b))$$

As long as the NMOS transistor is biased in the saturation region, the square root of the CMOS inverter current is a linear function of the input voltage.

When the PMOS transistor is biased in the saturation region, the current in the inverter is controlled by v_{SGP} and the NMOS drain-to-source voltage adjusts such that $i_{DP} = i_{DN}$. This condition is demonstrated in Equation (16.42). Using Equation (16.43), we can write that

$$i_{DN} = i_{DP} = K_p(V_{DD} - v_I + V_{TP})^2 \quad (16.45(a))$$

Taking the square root yields

$$\sqrt{i_{DN}} = \sqrt{i_{DP}} = \sqrt{K_p}(V_{DD} - v_I + V_{TP}) \quad (16.45(b))$$

As long as the PMOS transistor is biased in the saturation region, the square root of the CMOS inverter current is also a linear function of the input voltage.

Figure 16.28 shows plots of the square root of the inverter current for two values of V_{DD} bias. These curves are quasi-static characteristics in that no current is diverted into a capacitive load. At the inverter switching point, both transistors are biased in the saturation region and both transistors influence the current. At the switching point, the actual current characteristic does not have a sharp discontinuity in the slope. The channel length modulation parameter λ also influences the current characteristics at the peak value. However, the curves in Figure 16.28 are excellent approximations.

16.3.3 Power Dissipation

In the quiescent or static state, in which the input is either a logic 0 or a logic 1, power dissipation in the CMOS inverter is virtually zero. However, during the switching cycle from one state to another, current flows and power is dissipated. The CMOS

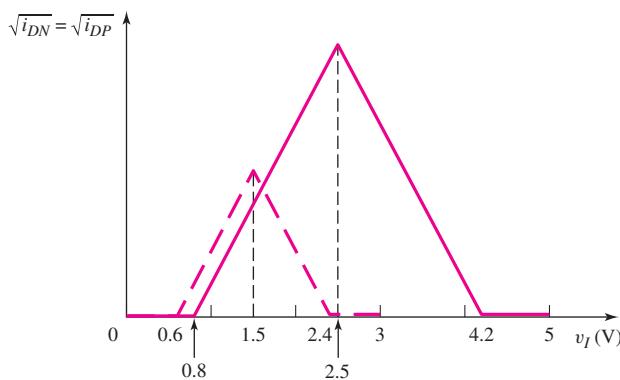


Figure 16.28 Square root of CMOS inverter current versus input voltage for CMOS inverters described in Example 16.7

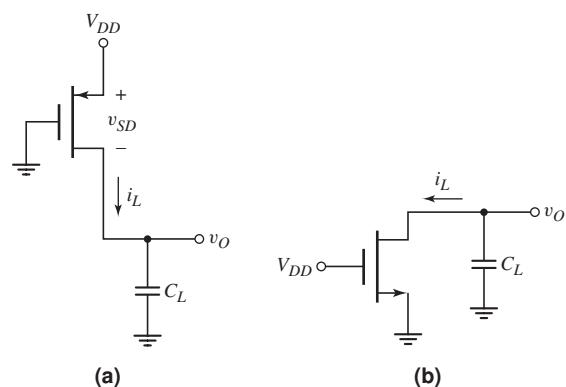


Figure 16.29 CMOS inverter when the output switches (a) low to high and (b) high to low

inverter and logic circuits are used to drive other MOS devices for which the input impedance is a capacitance. During the switching cycle, then, this load capacitance must be charged and discharged.

In Figure 16.29(a), the output switches from its low to its high state. The input is switched low, the PMOS gate is at zero volts, and the NMOS is cut off. The load capacitance C_L must be charged through the PMOS device. Power dissipation in the PMOS transistor is given by

$$P_P = i_L v_{SD} = i_L(V_{DD} - v_O) \quad (16.46)$$

The current and the output voltage are related by

$$i_L = C_L \frac{dv_O}{dt} \quad (16.47)$$

The energy dissipated in the PMOS device as the output switches from low to high is

$$\begin{aligned} E_P &= \int_0^\infty P_P dt = \int_0^\infty C_L(V_{DD} - v_O) \frac{dv_O}{dt} dt \\ &= C_L V_{DD} \int_0^{V_{DD}} dv_O - C_L \int_0^{V_{DD}} v_O dv_O \end{aligned} \quad (16.48)$$

which yields

$$E_P = C_L V_{DD} v_O \Big|_0^{V_{DD}} - C_L \frac{v_O^2}{2} \Big|_0^{V_{DD}} = \frac{1}{2} C_L V_{DD}^2 \quad (16.49)$$

After the output has switched high, the energy stored in the load capacitance is $(\frac{1}{2})C_L V_{DD}^2$. When the inverter input goes high, the output switches low, as shown in Figure 16.29(b). The PMOS device is cut off, the NMOS transistor conducts, and the load capacitance discharges through the NMOS device. All the energy stored in the load capacitance is dissipated in the NMOS device. As the output switches from high to low, the energy dissipated in the NMOS transistor is

$$E_N = \frac{1}{2} C_L V_{DD}^2 \quad (16.50)$$

The total energy dissipated in the inverter during one switching cycle is therefore

$$E_T = E_P + E_N = \frac{1}{2} C_L V_{DD}^2 + \frac{1}{2} C_L V_{DD}^2 = C_L V_{DD}^2 \quad (16.51)$$

If the inverter is switched at frequency f , the power dissipated in the inverter is

$$P = f E_T = f C_L V_{DD}^2 \quad (16.52)$$

Equation (16.52) shows that the power dissipated in a CMOS inverter is directly proportional to the switching frequency and to V_{DD}^2 . The drive in digital IC design is toward lower supply voltages, such as 3 V or less.

The power dissipation is proportional to V_{DD}^2 . In some digital circuits, such as digital watches, the CMOS logic circuits are biased at $V_{DD} = 1.5$ V, so the power dissipation is substantially reduced.

EXAMPLE 16.8

Objective: Calculate the power dissipation in a CMOS inverter.

Consider a CMOS inverter with a load capacitance of $C_L = 2$ pF biased at $V_{DD} = 5$ V. The inverter switches at a frequency of $f = 100$ kHz.

Solution: From Equation (16.52), power dissipation in the CMOS inverter is

$$P = f C_L V_{DD}^2 = (10^5)(2 \times 10^{-12})(5)^2 \Rightarrow 5 \mu\text{W}$$

Comment: Previously determined values of static power dissipation in NMOS inverters were on the order of $500 \mu\text{W}$; therefore, power dissipation in a CMOS inverter is substantially smaller. In addition, in most digital systems, only a small fraction of the logic gates change state during each clock cycle; consequently, the power dissipation in a CMOS digital system is substantially less than in an NMOS digital system of similar complexity.

EXERCISE PROBLEM

Ex 16.8: A CMOS inverter is biased at $V_{DD} = 3 \text{ V}$. The inverter drives an effective load capacitance of $C_L = 0.5 \text{ pF}$. Determine the maximum switching frequency such that the power dissipation is limited to $P = 0.10 \mu\text{W}$. (Ans. $f = 22.2 \text{ kHz}$)

16.3.4 Noise Margin

The word “noise” means transient, unwanted variations in voltages or currents. In digital circuits, if the magnitude of the noise at a logic node is too large, logic errors can be introduced into the system. However, if the noise amplitude is less than a specified value, called the **noise margin**, the noise signal will be attenuated as it passes through a logic gate or circuit, while the logic signals will be transmitted without error.

Noise signals are usually generated outside the digital circuit and transferred to logic nodes or interconnect lines through parasitic capacitances or inductances. The coupling process is usually time dependent, leading to dynamic conditions in the circuit. In digital systems, however, the noise margins are usually defined in terms of static voltages.

Noise Margin Definition

For static noise margins, the type of noise usually considered is called series-voltage noise. Figure 16.30 shows two inverters in series in which the output of the second is connected back to the input of the first. Also included are series-voltage noise sources δV_L and δV_H . This type of noise can be developed by inductive coupling. The input voltage levels are indicated by H (high) and L (low). The noise amplitudes δV_L and δV_H can be different, and the polarities may be such as to increase the low output and reduce the high output. The noise margins are defined as the maximum values of δV_L and δV_H at which the inverters will remain in the correct state.

The actual definitions of the noise margins NM_L and NM_H are not unique. In addition other types of noise, other than series-voltage source noise, may be present in the system. Dynamic noise sources also complicate the issue. However, in this text, in order to provide some measure of noise margin in a logic circuit, we will use the unity-gain approach to determine the logic threshold levels V_{IL} and V_{IH} and the corresponding noise margins.

Figure 16.31 shows a general voltage transfer function for an inverter. The expected logic 1 and logic 0 output voltages of the inverter are V_{OH} and V_{OL} , respec-

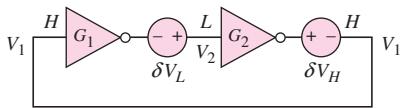


Figure 16.30 Two-inverter flip-flop, including series-voltage noise sources

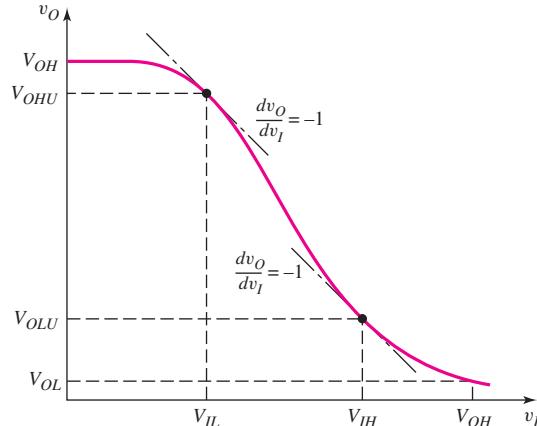


Figure 16.31 Generalized inverter voltage curve and defined voltage limits V_{IL} and V_{IH}

tively. The parameters V_{IH} and V_{IL} , which determine the noise margins, are defined as the points at which

$$\frac{dv_O}{dv_I} = -1 \quad (16.53)$$

For $v_I \leq V_{IL}$, the inverter gain magnitude is less than unity, and the output changes slowly with a change in the input voltage. Similarly, for $v_I \geq V_{IH}$, the output again changes slowly with input voltage since the gain magnitude is less than unity. However, when the input voltage is in the range $V_{IL} < v_I < V_{IH}$, the gain magnitude is greater than one, and the output signal changes rapidly. This region is called the **undefined range**. If the input voltage is inadvertently pushed into this range by a noise signal, the output may change logic states, and a logic error could be introduced into the system. The corresponding output voltages at the unity-gain points are denoted V_{OHU} and V_{OLU} , where the last subscript *U* signifies the unity-gain values.

The noise margins are defined as

$$NM_L = V_{IL} - V_{OLU} \quad (16.54(a))$$

and

$$NM_H = V_{OHU} - V_{IH} \quad (16.54(b))$$

Figure 16.32 shows the general voltage transfer function of a CMOS inverter. (The numbers in the figure are from Example 16.9 to be considered later.) The parameters V_{IH} and V_{IL} determine the noise margins and are defined as the points at which

$$\frac{dv_O}{dv_I} = -1 \quad (16.55)$$

For $v_I \leq V_{IL}$ and $v_I \geq V_{IH}$, the gain is less than unity and the output changes slowly with input voltage. However, when the input voltage is in the range $V_{IL} < v_I < V_{IH}$, the inverter gain is greater than unity, and the output signal changes rapidly with a change in the input voltage. This is the undefined range.

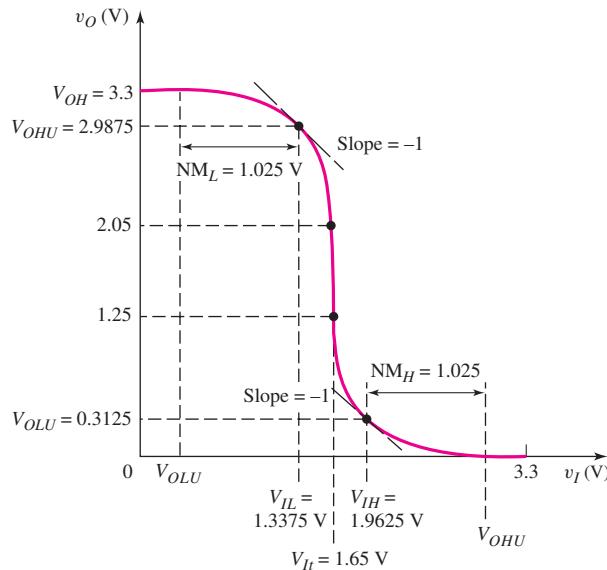


Figure 16.32 CMOS inverter voltage transfer characteristics with defined noise margins

Point V_{IL} occurs when the NMOS is biased in the saturation region and the PMOS is biased in the nonsaturation region. The relationship between the input and output voltages is given by Equation (16.35). Taking the derivative with respect to v_I yields

$$2K_n[v_I - V_{TN}] = K_p \left[-2(V_{DD} - v_O) - 2(V_{DD} - v_I + V_{TP}) \frac{dv_O}{dv_I} - 2(V_{DD} - v_O) \left(-\frac{dv_O}{dv_I} \right) \right] \quad (16.56)$$

Setting the derivative equal -1 , we have

$$K_n[v_I - V_{TN}] = -K_p[(V_{DD} - v_O) - (V_{DD} - v_I + V_{TP}) + (V_{DD} - v_O)] \quad (16.57)$$

Solving for v_O produces

$$v_O = V_{OHU} = \frac{1}{2} \left\{ \left(1 + \frac{K_n}{K_p} \right) v_I + V_{DD} - \left(\frac{K_n}{K_p} \right) V_{TN} - V_{TP} \right\} \quad (16.58)$$

Combining Equations (16.58) and (16.35), we see that voltage V_{IL} is

$$v_I = V_{IL} = V_{TN} + \frac{(V_{DD} + V_{TP} - V_{TN})}{\left(\frac{K_n}{K_p} - 1 \right)} \left[2 \sqrt{\frac{\frac{K_n}{K_p}}{\frac{K_n}{K_p} + 3}} - 1 \right] \quad (16.59)$$

If $K_n = K_p$, Equation (16.59) becomes indefinite, since a zero would exist in both the numerator and the denominator. However, when $K_n = K_p$, Equation (16.58) becomes

$$v_O = V_{OHU(K_n=K_p)} = \frac{1}{2} \{ 2v_I + V_{DD} - V_{TN} - V_{TP} \} \quad (16.60)$$

Substituting Equation (16.60) into Equation (16.35) yields a voltage V_{IL} of

$$v_I = V_{IL(K_n=K_p)} = V_{TN} + \frac{3}{8}(V_{DD} + V_{TP} - V_{TN}) \quad (16.61)$$

for $K_n = K_p$.

Point V_{IH} occurs when the NMOS is biased in the nonsaturation region and the PMOS is biased in the saturation region. The relationship between the input and output voltages is given by Equation (16.43). Taking the derivative with respect to v_I yields

$$K_n \left[2(v_I - V_{TN}) \frac{dv_O}{dv_I} + 2v_O - 2v_O \frac{dv_O}{dv_I} \right] = 2K_p(V_{DD} - v_I + V_{TP})(-1) \quad (16.62)$$

Setting the derivative equal to -1 , we find that

$$K_n[-(v_I - V_{TN}) + v_O + v_O] = -K_p[V_{DD} - v_I + V_{TP}] \quad (16.63)$$

The output voltage v_O is then

$$v_O = V_{OLU} = \frac{v_I \left(1 + \frac{K_n}{K_p} \right) - V_{DD} - \left(\frac{K_n}{K_p} \right) V_{TN} - V_{TP}}{2 \left(\frac{K_n}{K_p} \right)} \quad (16.64)$$

Combining Equations (16.64) and (16.43), yields voltage V_{IH} as

$$v_I = V_{IH} = V_{TN} + \frac{(V_{DD} + V_{TP} - V_{TN})}{\left(\frac{K_n}{K_p} - 1 \right)} \left[\frac{\frac{2K_n}{K_p}}{\sqrt{3\frac{K_n}{K_p} + 1}} - 1 \right] \quad (16.65)$$

Again, if $K_n = K_p$, Equation (16.65) becomes indefinite, since a zero would exist in both the numerator and the denominator. However, when $K_n = K_p$, Equation (16.64) becomes

$$v_O = V_{OLU(K_n=K_p)} = \frac{1}{2}\{2v_I - V_{DD} - V_{TN} - V_{TP}\} \quad (16.66)$$

Substituting Equation (16.66) into Equation (16.43) yields a voltage V_{IH} of

$$v_I = V_{IH(K_n=K_p)} = V_{TN} + \frac{5}{8}(V_{DD} + V_{TP} - V_{TN}) \quad (16.67)$$

EXAMPLE 16.9

Objective: Determine the noise margins of a CMOS inverter.

Consider a CMOS inverter biased at $V_{DD} = 3.3$ V. Assume the transistors are matched with $K_n = K_p$ and $V_{TN} = -V_{TP} = 0.4$ V.

Solution: From Equation (16.41), the input voltage at the transition points, or the inverter switching point, is 1.65 V. Since $K_n = K_p$, V_{IL} is, from Equation (16.61),

$$V_{IL} = V_{TN} + \frac{3}{8}(V_{DD} + V_{TP} - V_{TN}) = 0.4 + \frac{3}{8}(3.3 - 0.4 - 0.4) = 1.3375 \text{ V}$$

Point V_{IH} is, from Equation (16.67),

$$V_{IH} = V_{TN} + \frac{5}{8}(V_{DD} + V_{TP} - V_{TN}) = 0.4 + \frac{5}{8}(3.3 - 0.4 - 0.4) = 1.9625 \text{ V}$$

The output voltages at points V_{IL} and V_{IH} are determined from Equations (16.60) and (16.66), respectively. They are

$$\begin{aligned} V_{OHU} &= \frac{1}{2}[2V_{IL} + V_{DD} - V_{TN} - V_{TP}] \\ &= \frac{1}{2}[2(1.3375) + 3.3 - 0.4 + 0.4] = 2.9875 \text{ V} \end{aligned}$$

and

$$\begin{aligned} V_{OLU} &= \frac{1}{2}[2V_{IH} - V_{DD} - V_{TN} - V_{TP}] \\ &= \frac{1}{2}[2(1.9625) - 3.3 - 0.4 + 0.4] = 0.3125 \text{ V} \end{aligned}$$

The noise margins are therefore

$$NM_L = V_{IL} - V_{OLU} = 1.3375 - 0.3125 = 1.025 \text{ V}$$

and

$$NM_H = V_{OHU} - V_{IH} = 2.9875 - 1.9625 = 1.025 \text{ V}$$

Comment: The results of this example are shown in Figure 16.32. Since the two transistors are electrically identical, the voltage transfer curve and the resulting critical voltages are symmetrical. Also, $(V_{OH} - V_{OHU}) = 0.3125 \text{ V}$, which is less than $|V_{TP}|$, and $(V_{OLU} - V_{OL}) = 0.3125 \text{ V}$, which is less than V_{TN} . As long as the input voltage remains within the limits of the noise margin, no logic error will be transmitted through the digital system.

EXERCISE PROBLEM

Ex 16.9: A CMOS inverter is biased at $V_{DD} = 1.8 \text{ V}$. The transistor parameters are $V_{TN} = 0.4 \text{ V}$, $V_{TP} = -0.4 \text{ V}$, $K_n = 200 \mu\text{A}/\text{V}^2$, and $K_p = 80 \mu\text{A}/\text{V}^2$. (a) Determine the transition points. (b) Find the critical voltages V_{IL} and V_{IH} , and the corresponding output voltages. (c) Calculate the noise margins NM_L and NM_H . (Ans. (a) $V_{It} = 0.7874 \text{ V}$, $V_{OPT} = 1.187 \text{ V}$, $V_{ONt} = 0.3874 \text{ V}$; (b) $V_{IL} = 0.6323 \text{ V}$, $V_{IH} = 0.8767 \text{ V}$, $V_{OHU} = 1.7065 \text{ V}$, $V_{OLU} = 0.1337 \text{ V}$; (c) $NM_L = 0.4986 \text{ V}$, $NM_H = 0.8298 \text{ V}$)

Test Your Understanding

TYU 16.7 Consider a CMOS inverter biased at $V_{DD} = 5 \text{ V}$, with transistor threshold voltages of $V_{TN} = +0.8 \text{ V}$ and $V_{TP} = -0.8 \text{ V}$. Calculate the peak current in the inverter for: (a) $K_n = K_p = 50 \mu\text{A}/\text{V}^2$, and (b) $K_n = K_p = 200 \mu\text{A}/\text{V}^2$. (Ans. (a) $i_D(\max) = 145 \mu\text{A}$ (b) $i_D(\max) = 578 \mu\text{A}$)

TYU 16.8 Repeat Exercise Ex 16.9 for a CMOS inverter biased at $V_{DD} = 5 \text{ V}$ with transistor parameters of $V_{TN} = 0.8 \text{ V}$, $V_{TP} = -2 \text{ V}$, and $K_n = K_p = 100 \mu\text{A}/\text{V}^2$. (Ans. (a) $V_{It} = 1.9 \text{ V}$, $V_{OPT} = 3.9 \text{ V}$, $V_{ONt} = 1.1 \text{ V}$; (b) $V_{IL} = 1.625 \text{ V}$, $V_{IH} = 2.175 \text{ V}$, $V_{OLU} = 0.275 \text{ V}$, $V_{OHU} = 4.725 \text{ V}$; (c) $NM_L = 1.35 \text{ V}$, $NM_H = 2.55 \text{ V}$)



16.4 CMOS LOGIC CIRCUITS

Objective: • Analyze and design static CMOS logic gates.

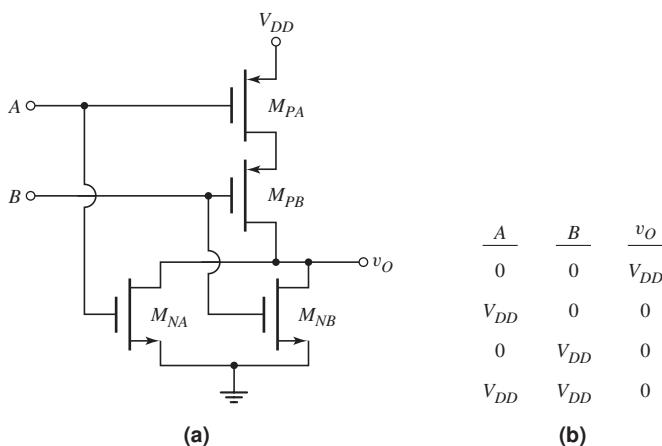
Large-scale integrated CMOS circuits are used extensively in digital systems, including watches, calculators, and microprocessors. We will look at the basic CMOS NOR and NAND gates, and will then analyze more complex CMOS logic circuits. Since there is no clock signal applied to these logic circuits, they are referred to as **static CMOS logic** circuits.

16.4.1 Basic CMOS NOR and NAND Gates

In the basic or classical CMOS logic circuits, the gates of a PMOS and an NMOS are connected together, and additional PMOS and NMOS transistors are connected in series or parallel to form specific logic circuits. Figure 16.33(a) shows a two-input CMOS NOR gate. The NMOS transistors are in parallel and the PMOS transistors are in series.

If $A = B = \text{logic 0}$, then both M_{NA} and M_{NB} are cut off, and the current in the circuit is zero. The source-to-gate voltage of M_{PA} is V_{DD} but the current is zero; therefore, v_{SD} of M_{PA} is zero. This means that the source-to-gate voltage of M_{PB} is also V_{DD} . However, since the current is zero, then v_{SD} of M_{PB} is also zero. The output voltage is therefore $v_O = V_{DD} = \text{logic 1}$.

If the input signals are $A = \text{logic 1} = V_{DD}$ and $B = \text{logic 0} = 0 \text{ V}$, then the source-to-gate voltage of M_{PA} is zero, and the current in the circuit is again zero. The gate-to-source voltage of M_{NA} is V_{DD} but the current is zero, so v_{DS} of M_{NA} is zero and $v_O = 0 = \text{logic 0}$. This result also holds for the other two possible input conditions, since at least one PMOS is cut off and at least one NMOS is in a conducting state. The NOR logic function is shown in the truth table of Figure 16.33(b).



A	B	v_O
0	0	V_{DD}
V_{DD}	0	0
0	V_{DD}	0
V_{DD}	V_{DD}	0

Figure 16.33 (a) Two-input CMOS NOR logic circuit and (b) truth table

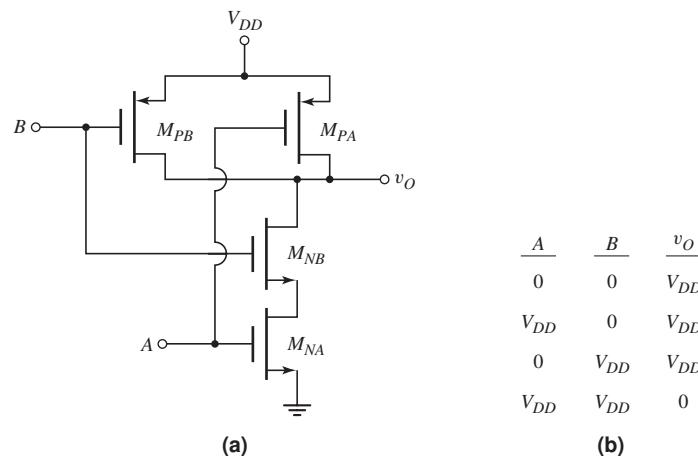


Figure 16.34 (a) Two-input CMOS NAND logic circuit and (b) truth table

A two-input CMOS NAND logic gate is shown in Figure 16.34(a). In this case, the NMOS transistors are in series and the PMOS transistors are in parallel. If $A = B = \text{logic 0}$, the two NMOS devices are cut off and the current in the circuit is zero. The source-to-gate voltage of each PMOS device is V_{DD} , which means that both PMOS transistors are in a conducting state. However, since the current is zero, v_{SD} for both M_{PA} and M_{PB} is zero and $v_O = V_{DD}$. This result applies if at least one input is a logic 0.

If the input signals are $A = B = \text{logic 1} = V_{DD}$, then both PMOS transistors are cut off, and the current in the circuit is zero. With $A = \text{logic 1}$, M_{NA} is in a conducting state; however, since the current is zero, then v_{DS} of M_{NA} is zero. This means that the gate-to-source voltage of M_{NB} is also V_{DD} and M_{NB} is also in a conducting state. However, since the current is zero, then v_{DS} of M_{NB} is zero, and $v_O = \text{logic 0} = 0 \text{ V}$. The NAND logic function is shown in the truth table in Figure 16.34(b).

In both the CMOS NOR and NAND logic gates, the current in the circuit is essentially zero when the inputs are in any quiescent state. Only very small reverse-bias pn junction currents exist. The quiescent power dissipation is therefore essentially zero. Again, this is the primary advantage of CMOS circuits.

16.4.2 Transistor Sizing

CMOS Inverter

We briefly discussed in Section 16.3.2 the sizing of transistors in the CMOS inverter in terms of symmetrical transfer curves. Other factors involved in the sizing of transistors are, for example, switching speed, power, area, and noise margin.

Since the standby power is very small in a CMOS inverter, the sizing can be based on switching speed. We will specify that the switching time in the pull-up mode should be the same as the switching time in the pull-down mode. Figure 16.35(a) shows the effective CMOS inverter in the pull-down mode. The PMOS is cutoff and the load capacitance is discharged through the NMOS device. The switching time is therefore a function of the current capability of the NMOS transistor. Figure 16.35(b) shows the effective CMOS inverter in the pull-up mode. The NMOS is cutoff and the

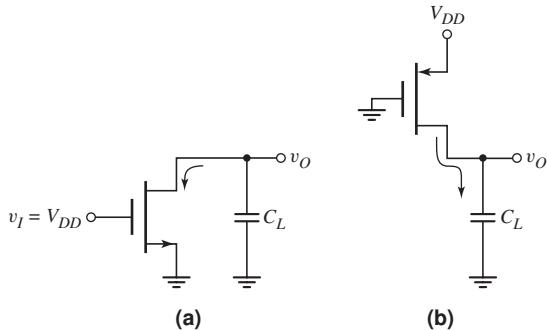


Figure 16.35 (a) Effective CMOS inverter in pull-down mode and (b) effective CMOS inverter in pull-up mode

load capacitance is charged through the PMOS device. The switching time is a function of the current capability of the PMOS transistor.

Assuming that $V_{TN} = |V_{TP}|$, equal switching times then implies that the conduction parameters of the NMOS and PMOS devices be equal, or

$$\frac{k'_n}{2} \left(\frac{W}{L} \right)_n = \frac{k'_p}{2} \left(\frac{W}{L} \right)_p \quad (16.68)$$

Assuming that $\mu_n \approx 2\mu_p$, we have

$$\frac{(W/L)_p}{(W/L)_n} = \frac{k'_n}{k'_p} = \frac{\mu_n}{\mu_p} \approx 2 \quad (16.69)$$

The width-to-length ratio of the PMOS device must be approximately twice as large as that of the NMOS device to obtain equal switching times.

In any given technology, the channel lengths of the NMOS and PMOS devices are the same. Therefore the channel widths are sized to the desired value. We can write that $W_n = W$ and $W_p = 2W$, where W_n and W_p are the channel widths of the NMOS and PMOS devices, respectively, and W is a standard width.

CMOS Logic Gates

We can now consider the sizing of transistors in the basic CMOS NAND and NOR logic gates. We will specify, again, equal pull-up and pull-down switching times, and we want the same switching times as the CMOS inverter with a load capacitance C_L . We will use the effective 2:1 ratio between PMOS and NMOS sizes from the CMOS inverter.

Consider the two-input CMOS NOR gate shown in Figure 16.33. Assume a load capacitance C_L is connected to the output. In the worst case during a pull-down operation, only one NMOS device will be turned on. To achieve the same switching time as the CMOS inverter, the NMOS channel widths should be $W_n = W$. If both NMOS devices are turned on, the effective channel width will be doubled (see Figure 16.14(a)) and the switching time will be shorter.

During a pull-up operation, both PMOS devices must be turned on. Since the PMOS devices are in series, the effective channel length doubles (see Figure 16.14(b)). Therefore, to maintain the same effective width-to-length ratio, the channel widths must be doubled. We must therefore have $W_p = 2(2W) = 4W$.

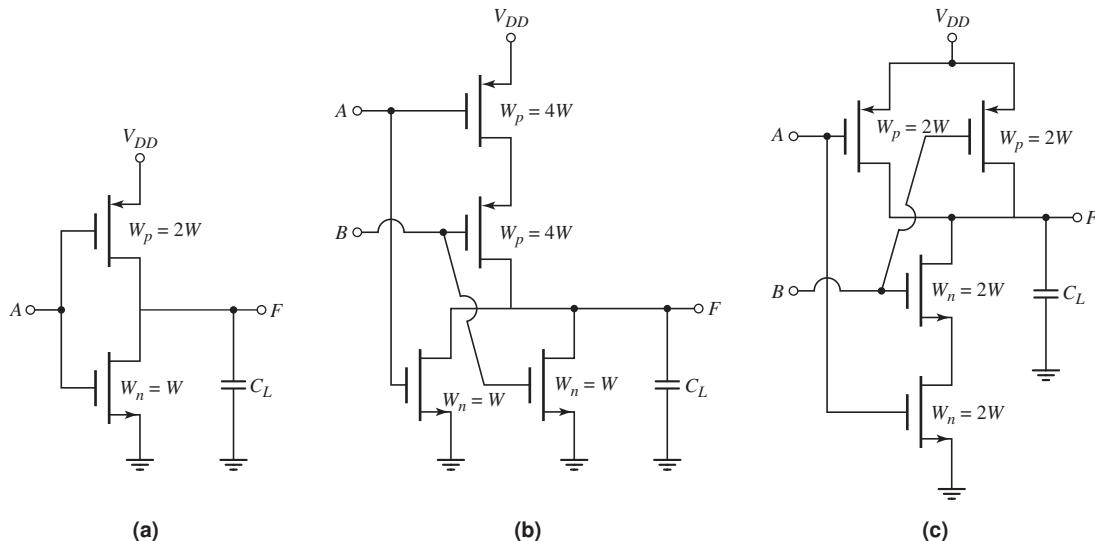


Figure 16.36 The width-to-length ratios of (a) the CMOS inverter, (b) the CMOS NOR gate, and (c) the CMOS NAND gate

Now consider the two-input NAND logic gate shown in Figure 16.34. Again, assume a load capacitance C_L is connected to the output. In the worst case during a pull-up operation, only one PMOS device will be turned on. This is equivalent to the CMOS inverter, so the channel width should be $W_p = 2W$. If both PMOS devices are turned on, the effective channel width is doubled and the switching time will be shorter.

During the pull-down operation, both NMOS devices must be turned on. Again, since the NMOS devices are in series, the effective channel length doubles. Therefore to maintain the same effective width-to-length ratio, the channel widths must be doubled. We must therefore have $W_n = 2(W) = 2W$.

The results of the transistor sizing for the CMOS inverter, and CMOS NOR and NAND gates are shown in Figure 16.36.

EXAMPLE 16.10

Objective: Determine the transistor width-to-length ratios of a three-input CMOS NAND logic gate.

Symmetrical switching times are desired and the switching times should correspond to the basic CMOS inverter.

Solution: There are three p-channel transistors in parallel for the three-input CMOS NAND gate. The worst case is when only one PMOS device is on in the pull-up mode. This corresponds to the basic CMOS inverter, so the effective width should be $W_p = 2W$.

There are three n-channel transistors in series for the three-input CMOS NAND gate. All three transistors must be turned on in the pull-down mode. For three transistors in series, the effective channel length triples. Therefore, to keep the effective NMOS width equal to W , we must have $W_n = 3(W) = 3W$.

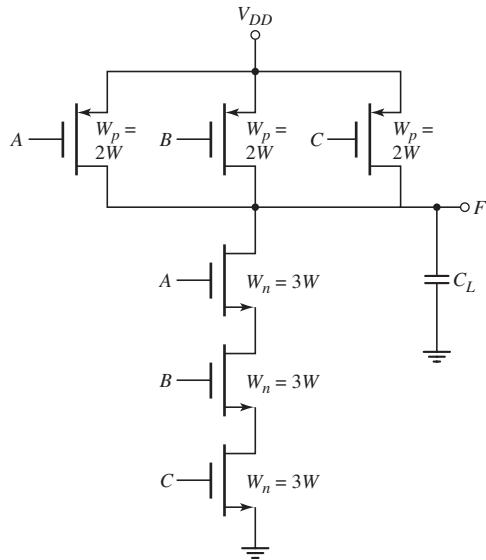


Figure 16.37 Width-to-length ratios for a three-input CMOS NAND logic gate

The results are shown in Figure 16.37.

Comment: As the number of inputs to a basic CMOS logic gate increases, the size of the transistors must increase. The increased area of the transistors means that the effective input capacitance increases so that switching times of cascaded logic gates will increase.

EXERCISE PROBLEM

Ex 16.10: Determine the transistor sizes of a 3-input CMOS NOR logic gate. Symmetrical switching times are desired and the switching times should correspond to the basic CMOS inverter. (Ans. $W_p = 6W$, $W_n = W$)

16.4.3 Complex CMOS Logic Circuits

Just as with NMOS logic designs, we can form complex logic gates in CMOS, which avoids connecting large numbers of NOR, NAND, and inverter gates to implement the logic function. There are formal methods that can be used to implement the logic circuit. However, we can use the knowledge gained in the analysis and design of the NOR and NAND circuits.

DESIGN EXAMPLE 16.11

Objective: Design a CMOS logic circuit to implement a particular logic function. Implement the logic function $Y = AB + C(D + E)$ in a CMOS design. The signals A , B , C , D , and E are available.

Design Approach: The general CMOS design is shown in Figure 16.38, in which the inputs are applied to both the PMOS and NMOS networks. We may start the

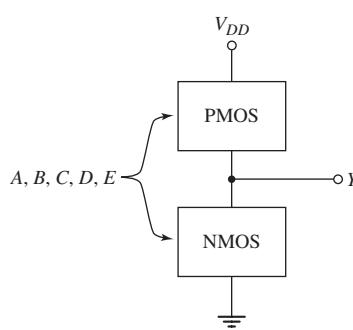


Figure 16.38 General CMOS design

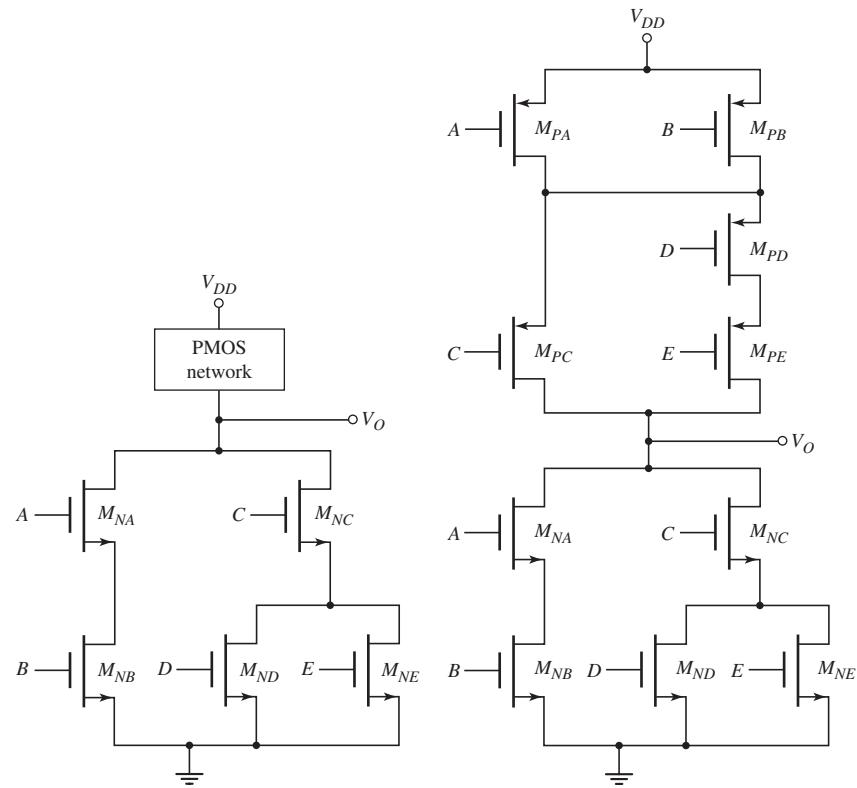


Figure 16.39 NMOS design for Example 16.11

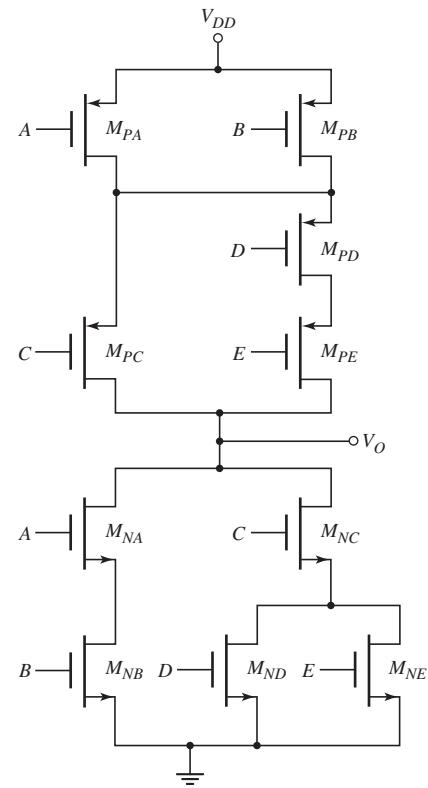


Figure 16.40 Complete CMOS design for Example 16.11

design by considering the NMOS portion of the circuit. To implement a basic OR (NOR) function, the n-channel transistors are in parallel (Figure 16.33) and to implement a basic AND (NAND) function, the n-channel transistors are in series (Figure 16.34). We will consider whether the function or its complement is generated at the end of the design.

Solution (NMOS Design): In the overall function, we note the logic OR between the functions AB and $C(D + E)$, so that the NMOS devices used to implement AB will be in parallel with the NMOS devices used to implement $C(D + E)$. There is a logic AND between the inputs A and B , so that the NMOS devices with these inputs will be in series. Finally, the NMOS devices with the D and E inputs will be in parallel and this combination will be in series with the NMOS device with the C input. The NMOS implementation of the function is shown in Figure 16.39.

Solution (PMOS Design): The arrangement of the PMOS devices is complementary to that of the NMOS devices. PMOS devices that perform the basic OR function are in series and PMOS devices that perform the basic AND function are in parallel. We then see that the PMOS devices used to implement AB will be in series with the devices used to implement $C(D + E)$. The two PMOS devices with the A and B inputs will be in parallel. The two PMOS devices with the D and E inputs will be in series and in turn will be in parallel with the PMOS device with the C input. The completed circuit is shown in Figure 16.40.

Final Solution: By considering various inputs, we may note that the output signal of the circuit shown in Figure 16.40 is actually the complement of the desired signal. We may then simply add a CMOS inverter to the output to obtain the desired function.

Comment: As mentioned, there are formal ways in which to design circuits. However, in many cases, these circuits can be designed by using the knowledge and intuition gained from previous work. The width-to-length ratios of the various transistors can be determined as we have done in previous examples.

EXERCISE PROBLEM

Ex 16.11: Design the width-to-length ratios of the transistors in the static CMOS logic circuit of Figure 16.40. Symmetrical switching times are desired and the switching times should correspond to the basic CMOS inverter. (Ans. All NMOS devices, $W_n = 2W$; $W_p(M_{PA}) = W_p(M_{PB}) = W_p(M_{PC}) = 4W$; $W_p(M_{PD}) = W_p(M_{PE}) = 8W$)

Another example of a CMOS logic gate is the exclusive-OR or XOR. The logic function can be written as

$$F_{XOR} = \bar{A}\bar{B} + A\bar{B} \quad (16.70)$$

We have noticed that the output of the CMOS gates is actually the complement of the input signal. We can therefore write

$$\bar{F}_{XOR} = F_{XNOR} = \bar{A}\bar{B} + AB \quad (16.71)$$

Assuming that input signals A , B , \bar{A} , and \bar{B} are available, Figure 16.41 shows a CMOS static implementation of the logic function.

We may note that $\bar{A}\bar{B}$ as well as AB means two NMOS devices in series and two PMOS devices in parallel. The OR function means the combination of NMOS

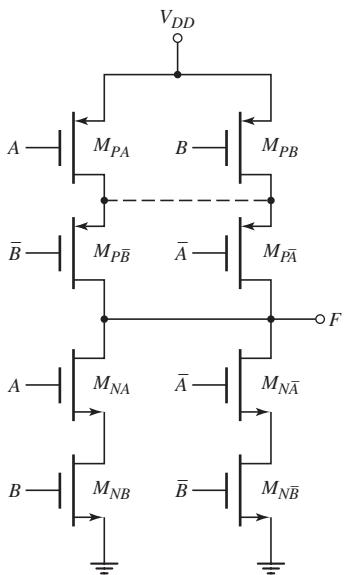


Figure 16.41 A CMOS static exclusive-OR logic gate

devices is in parallel and the combination of PMOS devices is in series. This design is shown in the figure. In considering the truth table for the exclusive-OR function, we may note that the output of the circuit in Figure 16.41 is indeed the exclusive-OR function. In the design of CMOS logic gates, then, we should actually design the complement of the desired function.

In the PMOS portion of the design, there should be an electrical connection between the drains of M_{PA} and M_{PB} . This connection is shown as a dotted line, but is not actually required. The only pull-up conditions are for $A = \bar{B} = 0$ and for $\bar{A} = B = 0$, which are achieved without this connection.

16.4.4 Fanout and Propagation Delay Time

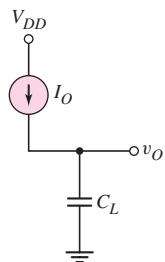


Figure 16.42 Constant-current source charging a load capacitor

Fanout

The term *fanout* refers to the number of load gates of similar design connected to the output of a driver gate. The maximum fanout is the maximum number of load gates that may be connected to the output. Since the CMOS logic gate will be driving other CMOS logic gates, the quiescent current required to drive the other CMOS gates is essentially zero. In terms of static characteristics, the maximum fanout is virtually limitless.

However, each additional load gate increases the load capacitance that must be charged and discharged as the driver gate changes state, and this places a practical limit on the maximum allowable number of load gates. Figure 16.42 shows a constant current charging a load capacitance. The voltage across the capacitance is

$$v_O = \frac{1}{C_L} \int I_O dt = \frac{I_O t}{C_L} \quad (16.72)$$

The load capacitance C_L is proportional to the number N of load gates and to the input gate capacitance of each load. The current I_O is proportional to the conduction parameter of the driver transistor. The switching time is therefore

$$t \propto \frac{N(W \cdot L)_L}{\left(\frac{W}{L}\right)_D} \quad (16.73)$$

where the gate capacitance is directly proportional to the gate area of the load $(W \cdot L)_L$, and the conduction parameter of the driver transistor is proportional to the width-to-length ratio. Equation (16.73) can be rewritten as

$$t \propto N(L_L L_D) \left(\frac{W_L}{W_D} \right) \quad (16.74)$$

The propagation delay time, which is proportional to the switching time, increases as the fanout increases. The propagation delay time could be reduced by increasing the size of the driver transistor. However, in any given driver logic circuit and load logic circuit, the sizes of the devices are generally fixed. Consequently, the maximum fanout is limited by the maximum acceptable propagation delay time.

Propagation delay times are typically measured with a specified load capacitance. The average propagation delay time of a two-input CMOS NOR gate (such as an SN74HC36) is 25 ns, measured with a load capacitance of $C_L = 50$ pF. Since the input capacitance is $C_I = 10$ pF, a fanout of five would produce a 50 pF load capacitance. A fanout larger than five would increase the load capacitance, and would also increase the propagation delay time above the specified value.

Propagation Delay Time

Although the propagation delay time of the CMOS inverter can be determined by analytical techniques, it can also be determined by computer simulation. This is especially true when more complex CMOS logic circuits are considered. Using the appropriate transistor models in the simulation, the transient response can be produced. Obtaining an accurate transient response depends on using the correct transistor parameters. Some computer simulation problems in the end-of-chapter problems deal with propagation delay times. However, we will not go into detail here.

Test Your Understanding

TYU 16.9 Design a static CMOS logic circuit that implements the logic function $Y = \overline{(ABC + DE)}$. (Ans. NMOS design: A, B, C inputs to three NMOS devices in series and D, E inputs to two NMOS devices in series; then, three NMOS and two NMOS in parallel)

TYU 16.10 Design the width-to-length ratios of the transistors in the static CMOS exclusive-OR logic gate in Figure 16.41. Symmetrical switching times are desired and the switching times should correspond to the basic CMOS inverter. (Ans. All NMOS, $W_n = 2W$; all PMOS, $W_p = 4W$)



16.5 CLOCKED CMOS LOGIC CIRCUITS

Objective: • Analyze and design clocked CMOS logic gates.

The CMOS logic circuits considered in the previous section are called static circuits. One characteristic of a static CMOS logic circuit is that the output node always has a low-resistance path to either ground or V_{DD} . This implies that the output voltage is well defined and is never left floating.

Static CMOS logic circuits can be redesigned with an added clock signal while at the same time eliminating many of the PMOS devices. In general, the PMOS devices must be larger than NMOS devices. Eliminating as many PMOS devices as possible reduces the required chip area as well as the input capacitance. The low-power dissipation of the CMOS technology, however, is maintained.

Clocked CMOS circuits are dynamic circuits that generally precharge the output node to a particular level when the clock is at a logic 0. Consider the circuit in Figure 16.43. When the clock signal is low, or $CLK = \text{logic 0}$, M_{N1} is cut off and the current in the circuit is zero. Transistor M_{P1} is in a conducting state, but since the current is zero, then v_{O1} charges to V_{DD} . A high input to the CMOS inverter means that $v_O = 0$. During this phase of the clock signal, the gate of M_{P2} is precharged.

During the next phase, when the clock signal goes high, or $CLK = \text{logic 1}$, transistor M_{P1} cuts off and M_{N1} is biased in a conducting state. If input $A = \text{logic 0}$, then M_{NA} is cut off and there is no discharge path for voltage v_{O1} ; therefore, v_{O1} remains charged at $v_{O1} = V_{DD}$. However, if $CLK = \text{logic 1}$ and $A = \text{logic 1}$, then both M_{N1} and M_{NA} are biased in a conducting state, providing a discharge path for voltage v_{O1} . As v_{O1} is pulled low, output signal v_O goes high.

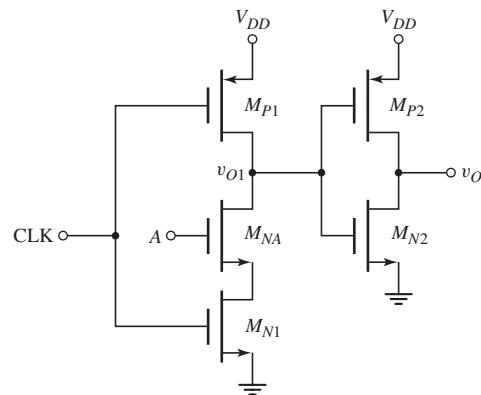


Figure 16.43 Simple clocked CMOS logic circuit

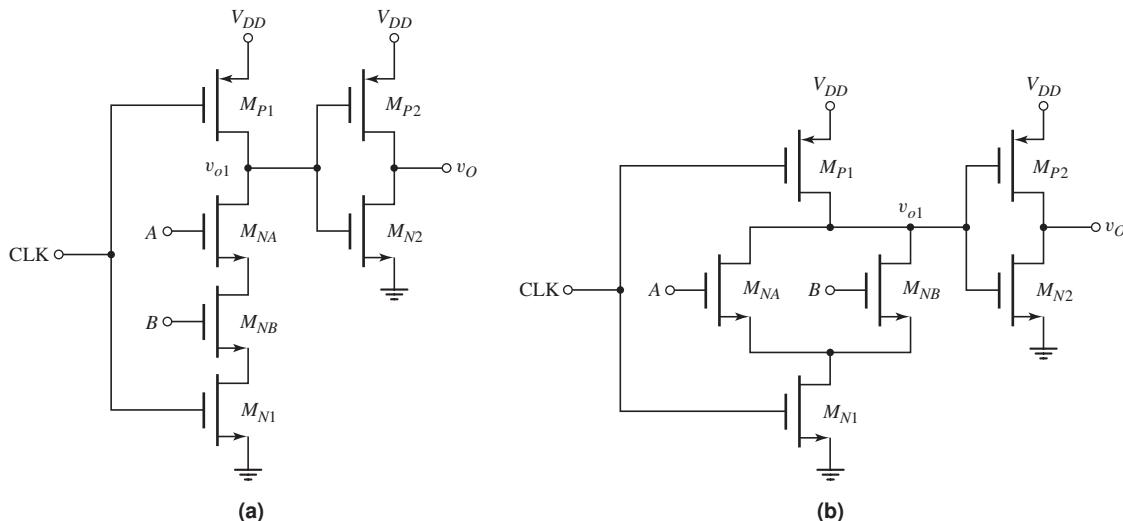


Figure 16.44 Clocked CMOS logic circuit: (a) AND function and (b) OR function

The quiescent power dissipation in this circuit is essentially zero, as it was in the standard CMOS circuits. A small amount of power is required to precharge output v_{O1} , if it had been pulled low during the previous half clock cycle.

The single NMOS transistor M_{NA} in Figure 16.43 can be replaced by a more complex NMOS logic circuit. Consider the two circuits in Figure 16.44. When $CLK = \text{logic 0}$, then M_{N1} cuts off and M_{P1} is in its conducting state in both circuits; then, v_{O1} is charged to $v_{O1} = V_{DD}$ and $v_O = 0$. For the circuit in Figure 16.44(a), when $CLK = \text{logic 1}$, voltage v_{O1} is discharged to ground or pulled low only when $A = B = \text{logic 1}$. In this case, v_O goes high. The circuit in Figure 16.44(a) performs the AND function. Similarly, the circuit in Figure 16.44(b) performs the OR function.

The advantage of the precharge technique is that it avoids the use of extensive pull-up networks: Only one PMOS and one NMOS transistor are required. This leads to an almost 50 percent savings in silicon area for larger circuits, and a reduction in capacitance resulting in higher speed. In addition, the static or quiescent power dissipation is essentially zero, so the circuit maintains the characteristics of CMOS circuits.

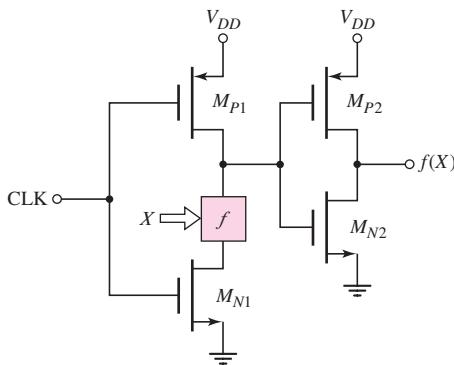


Figure 16.45 Generalized CMOS clocked logic circuit

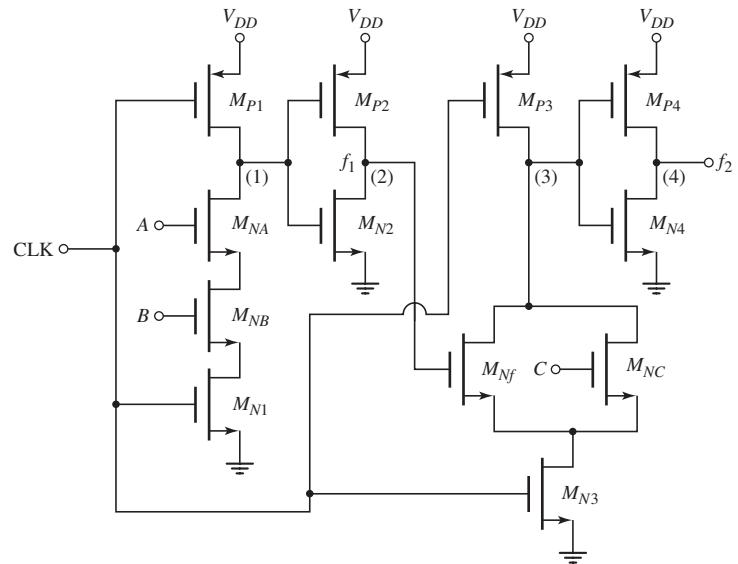


Figure 16.46 Cascaded clocked or domino CMOS logic circuit

The AND and OR logic transistors M_{NA} and M_{NB} in Figures 16.44(a) and 16.44(b) can be replaced by a generalized logic network as indicated in Figure 16.45. The box marked f is an NMOS pull-down network that performs a particular logic function $f(X)$ of n variables, where $X = (x_1, x_2, \dots, x_n)$. The NMOS circuit is a combination of series-parallel interconnections of n transistors. When the clock signal goes high, the CMOS inverter output is the logic function $f(X)$.

The set of X inputs to the logic circuits f is derived from the outputs of other CMOS inverters and clocked logic circuits. The means that when $CLK = \text{logic 0}$, the outputs of all CMOS inverters are a logic 0 during the precharge cycle. As a result, all n variables $X = (x_1, x_2, \dots, x_n)$ are a logic 0 during the precharge cycle. During this time, all NMOS transistors are cut off, which guarantees that output v_{O1} can be precharged to V_{DD} . There can then be only one possible transition at each node during the evaluation phase. The output of the CMOS buffer may change from a 0 to a 1.

An example of a cascaded domino CMOS circuit is shown in Figure 16.46. During the precharge cycle, in which $CLK = \text{logic 0}$, nodes 1 and 3 are charged high and nodes 2 and 4 are low. Also during this time, the inputs A , B , and C are all a logic 0. During the evaluation phase, in which $CLK = \text{logic 1}$, if $A = C = \text{logic 1}$ and $B = \text{logic 0}$, then node 1 remains charged high, $f_1 = \text{logic 0}$, and node 3 discharges through M_{NC} causing f_2 to go high. However, if, during the evaluation phase, $A = B = \text{logic 1}$ and $C = \text{logic 0}$, then node 1 is pulled low causing f_1 to go high, which in turn causes node 3 to go low and forces node 4 high. This chain of actions thus leads to the term **domino circuit**.

Test Your Understanding

TYU 16.11 Design a clocked CMOS domino logic circuit, such as shown in Figure 16.45, to generate an output $f(X) = A \cdot B \cdot C + D \cdot E$.

TYU 16.12 Sketch a clocked CMOS logic circuit that realizes the exclusive OR function.

16.6 TRANSMISSION GATES

Objective: • Analyze and understand the characteristics of NMOS and CMOS transmission gates.

Transistors can act as switches between driving circuits and load circuits. Transistors used to perform this function are called transmission gates. We will examine NMOS and CMOS transmission gates, which can also be configured to perform logic functions.

16.6.1 NMOS Transmission Gate

The NMOS enhancement-mode transistor in Figure 16.47(a) is a transmission gate connected to a load capacitance C_L , which could be the input gate capacitance of a MOS logic circuit. In this circuit, the transistor must be bilateral, which means it must be able to conduct current in either direction. This is a natural feature of MOS-FETs. Terminals a and b are assumed to be equivalent, and the bias applied to the transistor determines which terminal acts as the drain and which terminal acts as the source. The substrate must be connected to the most negative potential in the circuit, which is usually ground. Figure 16.47(b) shows a simplified circuit symbol for the **NMOS transmission gate** that is used extensively.

We assume that the NMOS transmission gate is to operate over a voltage range of zero-to- V_{DD} . If the gate voltage ϕ is zero, then the n-channel transistor is cut off and the output is isolated from the input. The transistor is essentially an open switch.

If $\phi = V_{DD}$, $v_I = V_{DD}$, and v_O is initially zero, then terminal a acts as the drain since its bias is V_{DD} , and terminal b acts as the source since its bias is zero. Current enters the drain from the input, charging up the capacitor. The gate-to-source voltage is

$$v_{GS} = \phi - v_O = V_{DD} - v_O \quad (16.75)$$

As the capacitor charges and v_O increases, the gate-to-source voltage decreases. The capacitor stops charging when the current goes to zero. This occurs when the gate-to-source voltage v_{GS} becomes equal to the threshold voltage V_{TN} . The maximum output voltage occurs when $v_{GS} = V_{TN}$, therefore, from Equation (16.75), we have

$$v_{GS}(\min) = V_{TN} = V_{DD} - v_O(\max) \quad (16.76(a))$$

or

$$v_O(\max) = V_{DD} - V_{TN} \quad (16.76(b))$$

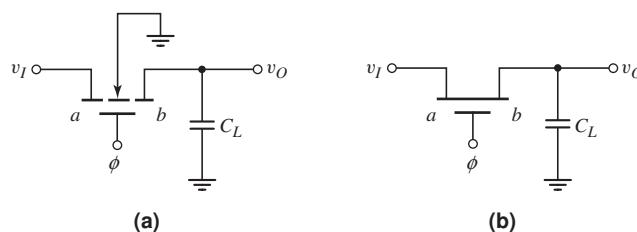


Figure 16.47 (a) NMOS transmission gate, showing substrate connection, and (b) simplified diagram

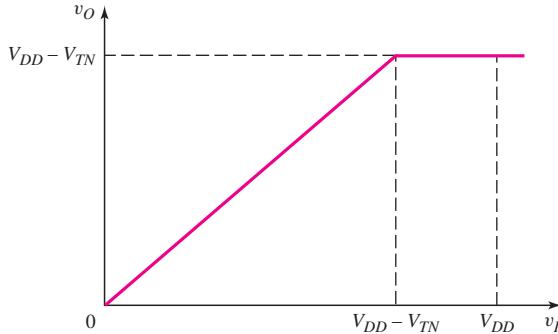


Figure 16.48 Output voltage versus input voltage characteristics of the NMOS transmission gate

where V_{TN} is the threshold voltage taking into account the body effect.

Equation (16.76(b)) demonstrates one disadvantage of an NMOS transmission gate. A logic 1 level degrades, or attenuates, as it passes through the transmission gate. However, this may not be a serious problem for many applications.

Figure 16.48 shows the quasi-static output voltage versus input voltage of the NMOS transmission gate. As seen in the figure, when $v_I = V_{DD}$, the output voltage is $v_O = V_{DD} - V_{TN}$ as we have discussed. For input voltages in the range $v_I < V_{DD} - V_{TN}$, the figure demonstrates that $v_O = v_I$. In this range of input voltages, the gate-to-source voltage is still greater than the threshold voltage. However, in steady-state, the current must be zero through the capacitor. In this case, the current becomes zero when the drain-to-source voltage is zero, or when $v_O = v_I$.

Now consider the situation in which $\phi = V_{DD}$, $v_I = 0$, and $v_O = V_{DD} - V_{TN}$ initially. Terminal *b* then acts as the drain and terminal *a* acts as the source. The gate-to-source voltage is

$$v_{GS} = \phi - v_I = V_{DD} - 0 = V_{DD} \quad (16.77)$$

The value of v_{GS} is a constant, and the capacitor discharges as current enters the NMOS transistor drain. The capacitor stops discharging when the current goes to zero. Since v_{GS} is a constant at V_{DD} , the drain current goes to zero when the drain-to-source voltage is zero, which means that the capacitor completely discharges to zero. This implies that a logic 0 is transmitted unattenuated through the NMOS transmission gate.

Using an NMOS transmission gate in a MOS circuit may introduce a dynamic condition. Figure 16.49 shows a cross section of the NMOS transistor in the transmission gate configuration. If $v_I = \phi = V_{DD}$, then the load capacitor charges to

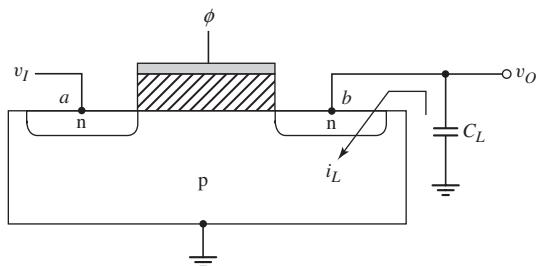


Figure 16.49 NMOS transmission gate with cross section of NMOS transistor

$v_O = V_{DD} - V_{TN}$. When $\phi = 0$, the NMOS device turns off and the input and output become isolated.

The capacitor voltage reverse biases the pn junction between terminal b and ground. A reverse-biased pn junction leakage current begins to discharge the capacitor, and the circuit does not remain in a static condition. This circuit is now dynamic in that the high output does not remain constant with time.

EXAMPLE 16.12

Objective: Estimate the rate at which the output voltage v_O in Figure 16.49 decreases with time when the NMOS transmission gate transistor is in its off state.

Assume the capacitor is initially charged to $v_O = 2.9$ V. Let $C_L = 0.2$ pF and assume the reverse-biased pn junction leakage current is a constant at $i_L = 100$ pA.

Solution: The voltage across the capacitor can be written as

$$v_O = -\frac{1}{C_L} \int i_L dt = -\frac{i_L}{C_L} t + K_1$$

where $K_1 = v_O(t = 0) = 2.9$ V is the initial condition. Therefore,

$$v_O(t) = 2.9 - \frac{i_L}{C_L} t$$

The rate at which the output voltage decreases is

$$\frac{dv_O}{dt} = -\frac{i_L}{C_L} = -\frac{100 \times 10^{-12}}{0.2 \times 10^{-12}} = -500 \text{ V/s} \Rightarrow -0.5 \text{ V/ms}$$

Therefore, in this example, the capacitor would completely discharge in 5.8 ms.

Comment: Even though the NMOS transmission gate may introduce a dynamic condition into a circuit, this gate is still useful in clocked logic circuits in which a clock signal is periodically applied to the NMOS transistor gate. If, for example, the clock frequency is 25 kHz, the clock pulse period is 40 μ s, which means that the output voltage would decay by only approximately 0.7 percent during a clock period.

EXERCISE PROBLEM

Ex 16.12: The threshold voltage of the NMOS transmission gate transistor in Figure 16.47(a) is $V_{TN} = 0.4$ V. Determine the output voltage v_O for: (a) $v_I = \phi = 2.5$ V; (b) $v_I = 1.8$ V, $\phi = 2.5$ V; (c) $v_I = 2.3$ V, $\phi = 2.5$ V; and (d) $v_I = 2.5$ V, $\phi = 1.5$ V. Neglect the body effect. (Ans. (a) $v_O = 2.1$ V, (b) $v_O = 1.8$ V, (c) $v_O = 2.1$ V, (d) $v_O = 1.1$ V)

EXAMPLE 16.13

Objective: Determine the output of an NMOS inverter driven by a series of NMOS transmission gates.

Consider the circuit shown in Figure 16.50. The NMOS inverter is driven by three NMOS transmission gates in series. Assume the threshold voltages of the NMOS transmission gate transistors and the NMOS driver transistor are

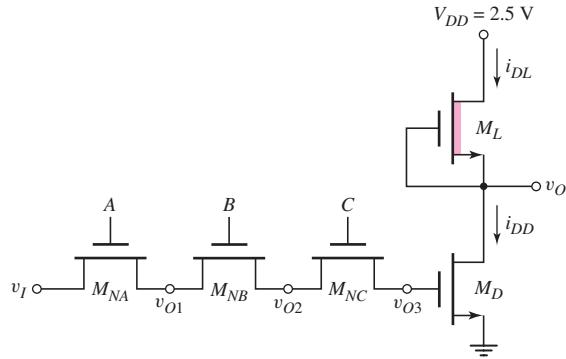


Figure 16.50 NMOS inverter driven by three NMOS transmission gates in series

$V_{TN} = 0.4$ V, and the threshold voltage of the load transistor is $V_{T_{NL}} = -0.6$ V. Let $K_D/K_L = 3$ for the inverter. Determine v_O for $v_I = 0$ and $v_I = 2.5$ V

Solution: The three NMOS transmission gates in series act as an AND/NAND function. If $v_I = 0$ and $A = B = C = \text{logic 1} = 2.5$ V, the gate capacitance to driver M_D becomes completely discharged, which means that $v_{O1} = v_{O2} = v_{O3} = 0$. Driver M_D is cut off and $v_O = 2.5$ V.

If $v_I = 2.5$ V and $A = B = C = \text{logic 1} = 2.5$ V, the three transmission gates are biased in their conducting state, and the gate capacitance of M_D becomes charged. For transistor M_{NA} , the current becomes zero when the gate-to-source voltage is equal to the threshold voltage, or, from Equation (16.76(b)),

$$v_{O1} = V_{DD} - V_{TN} = 2.5 - 0.4 = 2.1 \text{ V}$$

Transistors M_{NB} and M_{NC} also cut off when the gate-to-source voltages are equal to the threshold voltage; therefore

$$v_{O2} = v_{O3} = V_{DD} - V_{TN} = 2.5 - 0.4 = 2.1 \text{ V}$$

This result shows that the drain-to-source voltages of M_{NB} and M_{NC} are also zero. A threshold voltage drop is lost in the first transmission gate, but additional threshold voltage drops are not lost in subsequent NMOS transmission gates in series.

For a voltage of $v_{O3} = 2.1$ V applied to the gate of M_D , the driver is biased in the nonsaturation region and the load is biased in the saturation region. From $i_{DD} = i_{DL}$, we have

$$K_D [2(v_{O3} - V_{TN})v_O - v_O^2] = K_L [-V_{T_{NL}}]^2$$

The output voltage is found to be $v_O = 35.7$ mV.

If any one of the transmission gate voltages, A or B or C , switches to a logic 0, then v_{O3} will begin to discharge through a reverse-biased pn junction in the transmission gates, which means that v_O will increase with time.

Comment: In this example, the inverter is again in a dynamic condition; that is, when any transmission gate is cut off, the output voltage changes with time. However, this type of circuit can be used in clocked digital systems.

EXERCISE PROBLEM

Ex 16.13: Consider the NMOS inverter with enhancement load driven by an NMOS transmission gate in Figure 16.51. The threshold voltage of each n-channel transistor is $V_{TN} = 0.5$ V. Neglect the body effect. Design K_D/K_L of the inverter such that $v_O = 0.1$ V when: (a) $v_I = 2.8$ V, $\phi = 3.3$ V; and (b) $v_I = \phi = 2.8$ V. (Ans. (a) $K_D/K_L = 16.2$, (b) $K_D/K_L = 20.8$)

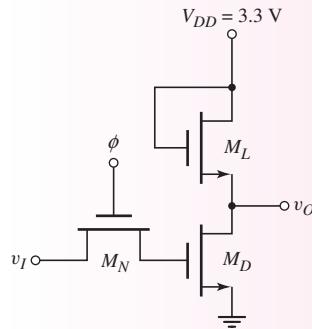


Figure 16.51 Figure for Exercise Ex16.13

16.6.2 NMOS Pass Networks

As integrated circuit technology advances, one emphasis is on increased circuit density. The maximum number of circuit functions per unit area is determined either by power dissipation density or by the area occupied by transistors and related devices.

One form of NMOS circuit logic that minimizes power dissipation and maximizes device density is called **pass transistor logic**. Pass transistor circuits use minimum-sized transistors, providing high density and high operating speed. The average power dissipation is due only to the switching power consumed by the driver circuits in charging and discharging the pass transistor control gates and driving the pass network inputs.

In this section, we present a few examples of NMOS pass transistor logic circuits. Consider the circuit in Figure 16.52. To determine the output response, we examine the conditions listed in Table 16.1 for the possible states of the input signals A and B . We assume that a logic 1 level is V_{DD} volts. In states 1 and 2, transmission gate M_{N2} is biased in its conducting state. For state 1, \bar{A} = logic 1 is transmitted to the output so f = logic $1'$, where the logic $1'$ level is $(V_{DD} - V_{TN})$. The logic 1 level is attenuated by one threshold voltage drop. For state 2, A = logic 0 is transmitted

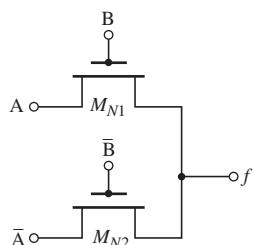


Figure 16.52 Simple NMOS pass logic network

Table 16.1 Input and output states for the circuit in Figure 16.52

State	A	B	\bar{A}	\bar{B}	M_{N1}	M_{N2}	f
1	0	0	1	1	off	on	$1'$
2	1	0	0	1	off	on	0
3	0	1	1	0	on	off	0
4	1	1	0	0	on	off	$1'$

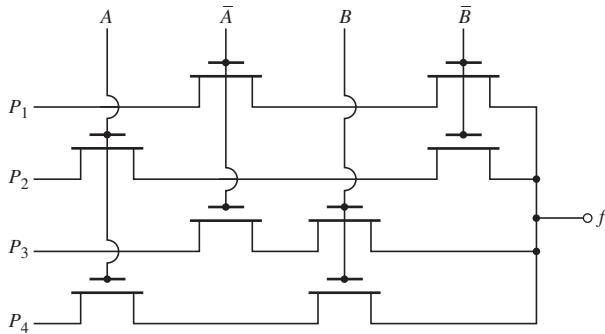


Figure 16.53 NMOS pass logic network example

Table 16.2 Input and output states for the circuit in Figure 16.53

State	A	B	\bar{A}	\bar{B}	f
1	0	0	1	1	P_1
2	1	0	0	1	P_2
3	0	1	1	0	P_3
4	1	1	0	0	P_4

unattenuated to the output. In states 3 and 4, transmission gate M_{N1} is biased in its conducting state. The $A = \text{logic 0}$ for state 3 is transmitted unattenuated to the output, and $A = \text{logic 1}$ for state 4 is attenuated during transmission; therefore, $f = \text{logic } 1'$. The output is thus the exclusive-NOR function.

Another example of an NMOS pass transistor logic circuit is shown in Figure 16.53. The output response as a function of the input gate controls A and B is shown in Table 16.2. This circuit is a multiplexer; that is, for a specific set of gate controls, the input signals P_i are individually passed to the output. By using both normal and inverted forms of A and B , four inputs can be controlled with just two variables.

A potential problem of NMOS pass transistor logic is that the output may be left floating in a high impedance state and charged high. Consider the circuit shown in Figure 16.54. If, for example, $B = C = \text{logic 0}$ and $A = \text{logic 1}$, then $f = \text{logic 1}'$, which is the logic 1 level attenuated by V_{TN} . When A is switched to logic 0, the output should be low, but there may not be a discharge path to ground, and the output may retain the logic 1' stored at the output capacitance.

The NMOS pass network must be designed to avoid a high impedance output by passing a logic 0 whenever a 0 is required at the output. A logic network that performs the logic function $f = A + \bar{B} \cdot C$, as indicated in Figure 16.54, is shown in Figure 16.55. The complementary function $\bar{f} = \bar{A} \cdot (B + \bar{C})$ attached at the output node drives the output to a logic 0 whenever $f = 0$.

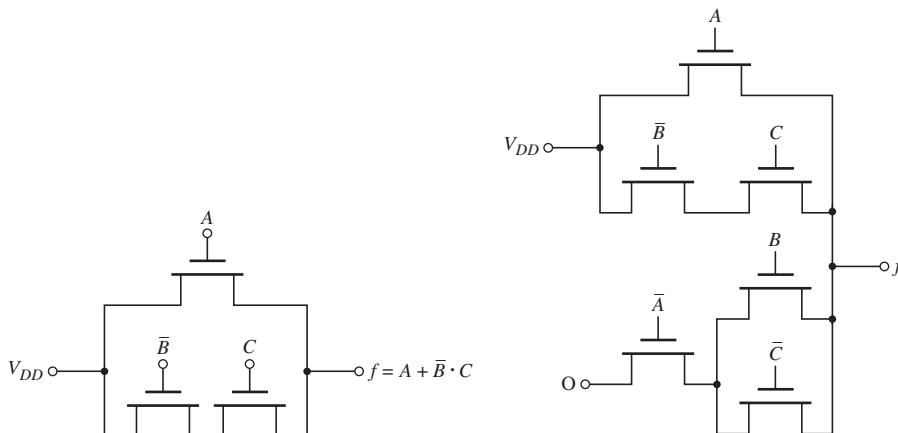


Figure 16.54 NMOS pass logic network with a potential problem

Figure 16.55 NMOS pass logic network with complementary function in parallel

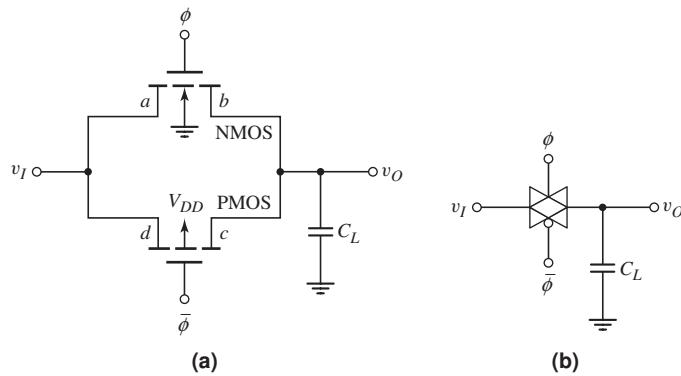


Figure 16.56 (a) CMOS transmission gate and (b) simplified circuit symbol

16.6.3 CMOS Transmission Gate

A **CMOS transmission gate** is shown in Figure 16.56(a). The parallel combination of NMOS and PMOS transistors, with complementary gate signals, allows the input signal to be transmitted to the output without the threshold voltage attenuation. Both transistors must be bilateral; therefore, the NMOS substrate is connected to the most negative potential in the circuit and the PMOS substrate is connected to the most positive potential (usually, ground and *V_{DD}*, respectively). Figure 16.56(b) shows a frequently used simplified circuit symbol for the CMOS transmission gate.

We again assume that the transmission gate is to operate over a voltage range of zero-to-*V_{DD}*. If the control voltages are *ϕ* = 0 and *ϕ̄* = *V_{DD}*, then both the NMOS and PMOS transistors are cut off and the output is isolated from the input. In this state, the circuit is essentially an open switch.

If *ϕ* = *V_{DD}*, *ϕ̄* = 0, *v_I* = *V_{DD}*, and *v_O* is initially zero, then for the NMOS device, terminal *a* acts as the drain and terminal *b* acts as the source, whereas for the PMOS device, terminal *c* acts as the drain and terminal *d* acts as the source. Current enters the NMOS drain and the PMOS source, as shown in Figure 16.57(a), to charge the load capacitor. The NMOS gate-to-source voltage is

$$v_{GSN} = \phi - v_O = V_{DD} - v_O \quad (16.78(a))$$

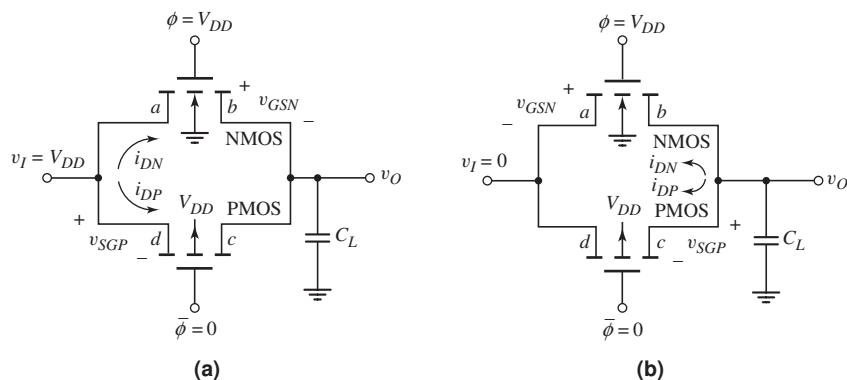


Figure 16.57 Currents and gate-source voltages in CMOS transmission gate for: (a) input high condition and (b) input low condition

and the PMOS source-to-gate voltage is

$$v_{SGP} = v_I - \bar{\phi} = V_{DD} - 0 = V_{DD} \quad (16.78(b))$$

As with the NMOS transmission gate, when $v_O = V_{DD} - V_{TN}$, the NMOS transistor cuts off and $i_{DN} = 0$ since $V_{GSN} = V_{TN}$. However, since the source-to-gate voltage of the PMOS device is a constant at $v_{SGP} = V_{DD}$, the PMOS transistor continues to conduct. The drain current i_{DP} goes to zero when the PMOS source-to-drain voltage goes to zero, or $v_{SDP} = 0$. This means that the load capacitor C_L continues to charge through the PMOS device until the output and input voltages are equal, or in this case, $v_O = v_I = V_{DD}$.

Consider what happens if $\phi = V_{DD}$, $\bar{\phi} = 0$, $v_I = 0$, and $v_O = V_{DD}$ initially. For the NMOS device, terminal a acts as the source and terminal b acts as the drain, whereas for the PMOS device, terminal c acts as the source and terminal d acts as the drain. Current enters the NMOS drain and the PMOS source, as shown in Figure 16.57(b), to discharge the capacitor. The NMOS gate-to-source voltage is

$$v_{GSN} = \phi - v_I = V_{DD} - 0 = V_{DD} \quad (16.79(a))$$

and the PMOS source-to-gate voltage is

$$v_{SGP} = v_O - \bar{\phi} = v_O - 0 = v_O \quad (16.79(b))$$

When $v_{SGP} = v_O = |V_{TP}|$, the PMOS device cuts off and i_{DP} goes to zero. However, since $v_{GSN} = V_{DD}$, the NMOS transistor continues conducting and capacitor C_L completely discharges to zero.

Using a CMOS transmission gate in a MOS circuit may introduce a dynamic condition. Figure 16.58 shows the CMOS transmission gate with simplified cross sections of the NMOS and PMOS transistors. If $\phi = 0$ and $\bar{\phi} = V_{DD}$, then the input and output are isolated. If $v_O = V_{DD}$, then the NMOS substrate-to-terminal b pn junction is reverse biased and capacitance C_L can discharge, as it did in the NMOS transmission gate. If, however, $v_O = 0$, then the PMOS terminal c -to-substrate pn junction is reverse biased and capacitance C_L can charge to a positive voltage. This circuit is therefore dynamic in that the output high or low conditions do not remain constant with time.

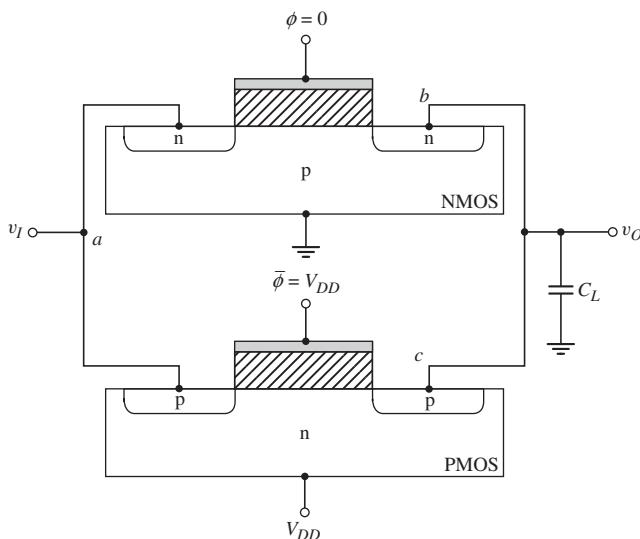


Figure 16.58 CMOS transmission gate showing cross sections of NMOS and PMOS transistors

16.6.4 CMOS Pass Networks

CMOS transmission gates may also be used in pass network logic design. CMOS pass networks use NMOS transistors to pass 0's, PMOS transistors to pass 1's, and CMOS transmission gates to pass a variable to the output. An example is shown in Figure 16.59. One PMOS transistor is used to transmit a logic 1, while transmission gates are used to transmit a variable that may be either a logic 1 or a logic 0. We can show that for any combination of signals, a logic 1 or logic 0 is definitely passed to the output.

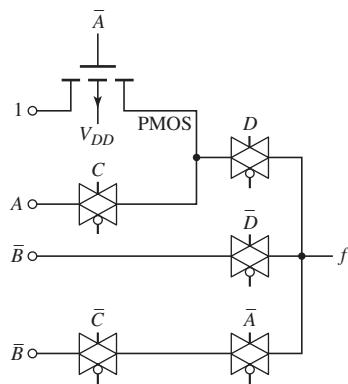


Figure 16.59 CMOS pass logic network

Test Your Understanding

TYU 16.13 Design an NMOS pass network to perform the logic function $f = A(B + C)$.

TYU 16.14 Consider the CMOS transmission gate in Figure 16.56(a). Assume transistor parameters of $V_{TN} = 0.4$ V and $V_{TP} = -0.4$ V. When $\phi = 2.5$ V, the input voltage v_I varies with time as $v_I = 2.5 - 0.2t$ for $0 \leq t \leq 12.5$ s. Let $v_O(t = 0) = 2.5$ V and assume $C_L = 0.2$ pF. Determine the range of times that the NMOS and PMOS devices are conducting. (Ans. NMOS conducting for $2 \leq t \leq 12.5$ s; PMOS conducting for $0 \leq t \leq 10.5$ s).



16.7 SEQUENTIAL LOGIC CIRCUITS

Objective: • Analyze and understand the characteristics of shift registers and various flip-flop designs.

In the logic circuits that we have considered in the previous sections, such as NOR and NAND logic gates, the output is determined only by the instantaneous values of the input signals. These circuits are therefore classified as combinational logic circuits.

Another class of circuits is called **sequential logic circuits**. The output depends not only on the inputs, but also on the previous history of its inputs. This feature gives sequential circuits the property of memory. Shift registers and flip-flops are typical examples of such circuits. We will also briefly consider a full-adder circuit. The characteristic of these circuits is that they store information for a short time until the information is transferred to another part of the system.

In this section, we introduce a basic shift register and the basic concept of a flip-flop. These circuits can become very complex and are usually described with logic diagrams. We will also introduce a CMOS full adder circuit in terms of its logic diagram and then provide the transistor implementation of this logic function. Additional information can be found in more advanced texts.

16.7.1 Dynamic Shift Registers

A **shift register** can be formed from transmission gates and inverters. Figure 16.60 shows a combination of NMOS transmission gates and NMOS depletion-load inverters. The clock signals applied to the gates of the NMOS transmission gates must be complementary, nonoverlapping pulses. The effective capacitances at the gates of M_{D1} and M_{D2} are indicated by the dotted connections to C_{L1} and C_{L2} .

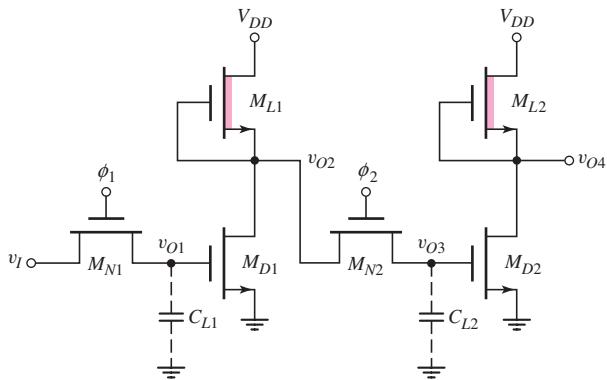


Figure 16.60 Dynamic shift register with NMOS inverters and transmission gates

If, for example, C_{L1} is initially uncharged when $v_{O1} = 0$ and if $v_I = V_{DD}$ when $\phi_1 = V_{DD}$, then a logic $1' = V_{DD} - V_{TN}$ voltage should exist at v_{O1} at the end of clock pulse ϕ_1 . The capacitance of C_{L1} charges through M_{N1} and the driving circuit of v_I . The effective RC time constant must be sufficiently small to achieve this charging effect. As v_{O1} goes high, v_{O2} goes low, but the low is not transmitted through M_{N2} as long as ϕ_2 remains low.

Figure 16.61 is used to determine the operation of this circuit and the voltages at various times. For simplicity, we assume that $V_{DD} = 5$ V and $V_{TN} = 1$ V for the NMOS drivers and transmission gate transistors.

At $t = t_1$, $v_I = \phi_1 = 5$ V, v_{O1} charges to $v_{DD} - V_{TN} = 4$ V, and v_{O2} goes low. At this time, M_{N2} is still cut off, which means that the values of v_{O3} and v_{O4} depend on the previous history. At $t = t_2$, ϕ_1 is zero, M_{N1} is cut off, but v_{O1} remains charged. At $t = t_3$, ϕ_2 is high, and the logic 0 at v_{O2} is transmitted to v_{O3} , which forces v_{O4} to 5 V. The input signal $v_I = 5$ V at $t = t_1$ has thus been transmitted to the output; therefore,

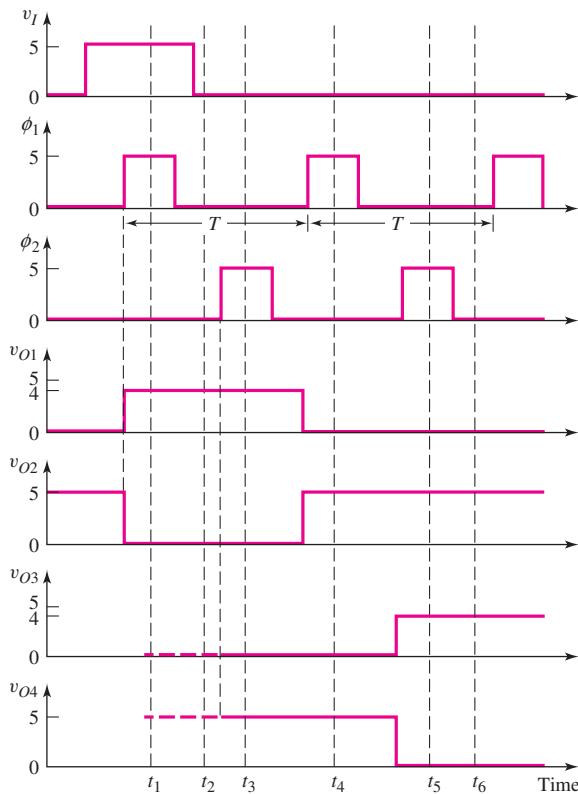


Figure 16.61 NMOS shift register voltages at various times

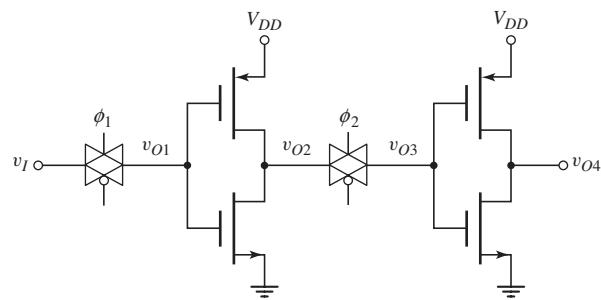


Figure 16.62 CMOS dynamic shift register

$v_{O4} = v_I = 5 \text{ V}$ at $t = t_3$. The input signal is transmitted, or *shifted*, from the input to the output during one clock cycle, making this circuit one stage of a shift register.

At $t = t_4$, $v_I = 0$, and $\phi_1 = 5 \text{ V}$, so that $v_{O1} = 0$ and $v_{O2} = 5 \text{ V}$. Since $\phi_2 = 0$, M_{N2} is cut off, and v_{O2} and v_{O3} are isolated. At $t = t_5$, $\phi_2 = 5 \text{ V}$, so that v_{O3} charges to $V_{DD} - V_{TN} = 4 \text{ V}$, and v_{O4} goes low (logic 0). At $t = t_6$, both NMOS transmission gates are cut off, and the two inverters remain in their previous states. It is important that ϕ_1 and ϕ_2 do not overlap, or the signal would propagate through the whole chain at once and we would no longer have a shift register.

In the dynamic condition of NMOS transmission gates, the high output voltage across the output capacitance does not remain constant with time; it discharges through the transmission gate transistor. This same effect applies to the shift register in Figure 16.60. For example, from Figure 16.61, at $t = t_2$, $v_{O1} = 4 \text{ V}$, $\phi_1 = 0$, and M_{N1} is cut off. Voltage v_{O1} will start to decay and v_{O2} will begin to increase. To prevent logic errors from being introduced into the system, the clock signal period T must be small compared to the effective RC discharge time constant. The circuit in Figure 16.60 is therefore called a **dynamic shift register**.

A dynamic shift register formed in a CMOS technology is shown in Figure 16.62. Operation of this circuit is very similar to that of the dynamic NMOS shift register, except for the voltage levels. For example, when $v_I = \phi_1 = V_{DD}$, then $v_{O1} = V_{DD}$ and $v_{O2} = 0$. When ϕ_2 goes high, then v_{O3} goes to zero, $v_{O4} = V_{DD}$, and the input signal is shifted to the output during one clock period.

16.7.2 R-S Flip-Flop

Flip-flops are bistable circuits usually formed by cross-coupling two NOR gates. Figure 16.63 shows an R-S flip-flop using NMOS NOR logic gates with depletion loads. As shown, M_1 , M_2 , and M_3 form one NOR gate, and M_4 , M_5 , and M_6 form the second. The outputs of the two NOR circuits are connected back to the inputs of the opposite NOR gates.

If we assume that S = logic 1 and R = logic 0, then M_1 is biased in its conducting state and output \bar{Q} is forced low. The inputs to both M_4 and M_5 are low, so output Q goes high to a logic 1 = V_{DD} . Transistor M_2 is then also biased in a conducting state. The two outputs Q and \bar{Q} are complementary and, by definition, the flip-flop is in the set state when Q = logic 1 and \bar{Q} = logic 0.

If S returns to logic 0, then M_1 turns off, but M_2 remains turned on so \bar{Q} remains low and Q remains high. Therefore, when S goes low, nothing in the circuit can force a change and the flip-flop stores this particular logic state.

When R = logic 1 and S = logic 0, then M_4 turns on so output Q goes low. With $S = Q$ = logic 0, then both M_1 and M_2 are cut off and \bar{Q} goes high. Transistor M_5 turns on, keeping Q low when R goes low. The flip-flop is now in the reset state.

If both S and R inputs were to go high, then both outputs Q and \bar{Q} would go low. However, this would mean that the outputs would not be complementary. Therefore, a logic 1 at both S and R is considered to be a forbidden or nonallowed condition. If both inputs go high and then return to logic 0, the state of the flip-flop is determined by whichever input goes low last. If both inputs go low simultaneously, then the outputs will flip into one state or the other, as determined by slight imbalances in transistor characteristics.

Figure 16.64 shows an R-S flip-flop using CMOS NOR logic gates. The outputs of the two NOR gates are connected back to the inputs of the opposite NOR gates to form the flip-flop.

If S = logic 1 and R = logic 0, then M_{N1} is turned on, M_{P1} is cut off, and \bar{Q} goes low. With $\bar{Q} = R$ = logic 0, then both M_{N3} and M_{N4} are cut off, both M_{P3} and M_{P4} are biased in a conducting state so that the output Q goes high. With Q = logic 1, M_{N2} is biased on, M_{P2} is biased off, and the flip-flop is in a set condition. When S goes low, M_{N1} turns off, M_{N2} remains conducting, so the state of the flip-flop does not change.

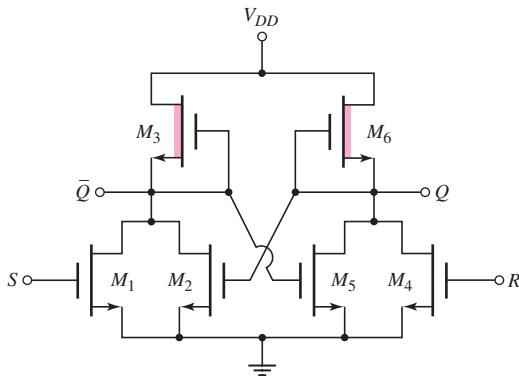


Figure 16.63 NMOS R-S flip-flop

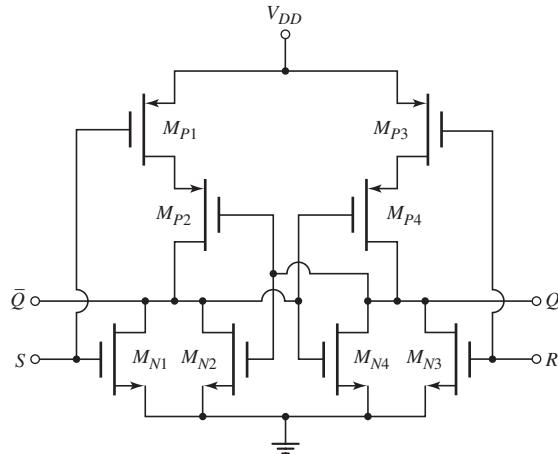


Figure 16.64 CMOS R-S flip-flop

When $S = \text{logic 0}$ and $R = \text{logic 1}$, then output Q is forced low, output \bar{Q} goes high, and the flip-flop is in a reset condition. Again, a logic 1 at both S and R is considered to be a forbidden or a nonallowed condition, since the resulting outputs are not complementary.

16.7.3 D Flip-Flop

A **D-type flip-flop** is used to provide a delay. The logic bit on the D input is transferred to the output at the next clock pulse. This flip-flop is used in counters and shift registers. The basic circuit is similar to the CMOS dynamic shift register in Figure 16.62, except that additional circuitry makes the D flip-flop a static circuit.

Consider the circuit in Figure 16.65. The CMOS inverter composed of M_{N2} and M_{P2} is driven by a CMOS transmission gate composed of M_{N1} and M_{P1} . A second CMOS inverter, M_{N3} and M_{P3} , is connected in a feedback configuration. If $v_I = \text{high}$, then v_{O1} goes high when the transmission gate is conducting, and output v_O , which is the input to the feedback inverter, goes low.

When the CMOS transmission gate turns off, the pn junction in the M_{N1} transmission gate transistor is reverse biased. In this case, however, voltage v_{O1} is not simply across the gate capacitance of inverter $M_{N2}-M_{P2}$. Transistor M_{P3} is biased in a conducting state, so the reverse-biased pn junction leakage current I_L is supplied through M_{P3} , as indicated in Figure 16.65. Since this leakage current is small, the source-to-drain voltage of M_{P3} will be small, and v_{O1} will remain biased at essentially V_{DD} . The circuit will therefore remain in this static condition.

Similarly, when v_{O1} is low and v_O is high, the pn junction in the M_{P1} transmission gate transistor is reverse biased and transistor M_{N3} is biased on. Transistor M_{N3} sinks the pn junction leakage current I'_L , and the circuit remains in this static condition until changed by a new input signal through the transmission gate.

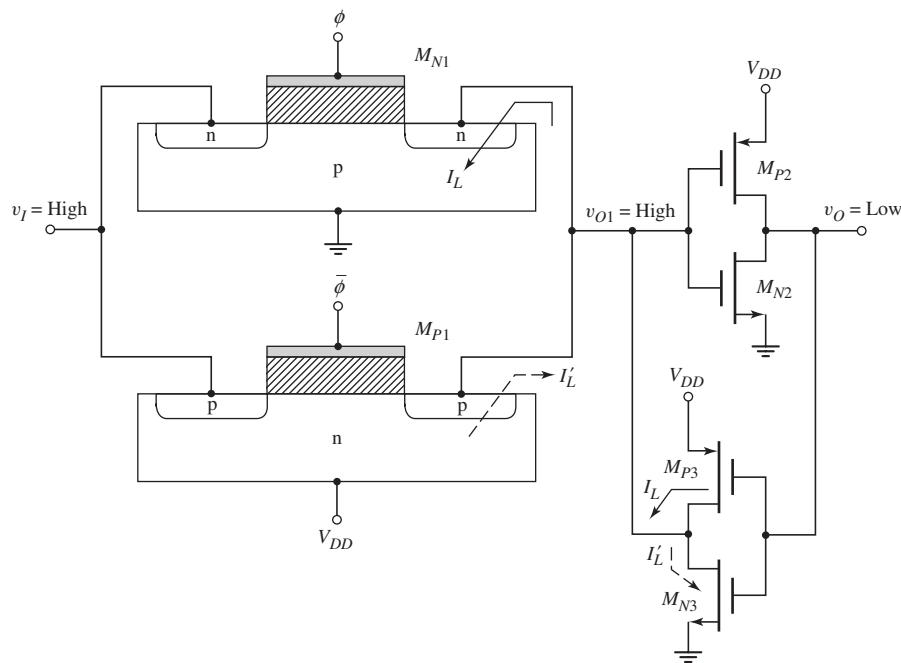


Figure 16.65 CMOS D-type flip-flop

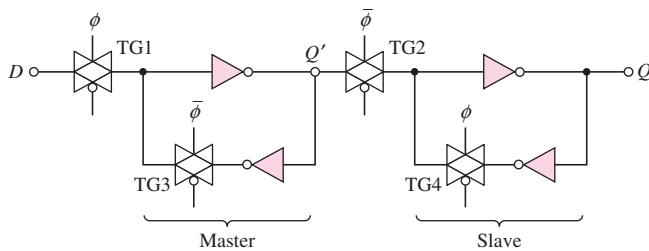


Figure 16.66 CMOS master–slave D flip-flop

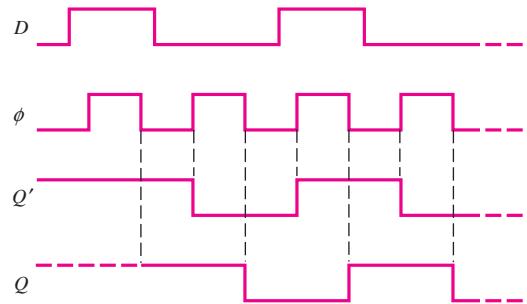


Figure 16.67 D flip-flop signals at various times

The circuit shown in Figure 16.66 is a master–slave configuration of a D flip-flop. When clock pulse ϕ is high, transmission gate TG1 is conducting, and data D goes through the first inverter, which means that $Q' = \bar{D}$. Transmission gate TG2 is off, so data stops at Q' . When clock pulse ϕ goes low, then TG3 turns on, and the master portion of the flip-flop is in a static configuration. Also when ϕ goes low, TG2 turns on, the data are transmitted through the slave portion of the flip-flop, and the output is $Q = Q' = D$. The data present when ϕ is high are transferred to the output of the flip-flop during the negative transition of the clock pulse. The various signals in the D flip-flop are shown in Figure 16.67.

Additional circuitry can be added to the D flip-flop in Figure 16.66 to provide a set and reset capability.

16.7.4 CMOS Full-Adder Circuit

One of the most widely used building blocks in arithmetic processing architectures is the one-bit full-adder circuit. We will first consider the logic diagram from the Boolean function and then consider the implementation in a conventional CMOS design.

Assuming that we have two input bits to be added plus a carry signal from a previous stage, the sum-out and carry-out signals are defined by the following two Boolean functions of three input variables A , B , and C .

$$\begin{aligned} \text{Sum-out} &= A \oplus B \oplus C \\ &= ABC + A\bar{B}\bar{C} + \bar{A}\bar{B}C + \bar{A}BC \end{aligned} \quad (16.80(\text{a}))$$

$$\text{Carry-out} = AB + AC + BC \quad (16.80(\text{b}))$$

The logic diagrams for these functions are shown in Figure 16.68. As we have seen previously, the implementation at the transistor level can be done with fewer transistors than would be used if all the NOR and NAND gates were actually connected as shown in the logic diagram.

Figure 16.69 is a transistor-level schematic of the one-bit full-adder circuit implemented in a conventional CMOS technology. We can understand the basic design from the logic diagram. For example, we may consider the NMOS portion of the carry-out signal. We see that transistors M_{NA1} and M_{NB1} are in parallel, to perform the basic OR function, and these transistors are in series with transistor M_{NC1} , to perform the basic AND function. These three transistors form the NMOS portion of the design of the two gates labeled G_1 and G_2 in Figure 16.68. We also have transistors M_{NA2} and M_{NB2} in series, to perform the basic AND function of gate G_3 . This set of

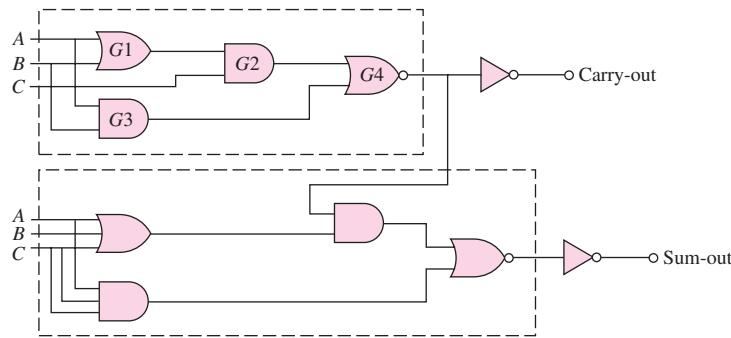


Figure 16.68 Gate configuration of the one-bit full adder

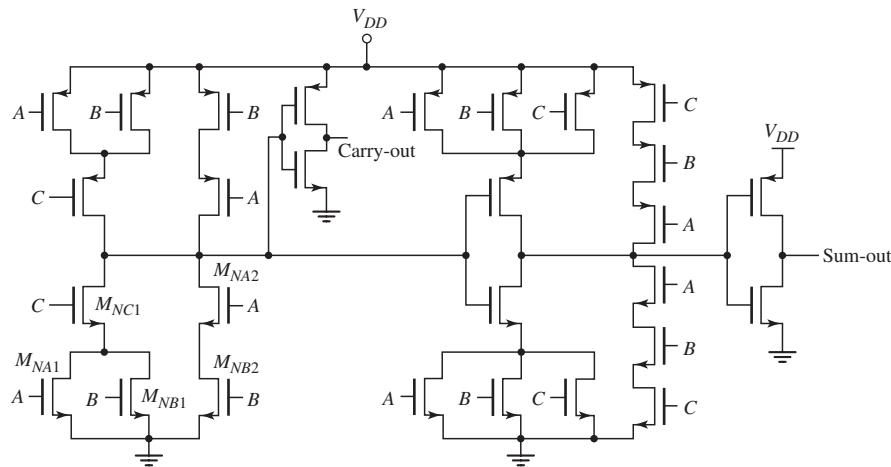


Figure 16.69 Transistor configuration of the CMOS one-bit full adder

two transistors is in parallel with the previous three transistors, and this configuration performs the basic OR function of gate G_4 . This output signal goes through an inverter to become the final carry-out signal.

We can go through the same discussion for the design of the NMOS portion of the sum-out signal. The PMOS design is then the complement of the NMOS design. As mentioned, the total number of transistors in the final design is considerably less than would have occurred if the basic OR and AND gates shown in the logic diagram were actually incorporated in the design.

16.8

MEMORIES: CLASSIFICATIONS AND ARCHITECTURES

Objective: • Discuss semiconductor memories.

In the previous sections of this chapter, various logic circuits were considered. Combinations of gates can be used to perform logic functions such as addition, multiplication, and multiplexing. In addition to these combinatorial logic functions, digital computers require some method of storing information. Semiconductor circuits form

one type of memory, considered in this chapter, and define a class of digital electronic circuits that are just as important as the logic gates.

A memory cell is a circuit, or in some cases just a single device, that can store a bit of information. A systematic arrangement of memory cells constitutes a memory. The memory must also include peripheral circuits to address and write data into the cells as well as detect data that are stored in the cells.

In this section, we define the various types of semiconductor memories, discuss the memory organization, and briefly consider address decoders. In the next section, we analyze in detail some of the basic memory cells and briefly discuss sense amplifiers.

16.8.1 Classifications of Memories

Two basic types of semiconductor memory are considered. The first is the **random access memory** (RAM), a read-write memory, in which each individual cell can be addressed at any particular time. The access time to each cell is virtually the same. Implicit in the definition of the RAM is that both the read and write operations are permissible in each cell with also approximately the same access time. Both static and dynamic RAM cells are considered.

A second class of semiconductor memory is the **read-only memory** (ROM). The set of data in this type of memory is generally considered to be fixed, although in some designs the data can be altered. However, the time required to write new data is considerably longer than the read access time of the memory cell. A ROM may be used, for example, to store the instructions of a system operating program.

A volatile memory is one that loses its data when power is removed from the circuit, while nonvolatile memory retains its data even when power is removed. In general, a random access memory is a volatile memory, while read-only memories are nonvolatile.

Random Access Memories

Two types of RAM are the static RAM (SRAM) and dynamic RAM (DRAM). A static RAM consists of a basic bistable flip-flop circuit that needs only a dc current or voltage applied to retain its memory. Two stable states exist, defined as logic 1 and logic 0. A dynamic RAM is an MOS memory that stores one bit of information as charge on a capacitor. Since the charge on the capacitor decays with a finite time constant (milliseconds), a periodic refresh is needed to restore the charge so that the dynamic RAM does not lose its memory.

The advantage of the SRAM is that this circuit does not need the additional complexity of a refresh cycle and refresh circuitry, but the disadvantage is that this circuit is fairly large. In general, SRAM requires six transistors. The advantage of a DRAM is that it consists of only one transistor and one capacitor, but the disadvantage is the required refresh circuitry and refresh cycles.

Read-Only Memories

There are two general types of ROM. The first is programmed either by the manufacturer (mask programmable) or by the user (programmable, or PROM). Once the ROM has been programmed by either method, the data in the memory are fixed and cannot be altered. The second type of ROM may be referred to as an alterable ROM in that the data in the ROM may be reprogrammed if desired. This type of ROM may

be called an EPROM (erasable programmable ROM), EEPROM (electrically erasable PROM), or flash memory. As mentioned, the data in these memories can be reprogrammed although the time involved is much longer than the read access time. In some cases, the memory chip may actually have to be removed from the circuit during the reprogramming process.

16.8.2 Memory Architecture

The basic memory architecture has the configuration shown in Figure 16.70. The terminal connections may include inputs, outputs, addresses, and read and write controls. The main portion of the memory involves the data storage. A RAM memory will have all of the terminal connections mentioned, whereas a ROM memory will not have the inputs and the write controls.

A typical RAM architecture, shown in Figure 16.71, consists of a matrix of storage bits arranged in an array with 2^M columns and 2^N rows. The array may be square, in which case M and N are equal. This particular array may be only one of several on a single chip. To read data stored in a particular cell within the array, a row address is inputted and decoded to select one of the row lines. All of the cells along this row are activated. A column address is also inputted and decoded to select one of the columns. The one particular memory cell at the intersection of the row and column addressed is then selected. The logic level stored in the cell is routed down a bit line to a sense amplifier.

Control circuits are used to enable or select a particular memory array on a chip and also to select whether data are to be read from or written into the memory cell. Memory chips or arrays are designed to be paralleled so that the memory capacity can be increased. The additional lines needed to address parallel arrays are called **chip select signals**. If a particular chip or array is not selected, then no memory cell is addressed in that particular array. The chip select signal controls the tristate output of the data-in and data-out buffers. In this way, the data-in and data-out lines to and from several arrays may be connected together without interfering with each other.

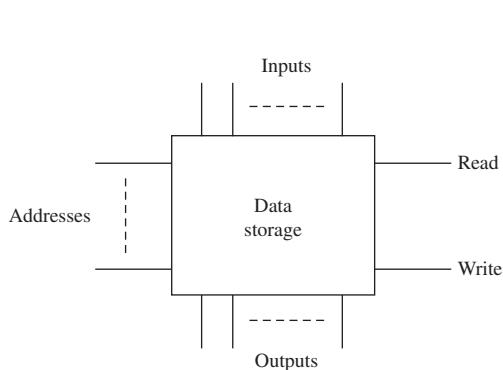


Figure 16.70 Schematic of a basic memory configuration

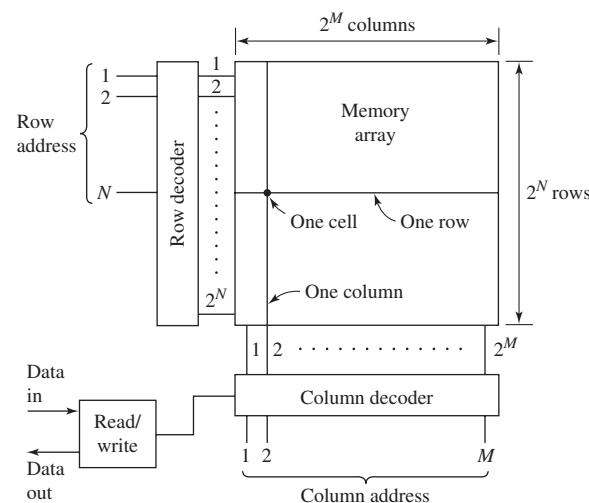


Figure 16.71 Basic random access memory architecture

16.8.3 Address Decoders

The row and column decoders in Figure 16.71 are essential elements in all memories. Access time and power consumption of memories may be largely determined by the decoder design. Figure 16.72 shows a simple decoder with a two-bit input. The decoder uses NAND logic circuits, although the same type of decoder may be implemented in NOR gates. The input word goes through input buffers that generate the complement as well as the signal.

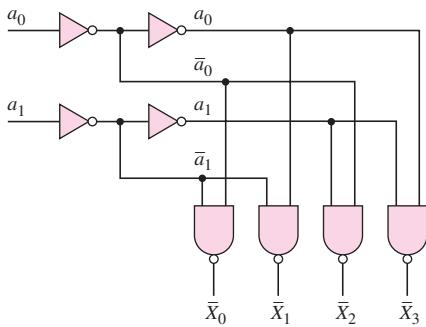


Figure 16.72 Simplified decoder with two-bit input

Another example of the direct implementation of a decoder is shown in Figure 16.73. Figure 16.73(a) shows a pair of NMOS input buffer-inverters, and Figure 16.73(b) shows a five-input NOR logic address decoder circuit using NMOS enhancement-mode drivers and a depletion load. A pair of input-buffer inverters is required for each input address line. The input signal is then required to drive only an inverter, while the buffer-inverter pair can be designed to drive the remainder of the logic circuits. The output of the NOR gate in Figure 16.73(b) would decode the address word 00110 and select the seventh row or column for a read or write operation. (Note: An input of 00000 is used to address the first row or column.)

As the size of the memory increases, the length of the address word must increase. For example, a 64-K (where 1 K = 1024 bits) memory whose cells are

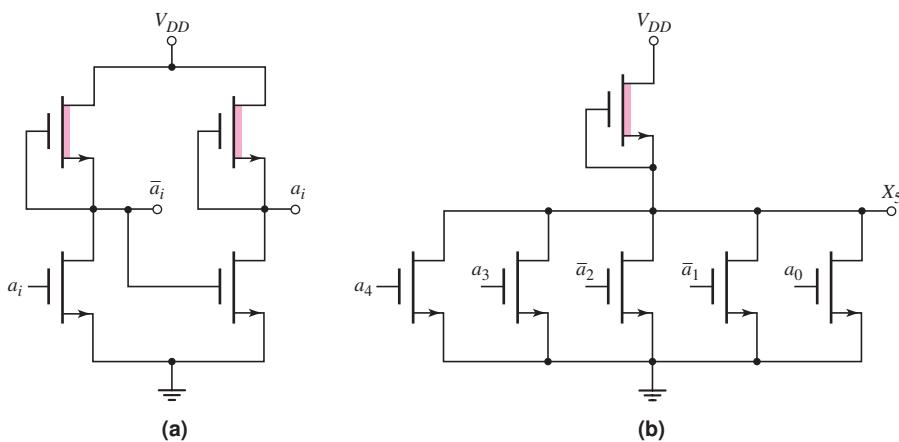


Figure 16.73 (a) Input buffer-inverter pair; (b) five-input NOR logic address decoder

arranged in a square array would require an 8-bit word for the row address and another 8-bit word for the column address. As the word size increases, the decoder becomes more complex, and the number of transistors and power dissipation may become large. In addition, the total capacitance of MOS decoder transistors and interconnect lines increase so that propagation delay times may become significant. The number of transistors required to design a decoder may be reduced by using a two-stage decoder using both NOR and NAND gates. These circuits may be found in more advanced textbooks on digital circuits.

Test Your Understanding

TYU 16.15 A NOR logic address decoder, such as shown in Figure 16.73(b), is used in both the row and column address decoders in a memory arranged in a square array. Calculate the number of decoder transistors required for a (a) 1-K, (b) 4-K, and (c) 16-K memory. (Ans. 384, 896, 2048 plus buffer transistors.)

16.9 RAM MEMORY CELLS

Objective: • Analyze and design random-access-memory (RAM) cells

In this section, we consider two designs of an NMOS static RAM (SRAM), one design of a CMOS static RAM, and one design of a dynamic RAM (DRAM). We also consider examples of sense amplifiers and read/write circuitry. This section is intended to present the basic concepts used in memory cell design. More advanced designs can again be found in advanced texts on digital circuits.

16.9.1 NMOS SRAM Cells

A static RAM cell is designed by cross-coupling the inputs and outputs of two inverters. In the case of an NMOS design, the load devices may be either depletion-mode transistors or polysilicon resistors, as shown in Figure 16.74. In either case, the

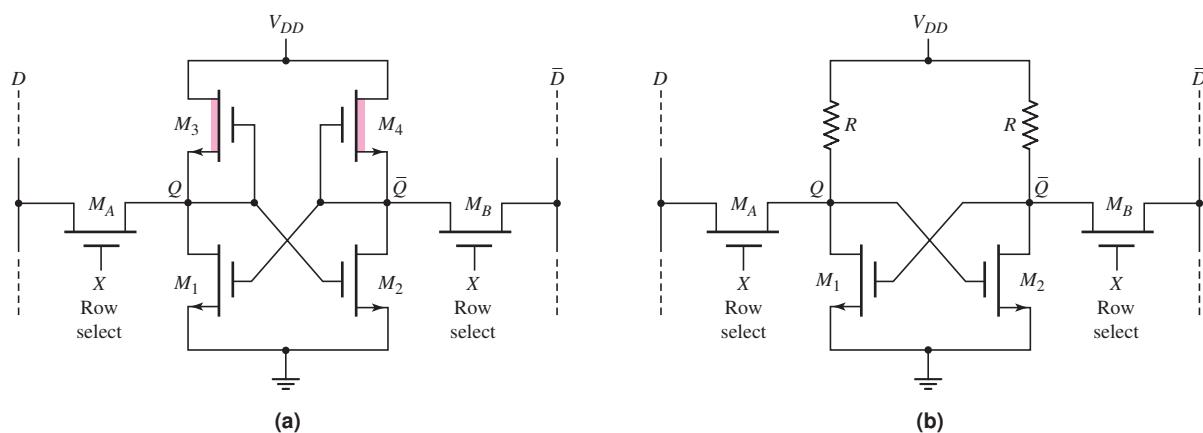


Figure 16.74 Static NMOS RAM cells with (a) depletion loads and (b) polysilicon resistor loads

inputs and outputs of the two inverters are cross-coupled to form a basic flip-flop. If transistor M_1 is turned on, for example, the output Q is low, which means that transistor M_2 is cut off. Since M_2 is cut off, the output \bar{Q} is high, ensuring that M_1 is turned on. Thus, we have a static situation as long as the bias voltage V_{DD} is applied to the circuit.

To access (read or write) the data contained in the memory cell, two NMOS transmission gate transistors, M_A and M_B , connect the memory cell to the complementary bit lines. When the word line signal or row select signal is low, both transmission gate transistors are cut off and the memory cell is isolated or in a standby condition. The data stored in the cell remain stored as long as power is applied to the cell. When the row select or word line signal goes high, the memory cell is then connected to the complementary data lines so that the data in the cell can be read or new data can be written into the cell.

One critical parameter in the design of RAM cells is power dissipation. As we will see in the following example, this is one situation in which incorporating a high-valued resistor as a load device improves the design. A lightly doped polysilicon load resistor is formed by ion implantation, which can accurately dope the polysilicon to produce the designed resistance value.

EXAMPLE 16.14

Objective: Determine the currents, voltages, and power dissipation in two NMOS SRAM cells. The first design uses a depletion-load device and the second design uses a resistor-load device.

Assume the following parameters: $V_{DD} = 3$ V and $k'_n = 60 \mu\text{A/V}^2$; driver transistors: $V_{TND} = 0.5$ V and $(W/L)_D = 2$; load devices: $V_{TNL} = -1.0$ V, $(W/L)_L = 1/2$, and $R = 2 \text{ M}\Omega$.

Solution (With Depletion Load): Assume M_2 is cut off in the circuit in Figure 16.74(a) so that $\bar{Q} = V_{DD} = 3$ V. M_1 is on in the nonsaturation region and M_3 is on in the saturation region. The drain current in M_1 and M_3 is then

$$i_D = \frac{k'_n}{2} \cdot \left(\frac{W}{L}\right)_L (V_{GSL} - V_{TNL})^2 = \frac{60}{2} \cdot \left(\frac{1}{2}\right) (0 - (-1))^2$$

or

$$i_D = 15 \mu\text{A}$$

The power dissipated in the circuit is then

$$P = i_D \cdot V_{DD} = (15)(3) = 45 \mu\text{W}$$

The logic 0 value of the Q output is found from

$$i_D = \frac{k'_n}{2} \cdot \left(\frac{W}{L}\right)_D [2(V_{GSD} - V_{TND})V_{DSD} - V_{DSD}^2]$$

or

$$15 = \frac{60}{2} \cdot (2)[2(3 - 0.5)Q - Q^2]$$

which yields

$$Q = 50.5 \text{ mV}$$

Solution (With Resistor Load): Again assume M_2 is cut off in the circuit in Figure 16.74(b) so that $\bar{Q} = V_{DD} = 3$ V. Again M_1 is on in the nonsaturation region. The drain current is found from

$$\frac{V_{DD} - Q}{R} = \frac{k'_n}{2} \cdot \left(\frac{W}{L}\right)_D [2(V_{GSD} - V_{TND})Q - Q^2]$$

or

$$\frac{3 - Q}{2} = \frac{60}{2} \cdot (2)[2(3 - 0.5)Q - Q^2]$$

[Note that dividing by megohms on the left agrees with microamperes on the right.]

We find

$$Q \cong 5 \text{ mV}$$

The drain current is then found:

$$i_D = \frac{V_{DD} - Q}{R} = \frac{3 - 0.005}{2} \cong 1.5 \mu\text{A}$$

The power dissipated in the circuit is then

$$P = i_D \cdot V_{DD} = (1.5)(3) = 4.5 \mu\text{W}$$

Comment: We see that the SRAM with the resistive load dissipates 10 times less power than the SRAM with the depletion-load device. Thus, for a given allowed power dissipation per chip, the memory with the resistive load could be 10 times larger than that using the depletion load device.

EXERCISE PROBLEM

Ex 16.14: A 16-K NMOS static RAM cell using a resistor load is to be designed. Each cell is to be biased at $V_{DD} = 2.5$ V. Assume transistor parameters as described in Example 16.14. The entire memory is to dissipate no more than 125 mW in standby. Design the value of R in each cell to meet this specification. (Ans. $R = 0.82 \text{ M}\Omega$)

Since the value of the load resistance R is, in general, very large, the memory must be designed so that the resistor R is not required to be a pull-up device. We will see this type of design later. The resistors can actually be fabricated on top of the NMOS transistors by a double-polysilicon technology, so that the cell with resistor load devices can be very compact, resulting in a high-density memory.

16.9.2 CMOS SRAM Cells

The basic six-transistor CMOS SRAM cell is shown in Figure 16.75. The inputs and outputs of the two CMOS inverters are cross-coupled so that the circuit will be in one of two static conditions. For example, if \bar{Q} is low, then M_{N1} is cut off so that Q is high, which in turn means that M_{P2} is cut off, ensuring that \bar{Q} remains low. The two NMOS transmission gate transistors again connect the basic memory cell to the complementary data lines.

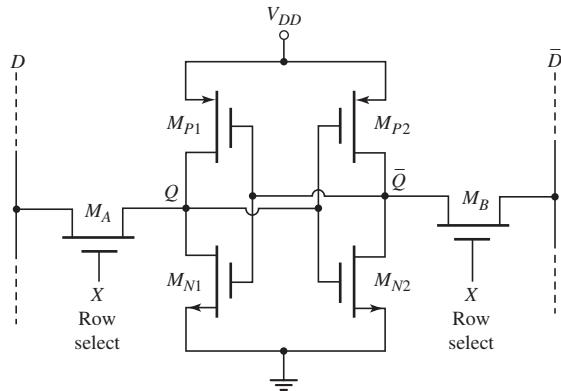


Figure 16.75 A CMOS static RAM cell

The traditional advantages of CMOS technology include low static power dissipation, superior noise immunity to either bipolar or NMOS, wide operating temperature range, sharp transfer characteristics, and wide voltage supply tolerance.

CMOS is inherently lower power than NMOS, since conducting paths between power and ground do not arise when the circuit is in one logic state or the other. In standard CMOS, the p- and n-channel devices in the memory cell and in the periphery circuits are in series and on at the same time only during switching. Current is, therefore, drawn only during switching. This makes SRAMs and CMOS extremely low power in standby, when there are only surface, junction, and channel leakage currents.

A more complete circuit of the CMOS static RAM is shown in Figure 16.76, which includes PMOS data line pull-up transistors on the complementary bit lines. If all word line signals are zero, then all pass transistors are turned off. The two data lines with the relatively large column capacitances are charged up by the column pull-up transistors, M_{P3} and M_{P4} , to the full V_{DD} voltage.

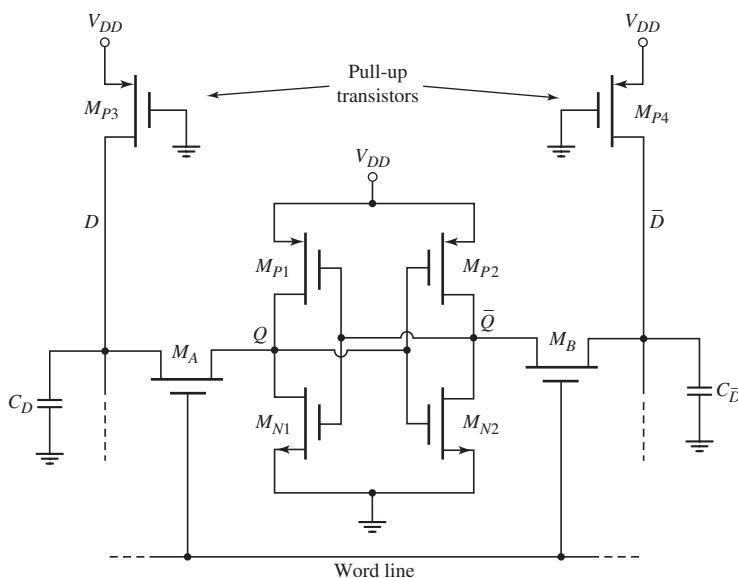


Figure 16.76 CMOS RAM cell including PMOS pull-up transistors

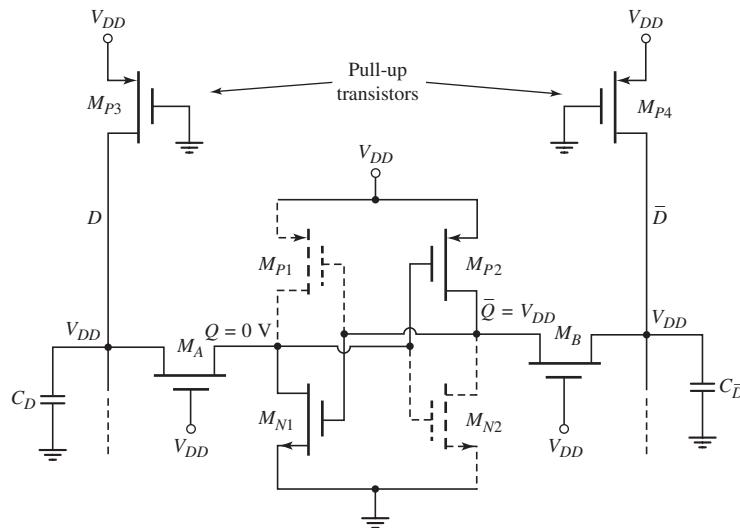


Figure 16.77 Voltage levels and “on” transistors in CMOS RAM cell at the beginning of the read cycle

To determine the (W/L) ratios of the transistors in a typical CMOS SRAM cell, two basic requirements must be taken into consideration. First, the read operation should not destroy the information stored in the cell, and second, the cell should allow for the modification of the data stored during a write operation. Consider a read operation in which a logic 0 ($Q = 0$ and $\bar{Q} = V_{DD}$) is stored in the cell. The voltage levels in the cell and on the data lines just prior to the read operation are shown in Figure 16.77. Transistors M_{P1} and M_{N2} are turned off while transistors M_{N1} and M_{P2} are biased in the nonsaturation region.

Immediately after the word select signal is applied to the pass transistors M_A and M_B , the voltage on the \bar{D} data line will not change significantly, since the pass transistor M_B is actually not conducting and no current flows. On the opposite side of the cell, current will flow through M_A and M_{N1} so that the voltage on the D data line will drop and the voltage Q will increase above its initial zero value. The key design point is that Q must not become larger than the threshold voltage of M_{N2} , so that M_{N2} remains cut off during the read phase. This will ensure that there is not a change in the data stored in the cell.

At the initial time the cell is addressed, we can assume that the D bit line remains at approximately V_{DD} , since the line capacitance cannot change instantaneously. The pass transistor M_A is biased in the saturation region and the transistor M_{N1} is biased in the nonsaturation region. Setting the drain currents through M_A and M_{N1} equal, we have

$$K_{nA}(V_{DD} - Q - V_{TN})^2 = K_{n1}[2(V_{DD} - V_{TN})Q - Q^2] \quad (16.81)$$

Setting $Q = Q_{\max} = V_{TN}$ as our design limit, then from Equation (16.81), we find the relation between the transistor width-to-length ratios to be

$$\frac{(W/L)_{nA}}{(W/L)_{n1}} < \frac{2(V_{DD}V_{TN}) - 3V_{TN}^2}{(V_{DD} - 2V_{TN})^2} \quad (16.82)$$

Assuming that $V_{DD} = 3$ V and $V_{TN} = 0.5$ V, we find that $(W/L)_{nA}/(W/L)_{n1} < 0.56$. So the width-to-length of the pass transistor should be approximately one-half that of

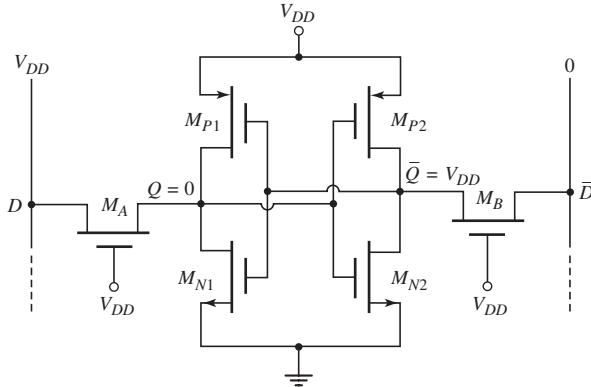


Figure 16.78 Voltage levels in the CMOS RAM at the beginning of a write cycle

the NMOS device in the memory cell. By symmetry, the same condition applies to the transistors M_{N2} and M_B .

We now need to consider the write operation. Assume that a logic 0 is stored and we want to write a logic 1 into the memory cell. Figure 16.78 shows the initial voltage levels in the CMOS SRAM cell when the cell is first addressed at the beginning of the write cycle. Transistors M_{P1} and M_{N2} are initially turned off, and M_{N1} and M_{P2} are biased in the nonsaturation region. The cell voltages are $Q = 0$ and $\bar{Q} = V_{DD}$ just before the pass transistors are turned on. The data line D is held at V_{DD} and the complementary data line \bar{D} is forced to a logic 0 value by the write circuitry. We may assume that $\bar{D} = 0$ V for analysis purposes. The voltage Q will remain below the threshold voltage of M_{N2} because of the condition given by Equation (16.82). Consequently, the voltage at Q is not sufficient to switch the state of the memory cell. To switch the state of the cell, the voltage at \bar{Q} must be reduced below the threshold voltage of M_{N1} , so that M_{N1} will turn off. When $\bar{Q} = V_{TN}$, then M_B is biased in the nonsaturation region and M_{P2} is biased in the saturation region. Equating drain currents, we have

$$K_{p2}(V_{DD} + V_{TP})^2 = K_{nB}[2(V_{DD} - V_{TN})V_{TN} - V_{TN}^2] \quad (16.83(a))$$

which can be written in the form

$$\frac{K_{p2}}{K_{nB}} < \frac{2(V_{DD}V_{TN}) - 3V_{TN}^2}{(V_{DD} + V_{TP})^2} \quad (16.83(b))$$

Considering the width-to-length ratios, we find

$$\frac{(W/L)_{p2}}{(W/L)_{nB}} < \frac{k'_n}{k'_p} \cdot \frac{2(V_{DD}V_{TN}) - 3V_{TN}^2}{(V_{DD} + V_{TP})^2} \quad (16.84)$$

Assuming that $V_{DD} = 3$ V, $V_{TN} = 0.5$ V, $V_{TP} = -0.5$ V, and $(k'_n/k'_p) = (\mu_n/\mu_p) = 2$, we find that $(W/L)_{p2}/(W/L)_{nB} < 0.72$.

From previous results, if we assume that the width-to-length of the pass transistor is one-half that of the NMOS in the memory cell, and if we assume that the width-to-length of the PMOS in the memory cell is 0.7 that of the pass transistor, then the width-to-length of the PMOS in the cell should be approximately 0.35 that of the NMOS in the memory cell.

16.9.3 SRAM Read/Write Circuitry

An example of a read/write circuit at the end of a column is shown in Figure 16.79. We may consider the write portion of the circuit as shown in Figure 16.80(a). We may note that if the column is not selected, then M_3 is cut off and the two data lines are held at their precharged value of V_{DD} . When $X = Y = 1$, then the one-bit cell shown is addressed. If $\bar{W} = 1$ then the write cycle is deselected and both M_1 and M_2 are cut off. For $\bar{W} = 0$ and $D = 1$, M_1 is cut off and M_2 is turned on so that the \bar{D} data line is pulled low while the D data line remains high. The logic 1 is then written into the cell. For $\bar{W} = 0$ and $D = 0$, the D data line is pulled low and the \bar{D} data line is held high so that logic 0 is written into the cell.

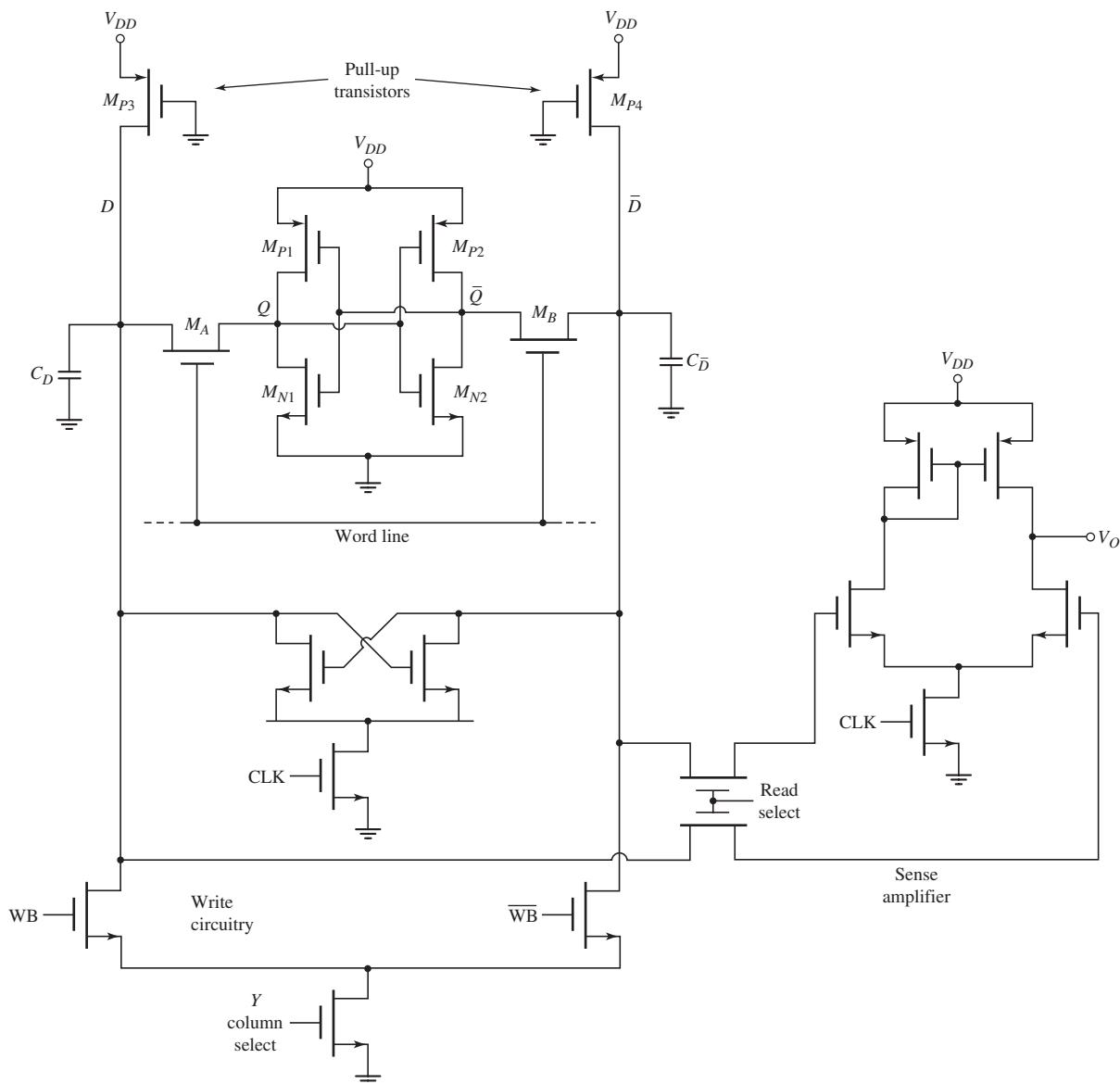


Figure 16.79 Complete circuit diagram of a CMOS RAM cell with write and read circuitry

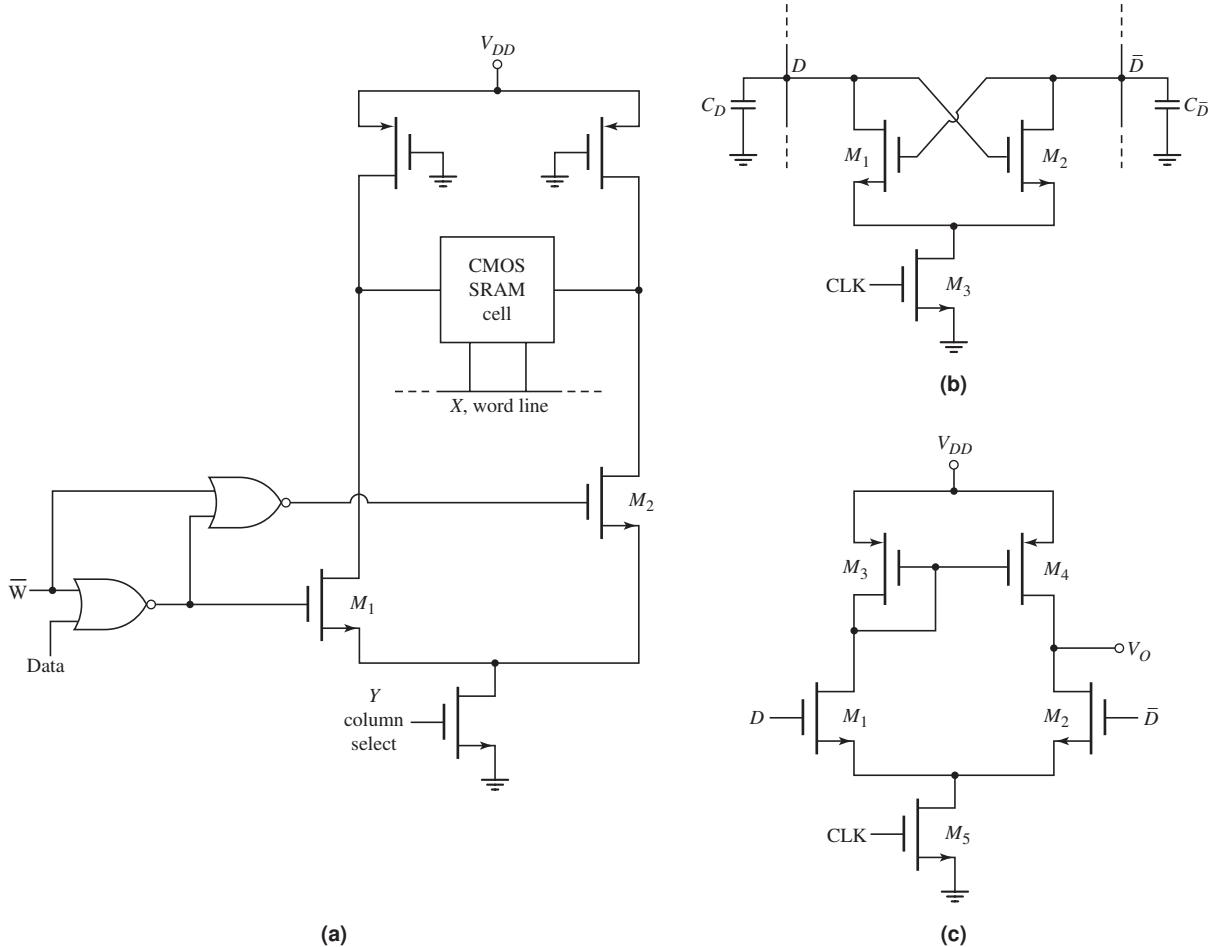


Figure 16.80 (a) Write circuitry associated with CMOS RAM cell; (b) cross-coupled NMOS sense amplifier; (c) CMOS differential sense amplifier

Figure 16.80(b) shows the NMOS cross-coupled sense amplifier that is in the complete circuit of Figure 16.79. This circuit does not generate an output signal, but rather amplifies the small difference in the data bit lines. Suppose that a logic 1 is to be read from the memory cell. When the cell is addressed, the D bit line is high and the \bar{D} bit line voltage begins to decrease. This means that when the M_3 transistor turns on, the M_2 transistor turns on harder than M_1 so that the \bar{D} bit line voltage is pulled low and the M_1 transistor will eventually turn off.

Figure 16.80(c) shows the differential amplifier that senses the output of the memory cell. Note that this sense amplifier is connected to the bit lines through a couple of pass transistors, as seen in Figure 16.79. If the input signal to the pass transistors is also a function of the column select signal, then this configuration enables the use of one main sense amplifier to read the data out of several columns, one at a time. When the clock signal is zero, the M_3 transistor in the differential amplifier is cut off and the common source node of M_1 and M_2 is pulled high, which means the output voltage is pulled high. When a memory cell is selected and the clock goes high, M_3 turns on. If a logic 1 level is to be read, then D remains high and the \bar{D} line

voltage decreases. This means that the M_2 transistor will turn off and the output voltage remains high. If a logic 0 is to be read, then the D line voltage decreases and \bar{D} remains high. The transistor M_1 will turn off while M_2 is turned on so that the output voltage goes low.

16.9.4 Dynamic RAM (DRAM) Cells

The CMOS RAM cell just considered requires six transistors and five lines connecting each cell, including the power and ground connections. A substantial area, then, is required for each memory cell. If the area per cell could be reduced, then higher-density RAM arrays would be possible.

In a dynamic RAM cell, a bit of data is stored as charge on a capacitor, where the presence or absence of charge determines the value of the stored bit. Data stored as charge on capacitors cannot be retained indefinitely, since leakage currents will eventually remove the stored charge. Thus the name *dynamic* refers to the situation in which a periodic refresh cycle is required to maintain the stored data.

One design of a DRAM cell is the one-transistor cell that includes a pass transistor M_S plus a storage capacitor C_S , shown in Figure 16.81. Binary information is stored in the form of zero charge on C_S (logic 0) and stored charge on C_S (logic 1). The cell is addressed by turning on the pass transistor via the word line signal WL and charges are transferred into or out of C_S on the bit line BL. The storage capacitor is isolated from the rest of the circuit when M_S is off, but the stored charge on C_S decreases because of the leakage current through the pass transistor. This effect was discussed in detail in Section 16.6 during the analysis of the NMOS pass transistor. As a result of this leakage, the cell must be refreshed regularly to restore its original condition.

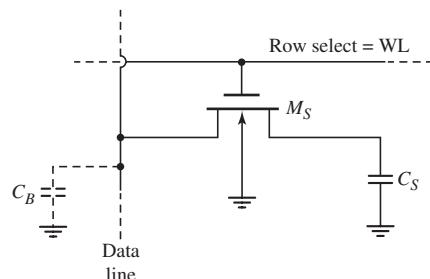


Figure 16.81 One-transistor dynamic RAM cell

An example of a sense amplifier to detect the charge stored in the memory cell is shown in Figure 16.82. On one side of the amplifier is a memory cell that either stores a full charge or is empty, depending on the binary value of the data. On the other side of the amplifier is a reference cell with a reference or dummy storage capacitor C_R that is one-half the value of the storage capacitor. The charge on C_R will then be one-half the logic 1 charge on C_S . A cross-coupled dynamic latch circuit is used to detect the small voltage differences and to restore the signal levels. The capacitors C_D and C_{DR} represent the relatively large parasitic bit line and reference bit line capacitances.

In the standby mode, the bit lines on both sides of the sense amplifier are precharged to the same potential. During the read cycle, both the WL and $D-WL$

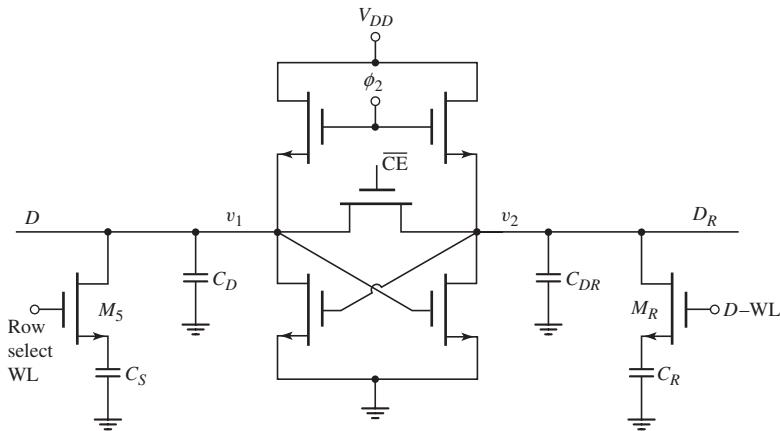


Figure 16.82 Sense amplifier configuration for dynamic RAM cell

address signals go high allowing the charges in the cells to be redistributed along the bit lines. After the charge equalization and since the charge in the dummy cell is half the full charge, then $v_1 < v_2$ when the memory cell is empty or a logic 0, and $v_1 > v_2$ when the memory cell is full or a logic 1. The sense amplifier detects and amplifies the voltage difference between the bit lines, and will latch at the logic level stored in the basic memory cell.

Test Your Understanding

TYU 16.16 A six-transistor CMOS SRAM cell is biased at $V_{DD} = 2.5$ V. The transistor parameters are $V_{TN} = +0.4$ V, $V_{TP} = -0.4$ V, and $(\mu_n/\mu_p) = 2.5$. Determine the relative width-to-length ratios such that Equations (16.81) through (16.84) are satisfied in terms of read/write requirements. (Ans. $[(W/L)_{nA}/(W/L_{n1})] = 0.526$, $[(W/L)_p/(W/L)_{nB}] = 0.862$)

TYU 16.17 A one-transistor DRAM cell is composed of a 0.05 pF storage capacitor and an NMOS transistor with a 0.5 V threshold voltage. A logic 1 is written into the cell when both the data line and row-select line are raised to 3 V. Sensing circuitry permits the stored charge to decay to 50 percent of its original value. Refresh occurs every 1.5 ms. Determine the maximum allowed leakage current that can exist. (Ans. $I = 41.7$ pA)



16.10 READ-ONLY MEMORY

Objective: • Analyze read-only memories (ROM)

We consider several examples of read-only memories in this section. The intent is again to provide an introduction to this type of memory. In the case of EPROMs and EEPROMs, the development effort has been directed toward the characteristics of the basic memory cell.

16.10.1 ROM and PROM Cells

We consider two types of ROMs. The first example is a mask-programmed ROM, in which contacts to devices are selectively included or excluded in the final manufacturing process to obtain the desired memory pattern. Figure 16.83 shows an example of an NMOS 16×1 mask-programmed ROM. Enhancement-mode NMOS transistors are fabricated in each of the 16 cell positions (the substrate connections are omitted for clarity). However, gate connections are fabricated only on selected transistors. The transistors M_1 – M_4 are column-select transistors and M_0 is a depletion-mode load device.

The inputs X_O , X_1 , Y_O , and Y_1 are the row- and column-select signals. If, for example, $X_O = \bar{X}_1 = Y_O = Y_1 = 1$, then the M_{12} transistor is addressed. Transistors M_{12} and M_3 turn on with this address, forcing the output to a logic 0. If the address changes, for example, to $\bar{X}_O = X_1 = \bar{Y}_O = \bar{Y}_1 = 1$, then the transistor M_{23} is addressed. However, this transistor does not have a gate connection and consequently never turns on, so the output is a logic 1.

The mask-programmed memory discussed is only a 16×1 -bit ROM, while a more useful memory would contain many more bits. Memories can be organized in any desired manner, such as a 2048×8 for a 16-K memory. This ROM is a non-volatile memory, since the data stored are not lost when power is removed.

The second example of a ROM is a user-programmed ROM. The data pattern is defined by the user after the final manufacture rather than during the manufacture.

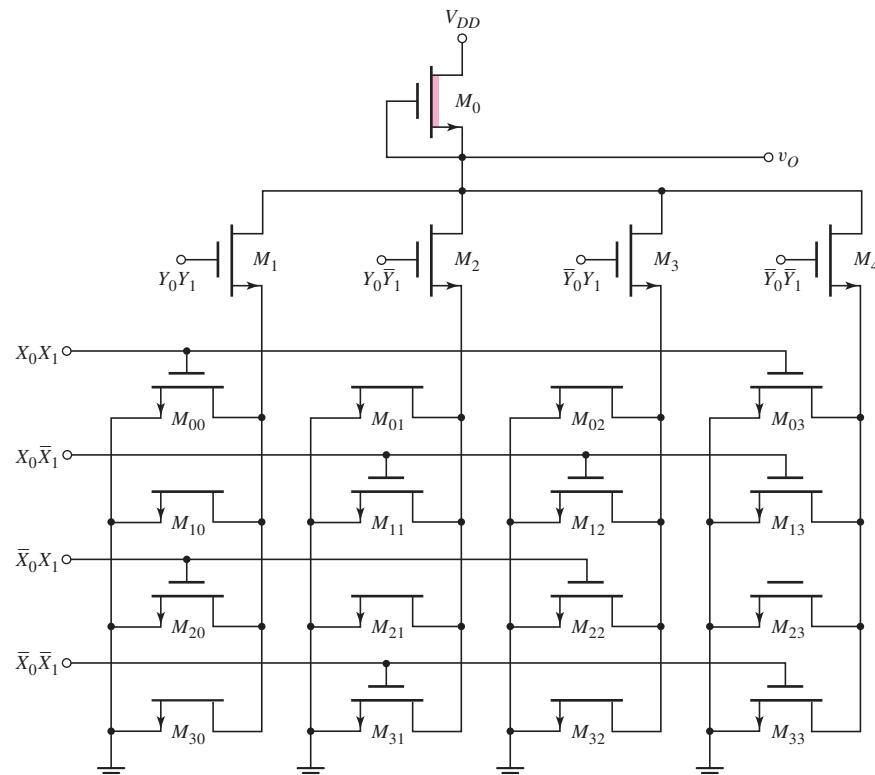


Figure 16.83 An NMOS 16×1 mask-programmable ROM

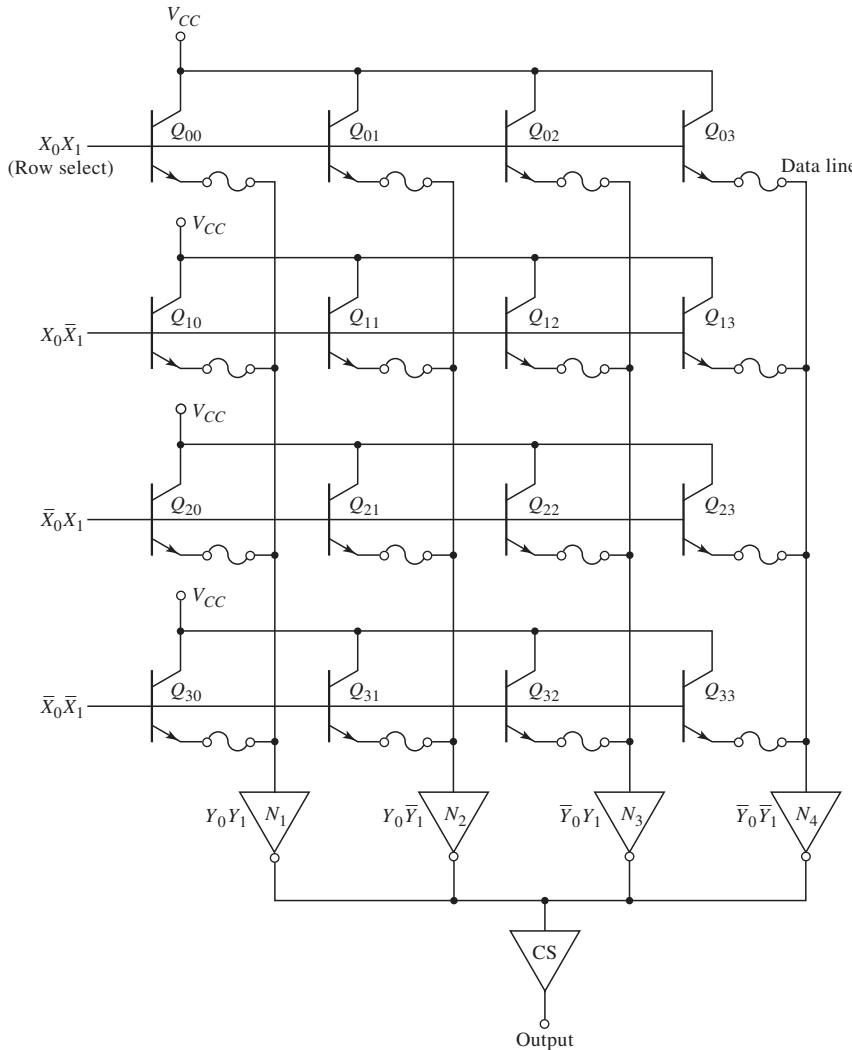
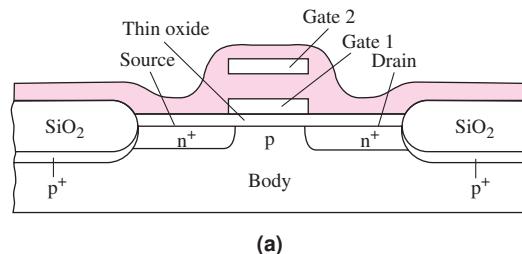


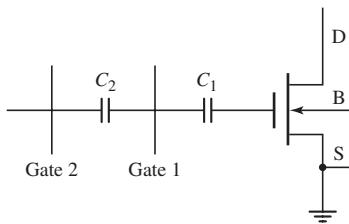
Figure 16.84 A bipolar fuse-linked user-programmable ROM

One specific type is shown schematically in Figure 16.84. A small fuse is in series with each emitter and can be selectively “blown” or left in place by the user. If, for example, the fuse in Q_{00} is left in place and this transistor is addressed by $X_O = X_1 = Y_O = Y_1 = 1$, then Q_{00} turns on, raising the data line voltage at the emitter of Q_{00} . The inverter N_1 is enabled, making the output a logic 0. If the fuse is blown in this transistor, then the input to the inverter is a logic 0, so the output is a logic 1.

The polysilicon fuse in the emitter of an npn bipolar transistor has a fairly low resistance, so with the fuse in place and at low currents, there is very little voltage drop across the fuse. When the current through the fuse is increased to the 20 to 30 mA range, the heating of the polysilicon fuse causes the temperature to increase. The silicon oxidizes, forming an insulator that effectively opens the path between the data line and the emitter. The bipolar ROM circuit with the fuses either in place or “blown” form a permanent ROM that is not alterable and is also nonvolatile.



(a)



(b)

Figure 16.85 (a) Cross section of erasable programmable ROM; (b) equivalent circuit

16.10.2 EPROM and EEPROM Cells

An EPROM transistor is shown in Figure 16.85. The device has a double gate, with gate 1 being a “floating gate” that has no electrical contact. Gate 2 is used for cell selection, taking the role of the single gate of an MOS transistor.

Operation of this EPROM cell relies on being able to store charge on the floating gate. Initially, we assume no charge on the floating gate so that with gate 2, drain, and source grounded, the potential of gate 1 is also zero. As the voltage on gate 2 increases, the gate 1 voltage rises also, but at a lower rate as determined by the capacitive divider. The net effect of this is to effectively raise the threshold voltage of this MOSFET as seen from gate 2. However, when the gate 2 voltage is raised sufficiently (approximately twice the normal threshold voltage), a channel forms. Under these conditions, the device provides a stored logic 0 when used in the NOR array.

To write a logic 1 into this cell, both gate 2 and drain are raised to about 25 V while the source and substrate remain at ground potential. A relatively large drain current flows because of normal device conduction characteristics. In addition, the high field in the drain–substrate depletion region results in avalanche breakdown of the drain–substrate junction, with a considerable additional flow of current. The high field in the drain depletion region accelerates electrons to high velocity such that a small fraction traverse the thin oxide and become trapped on gate 1. When the gate 2 and drain potentials are reduced to zero, the negative charge on gate 1 forces its potential to approximately -5 V. If the gate 2 voltage for reading is limited to +5 V, then a channel never forms. Thus a logic 1 is stored in the cell.

Gate 1 is completely surrounded by silicon dioxide (SiO_2), an excellent insulator, so charge can be stored for many years. Data can be erased, however, by exposing the cells to strong ultraviolet (UV) light. The UV radiation generates electron–hole pairs in the SiO_2 making the material slightly conductive. The negative charge on the gate can then leak off, restoring the transistor to its original uncharged condition. These EPROMs must be assembled in packages with transparent covers so

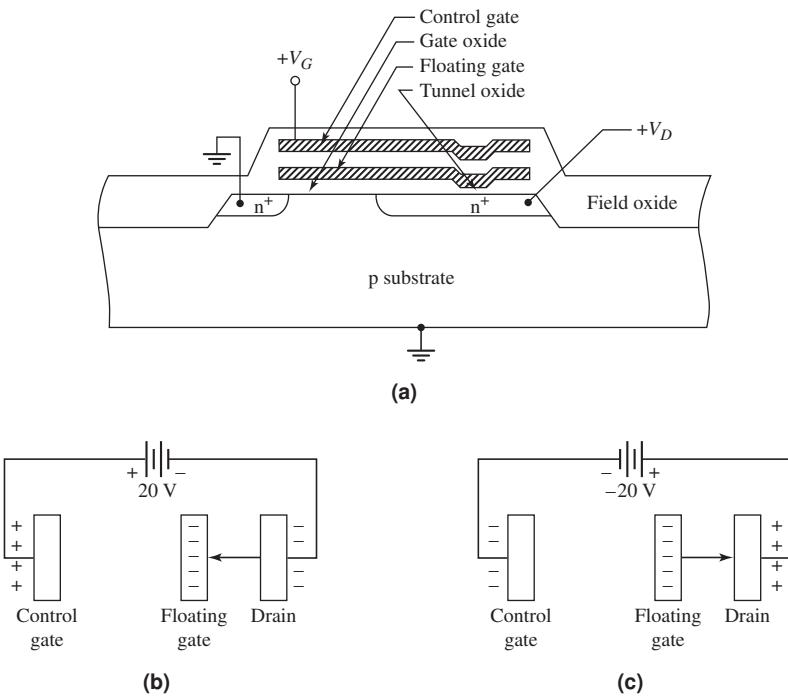


Figure 16.86 (a) Cross section of a floating-gate electrically erasable programmable ROM; (b) charging the floating gate; (c) discharging the floating gate

the silicon chip may be exposed to UV radiation. One disadvantage is that the entire memory must be erased before any reprogramming can be done. In general, reprogramming must also be done on specialized equipment; therefore, the EPROM must be removed from the circuit during this operation.

In the EEPROM, each individual cell can be erased and reprogrammed without disturbing any other cell. The most common form of EEPROM is also a floating gate structure; one example is shown in Figure 16.86(a). The memory transistor is similar to an n-channel MOSFET, but with a physical difference in the gate insulator region. Charge may exist on the floating gate that will alter the threshold voltage of the device. If a net positive charge exists on the floating gate, the n-channel MOSFET is turned on, whereas if zero or negative charge exists on the floating gate, the device is turned off.

The floating gate is capacitively coupled to the control gate with the tunnel oxide thickness less than 200 Å. If 20 V is applied to the control gate while keeping $V_D = 0$, electrons tunnel from the n⁺-drain region to the floating gate as demonstrated in Figure 16.86(b). This puts the MOSFET in the enhancement mode with a threshold voltage of approximately 10 V, so the device is effectively off. If zero volts is applied to the control gate and 20 V is applied to the drain terminal, then electrons tunnel from the floating gate to the n⁺-drain terminal as demonstrated in Figure 16.86(c). This leaves a net positive charge on the floating gate that puts the device in the depletion mode with a threshold voltage of approximately -2 V, so the device is effectively on. If all voltages are kept to within 5 V during the read cycle, this structure can retain its charge for many years.

16.11 DATA CONVERTERS

Objective: • Discuss the basic concepts in A/D and D/A converters.

Most physical signals exist in analog form. These signals include, for example, audio or speech and the output of transducer circuits. Some analog signal processing, such as amplifying the output of a microphone prior to the connection to speakers, may occur. However, digital signal processing may be required to convert an analog signal into digital form prior to transmission of the signal to a satellite receiver, for example. Therefore, analog-to-digital (A/D) and digital-to-analog (D/A) converters are an important class of integrated circuits.

16.11.1 Basic A/D and D/A Concepts

In this section, we briefly consider a few basic concepts used in A/D and D/A conversions. Figure 16.87 shows the block diagram representations of A/D and D/A converters. An analog signal v_A is applied to the input of the A/D converter and the output is an N -bit digital signal that can be represented as

$$v_D = \frac{b_1}{2^1} + \frac{b_2}{2^2} + \frac{b_3}{2^3} + \cdots + \frac{b_N}{2^N} \quad (16.85)$$

where b_1 , b_2 , etc. are the bit coefficients that are either a 1 or 0. The bit b_1 is the most significant bit (MSB) and the bit b_N is the least significant bit (LSB). The input to the D/A converter is the N -bit digital signal and the output is an analog signal v'_A . Ideally, the output analog signal v'_A is an exact replication of the input analog signal v_A .

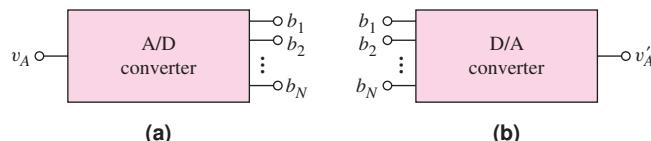


Figure 16.87 Block diagram representations of (a) A/D converter and (b) D/A converter

The analog signal is to be converted to a digital form as indicated in Equation (16.85). Consider, for example, an analog signal represented by a voltage in the range $0 \leq v_A \leq 5$ V. Assume the digital signal is a 6-bit word. The 6-bit word represents 64 discrete values. The analog signal will then be divided into 64 values, with each bit representing $5 \text{ V}/64 = 0.078125 \text{ V}$. The analog-to-digital conversion can be visualized in Figure 16.88.

When the analog input voltage is, for example, $v_A = \frac{5}{64}$ V, the digital output is 000001 and when the analog input voltage is $v_A = 2(\frac{5}{64})$ V, the digital output is 000010. However, we see that when the input is in the range $\frac{1}{2}(\frac{5}{64}) < v_A < \frac{3}{2}(\frac{5}{64})$ V, the digital output is constant at $v_D = 000001$. There is an inherent quantization error in the A/D conversion. A larger number of bits in the digital signal reduces the quantization error, but requires a more complex circuit.

The same effect occurs at the output of the D/A converter. Since the digital input signal exists in discrete steps or increments, the output signal will also occur in discrete steps or increments. An example is shown in Figure 16.89. The output signal

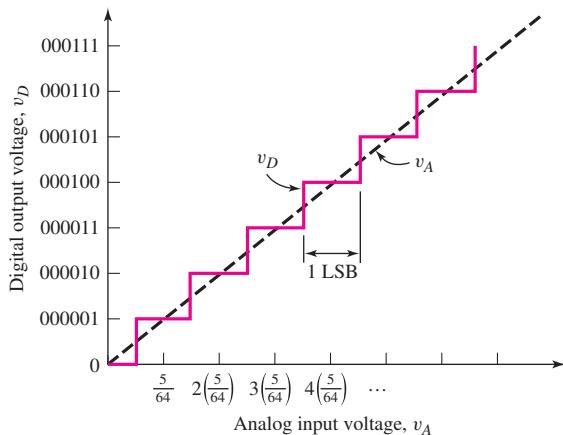


Figure 16.88 Digital output versus analog input for a 6-bit A/D converter

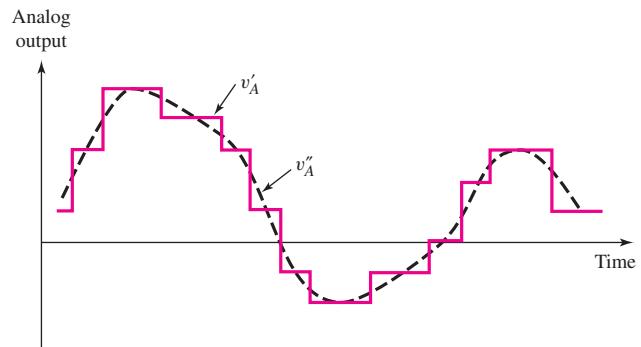


Figure 16.89 Discrete analog output v'_A and smoothed output v''_A versus time from a D/A converter

v'_A is in the form of stair steps. Normally, this signal will be fed through a low-pass filter to smooth out the signal to produce the dotted signal v''_A in the figure. The desired result is that the signal v''_A be as close to the original signal v_A as possible.

16.11.2 Digital-to-Analog Converters

We will consider a few basic D/A converters to gain an appreciation of the techniques used in these circuits.

Weighted-Resistor 4-Bit D/A

A simple circuit for a 4-bit D/A converter was shown in Chapter 9 in Figure P9.37. This circuit is repeated here in Figure 16.90 for convenience. The circuit is a summing amplifier and includes a reference voltage V_R , four weighted input resistors, four switches, and an op-amp with a feedback resistor. With $R_F = 10 \text{ k}\Omega$, we find the output voltage to be

$$v_O = \left(\frac{b_1}{2} + \frac{b_2}{4} + \frac{b_3}{8} + \frac{b_4}{16} \right) (5) \text{ V} \quad (16.86)$$

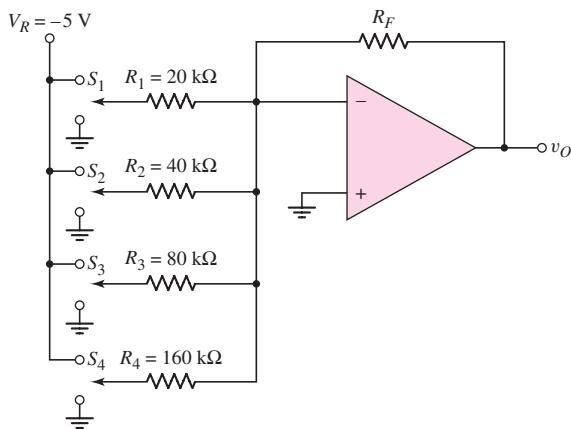


Figure 16.90 A 4-bit weighted-resistor D/A converter

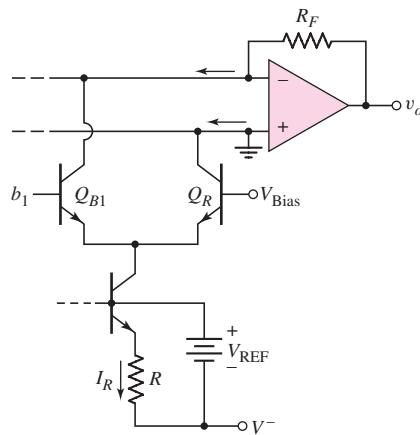


Figure 16.91 Example of a current switch in the MSB position in a weighted resistor of D/A converter

One factor that determines the accuracy of the circuit is the precision of the weighted input resistors and the feedback resistor. As the number of bits increases, the size of the weighted input resistance increases for the lesser significant bits. The accuracy for large resistance values becomes more difficult to maintain. The size of this D/A converter is in general limited to a 4-bit input.

Another factor that determines the accuracy of the D/A circuit is the precision of the switches. An example of a current switch, showing only the MSB, is shown in Figure 16.91. If the bit b_1 is a logic 1 ($>V_{\text{Bias}}$), then Q_{B1} is turned on and Q_R is turned off so that the current I_R is switched through Q_{B1} . This current becomes a component of the current through the feedback resistor. If b_1 is a logic 0 ($<V_{\text{Bias}}$), then Q_{B1} is turned off and Q_R is turned on so that the current is switched to ground. Because of the virtual ground, we may note that the collector voltages of Q_{B1} and Q_R are essentially identical.

For the circuit to operate properly, the base-emitter voltage of all the transistors must be the same. Since the currents are smaller for the lesser significant bits, the base-emitter areas must be reduced in order to maintain the same current density. The type of circuit shown in Figure 16.91 then requires a wide range of base-emitter areas in the same way that it requires a wide range of weighted resistor values.

R-2R Ladder Network D/A

A circuit that eliminates the wide spread in weighted input resistor values in the previous circuit is an $R-2R$ ladder resistive network. Consider the circuit shown in Figure 16.92. Assuming the switches are ideal, the current in each resistor is a constant because of the virtual ground concept.

Consider node X in the circuit. We can note that the resistance denoted as R_X is $R_X = 2R$. This same resistance occurs at every node in the circuit as indicated on the figure. Therefore, the current entering each node splits evenly as shown at node X . We have that $I_{N-1} = \frac{1}{2}I_{N-2}$. This effect again occurs at every node in the circuit. Therefore, we have

$$I_1 = 2I_2 = 4I_3 = \dots = 2^{N-2}I_{N-1} = 2^{N-1}I_N \quad (16.87)$$

Setting the feedback resistance to $R_F = R$, we have the output voltage given by

$$v_o = (-V_{\text{REF}}) \left(\frac{b_1}{2} + \frac{b_2}{4} + \dots + \frac{b_N}{2^N} \right) \quad (16.88)$$

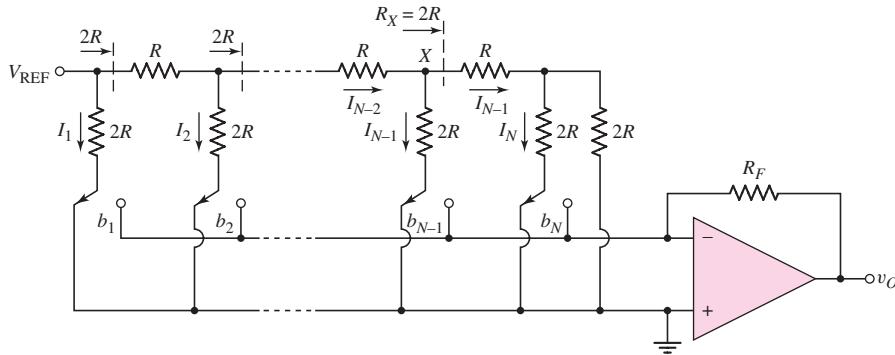


Figure 16.92 Example of an $R-2R$ ladder network in an N -bit D/A converter

The circuit in Figure 16.92 requires only two resistance values. These resistance values can then be held to tight tolerances.

There are a variety of D/A converter designs that will not be pursued in this text. This short discussion provides a brief introduction to D/A design.

16.11.3 Analog-to-Digital Converters

As in the previous section, we will consider a few basic A/D converters to gain an appreciation of the techniques used in these circuits.

Parallel or Flash A/D

The fastest A/D converter, and perhaps simplest in concept, is the parallel A/D or flash converter. Figure 16.93 shows a 3-bit flash A/D converter. The analog input signal v_A is applied to seven comparators at the noninverting terminals. A reference voltage is applied to a resistive ladder network. The outputs of the ladder network are applied to the inverting terminals of the comparators.

The total resistance in the ladder network is $8R$ so $V_{\text{REF}}/8R$ represents 1 LSB in terms of current. The smallest output voltage is

$$v_1 = \frac{V_{\text{REF}}}{8R} \left(\frac{R}{2} \right) = \frac{V_{\text{REF}}}{16} \quad (16.89)$$

which represents $\frac{1}{2}$ LSB. The second output voltage is

$$v_2 = \frac{V_{\text{REF}}}{8R} \left(\frac{3R}{2} \right) = 3 \left(\frac{V_{\text{REF}}}{16} \right) \quad (16.90)$$

which represents $1\frac{1}{2}$ LSB.

If the analog input is $v_A < \frac{1}{2}$ LSB, the output of all comparators will be low. If the analog input is $\frac{1}{2}$ LSB $< v_A < 1\frac{1}{2}$ LSB, the output of the first comparator goes high. As the analog input voltage increases, the outputs of additional comparators go high. The combinational logic network then produces the desired 3-bit output word. We can note that a complete conversion is obtained during one clock period.

A disadvantage of the flash A/D converter is that the number of resistors and comparators increases rapidly as the desired number of output bits increases. We see that 2^N resistors and $2^N - 1$ comparators are required. Thus, for a 10 bit word,

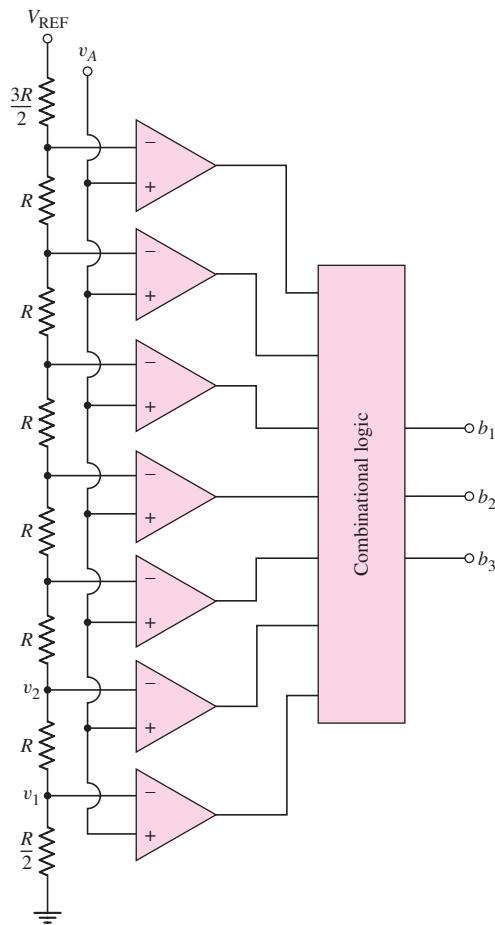


Figure 16.93 A 3-bit flash or parallel A/D converter

1024 resistors and 1023 comparators are required. However, 10-bit resolution A/D flash converters have been fabricated as ICs.

Counting A/D

A second type of A/D converter is the counting converter. This system contains a comparator, a counter, and a D/A converter in a feedback configuration as shown in Figure 16.94(a). Additional control circuitry is not shown for simplicity.

Initially the output of the counter is set equal to zero and the output of the D/A converter is set to $v_O = \frac{1}{2}$ LSB. When an analog input voltage v_A is applied, the output of the comparator is high (unless $v_A < \frac{1}{2}$ LSB), which enables the counter. Then for each clock pulse, the output of the counter increases by one, producing an N -bit digital output. When the output of the D/A becomes just greater than the analog input voltage, the output of the comparator goes low and the counter is disabled. The N -bit digital output then corresponds to the analog input signal.

Figure 16.94(b) shows the timing diagram of a counting converter for a 4-bit digital output. Assume that the analog input signal is in the range $0 \leq v_A \leq 5$ V. A 1 LSB then corresponds to $\frac{5}{16}$ V.

Assume the analog input signal is $v_A = 5.2(\frac{5}{16})$ V. The initial output of the D/A, as mentioned, is a $\frac{1}{2}$ LSB offset voltage. By including the offset voltage, the maximum

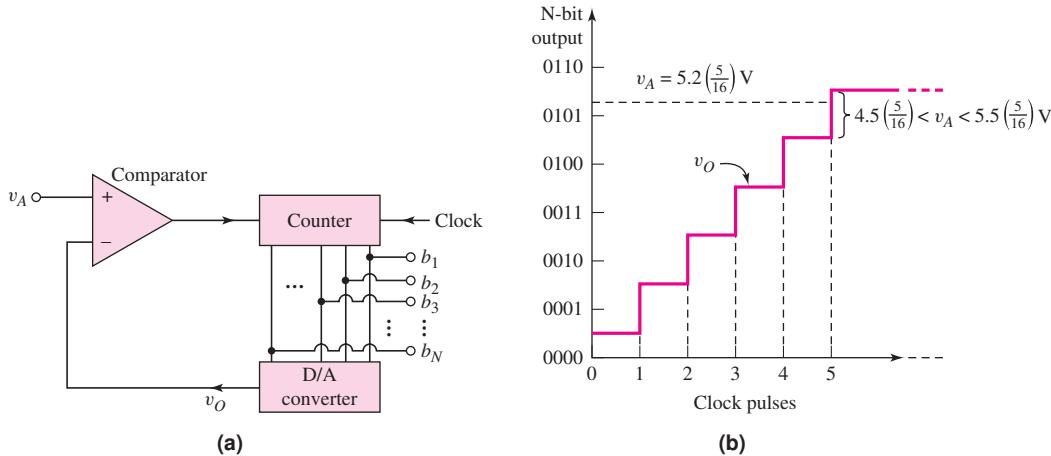


Figure 16.94 (a) Block diagram of a counting A/D converter and (b) the timing diagram of a 4-bit A/D counting converter for a specific input voltage

quantization error will then be $\pm\frac{1}{2}$ LSB. We see from Figure 16.94(b) that after the fifth clock pulse, the output of the D/A is

$$v_O = \frac{1}{2}\left(\frac{5}{16}\right) + 5\left(\frac{5}{16}\right) = 5.5\left(\frac{5}{16}\right)V \quad (16.91)$$

which is larger than v_A . The counter then stops counting and the digital output is 0101. We can note that the digital output corresponds to $5\left(\frac{5}{16}\right)V$, which is within $\frac{1}{2}$ LSB of the analog input signal.

To complete the conversion process, the clock must go through its complete cycle, which for a 4-bit output is 16 clock periods.

Dual-Slope A/D

Another type of A/D conversion scheme is the dual-slope A/D converter shown in Figure 16.95(a). This type of converter is found in high-resolution data acquisition systems, for example, since 20-bit conversions can be achieved.

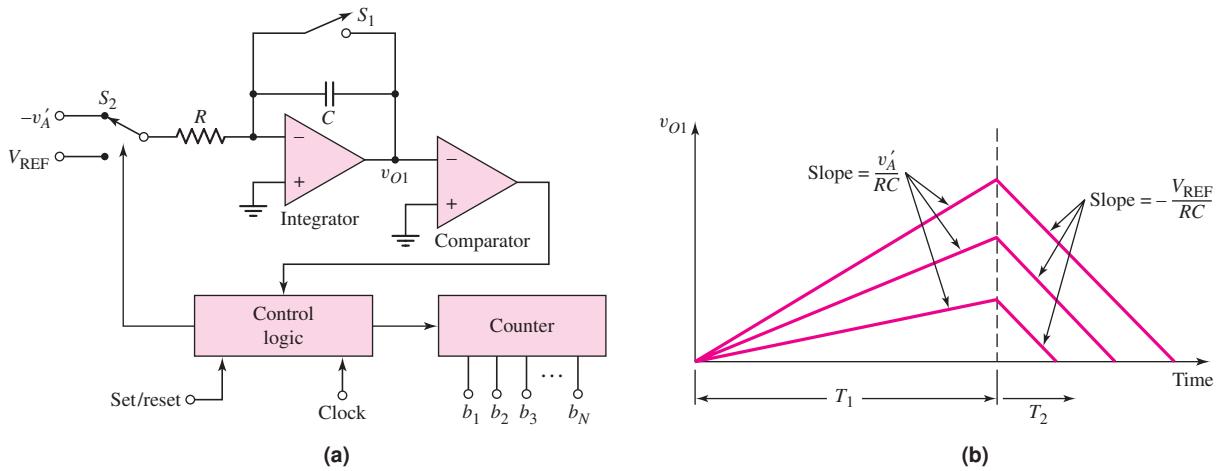


Figure 16.95 (a) Block diagram and (b) timing diagram of a dual-slope A/D converter

From Figure 16.95(a), at $t = 0$, the reset switch S_1 opens and a negative input signal ($-v'_A$) is applied to the integrator. The input signal v'_A is a sampled portion of the analog signal v_A and hence is a constant during the conversion process. The output v_{O1} of the integrator is a positive linear signal as shown in the timing diagram in Figure 16.95(b). The slope of the signal is proportional to the value of v'_A . This portion of the conversion process continues for a fixed time T_1 , at which time the counter has reached its maximum value and overflows.

At this time, the input switch S_2 changes to a positive input reference voltage V_{REF} . The output of the integrator starts at the peak output voltage reached at T_1 and now has a negative slope. The counter has been reset and is now counting. The counting stops when the output voltage v_{O1} reaches zero.

The time T_2 is related to T_1 and v'_A by

$$T_2 = T_1 \left(\frac{v'_A}{V_{\text{REF}}} \right) \quad (16.92)$$

The counter reading at T_2 is given by

$$n = 2^N \left(\frac{v'_A}{V_{\text{REF}}} \right) \quad (16.93)$$

The output of the counter is then the digital equivalent of v'_A .

The output of the dual-slope A/D converter is independent of the actual values of R and C and hence is very accurate. The disadvantage of this data converter is that it is a fairly slow system. The time T_1 requires 2^N clock pulses and the maximum possible time T_2 would also require 2^N clock pulses. For example, a 12-bit A/D converter would require a total of 8192 clock pulses. This corresponds to a conversion time of 8.2 ms for a 1 MHz clock.



16.12 DESIGN APPLICATION: A STATIC CMOS LOGIC GATE

Objective: • Design a static CMOS logic gate to implement a specific logic function.

Specifications: A static CMOS logic gate is to be designed that implements the function of a three-input odd-parity checker. The output is to be high when an odd number of inputs are high. The size of each transistor is to be determined so that the switching speed is the same as that of a basic CMOS inverter with $W_n = W$ and $W_p = 2W$. A minimum number of transistors are to be used in the NMOS pull-down and PMOS pull-up portions of the circuit.

Choices: We will assume that input signals A , B , and C as well as the complements \bar{A} , \bar{B} , and \bar{C} are available.

Solution (Logic Function): The output of the logic gate is to be high when one input is high or when all three inputs are high. The output is to be high, for example, when the inputs are $A = 1$ and $B = C = 0$. The output would be high, then, for $A\bar{B}\bar{C} = 1$. Considering the other possibilities, the logic function can be written as

$$F = A\bar{B}\bar{C} + \bar{A}B\bar{C} + \bar{A}\bar{B}C + ABC \quad (16.94)$$

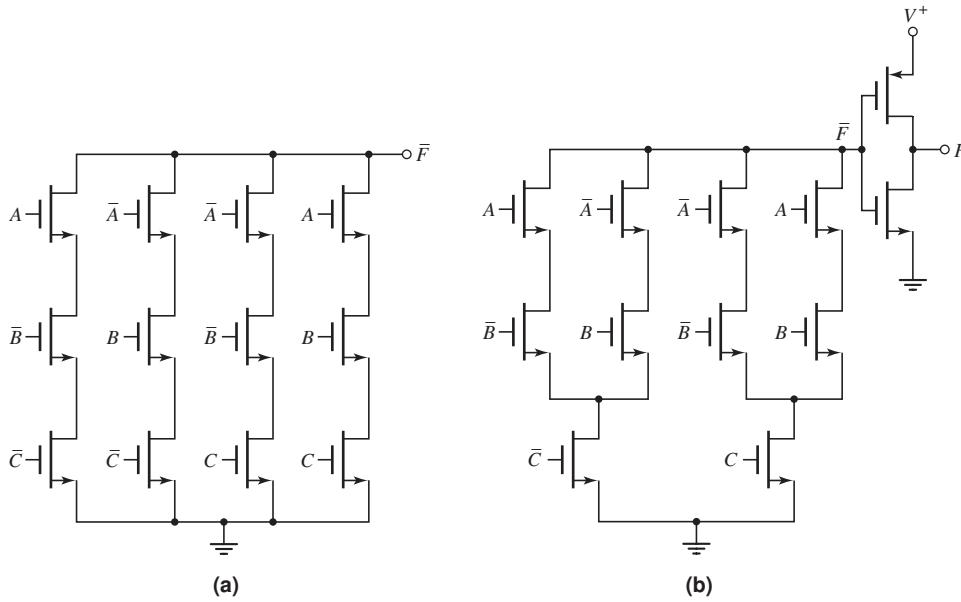


Figure 16.96 (a) The basic NMOS pull-down portion of the logic gate derived from the logic function; (b) the modified NMOS pull-down portion of the logic gate for the design application

Solution (NMOS Pull-Down): Figure 16.96(a) shows the basic NMOS pull-down portion of the logic gate derived from the logic function given in Equation (16.94). However, we may note that the two transistors at the bottom of the first two columns have a common input of \bar{C} and the two transistors at the bottom of the last two columns have a common input of C . The two transistors with common inputs can be combined into a single transistor. The final design of the NMOS pull-down portion of the logic gate is shown in Figure 16.96(b).

In order for the NMOS portion of the circuit to be in the pull-down mode, three NMOS devices in series must be turned on. In order for this circuit to be equivalent to the NMOS in the CMOS inverter, each NMOS device must have a width of $W_n = 3W$.

Solution (PMOS Pull-Up): Figure 16.97(a) shows the basic PMOS pull-up portion of the logic gate. This circuit is the complement of the NMOS circuit shown in Figure 16.96(a). We may note that two transistors on the right side of the circuit have common inputs C and \bar{C} . Each pair of transistors is effectively in series and hence can be replaced by a single transistor. The resulting circuit is shown in Figure 16.97(b). The complete three-input odd-parity checker circuit is then the combination of Figures 16.96(b) and 16.97(b) along with a CMOS inverter on the output.

Comment: The basic logic circuit can be derived from the logic function. However, as we have seen, some simplifications can be made in the design. These simplifications can also be obtained from simplifications in the basic logic function also.

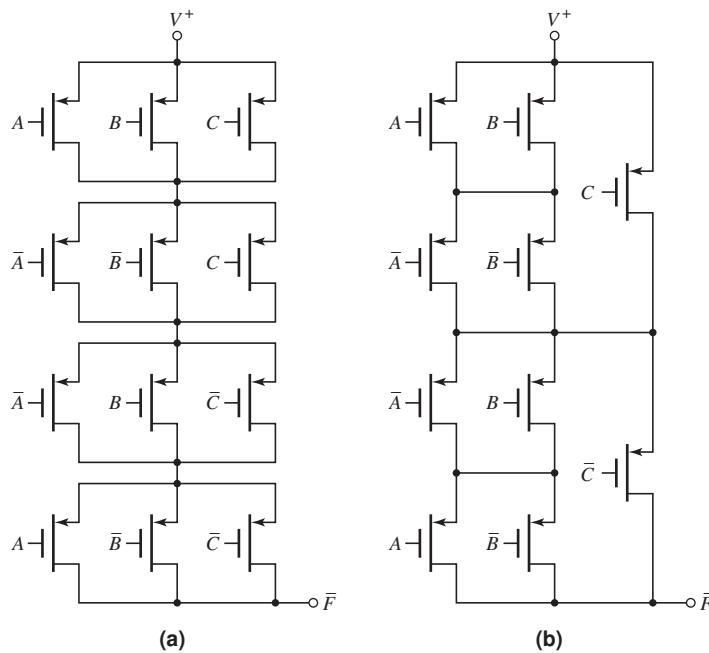


Figure 16.97 (a) The basic PMOS pull-up portion of the logic gate derived as the complement of the basic NMOS pull-down circuit; (b) the modified PMOS pull-up portion of the logic gate for the design application

16.13 SUMMARY

- In this chapter, NMOS and CMOS digital logic circuits were analyzed and designed. These circuits include basic logic gates, shift registers, flip-flops, and memories.
- The NMOS inverter is the basis of NMOS logic circuits. The quasi-static voltage transfer characteristics of NMOS inverters with resistive-load, enhancement-load, and depletion-load devices were generated.
- The basic NMOS NOR and NAND logic gates were analyzed. More sophisticated logic functions can be implemented by combining driver transistors in parallel and series combinations.
- The CMOS inverter is the basis of CMOS logic circuits. The quasi-static voltage transfer characteristics were generated. For the CMOS circuit, the quiescent power dissipation is essentially zero when the input is in either logic state.
- The basic CMOS NOR and NAND logic gates were analyzed. More sophisticated CMOS logic circuits were also analyzed and designed. Transistor width-to-length ratios were designed to provide equal current drive in the NMOS pull-down and PMOS pull-up portions of the circuit.
- CMOS clocked logic circuits can be designed to reduce the number of required PMOS devices. A generalized NMOS logic circuit is inserted between clocked PMOS and NMOS devices. The advantage of low static power dissipation is maintained.
- Sequential logic circuits, such as shift registers, flip-flops, and a full one-bit adder, were analyzed.

- A whole classification of circuits called memories was considered. In the NMOS SRAM, static power is continuously dissipated in the cell, which limits the size of the memory because of the total chip power limitation. The primary advantage of a CMOS SRAM is that there is essentially no static power dissipation. The size of a CMOS memory is limited primarily by chip area requirements.
- Read-only-memory (ROM and PROM) contains fixed data that are implemented by the manufacturer or by the user. In both cases, the data cannot be altered. EEPROM and EEPROM cells contain MOSFETs with floating gates that can be either charged or left uncharged by the user depending on whether logic 1 or a logic 0 is to be stored.
- The basic concepts used in A/D and D/A converters were discussed. A few examples of D/A converter circuits and A/D converter systems were analyzed.
- As an application, a static CMOS logic circuit to implement a specific logic function is designed.

CHECKPOINT

After studying this chapter, the reader should have the ability to:

- ✓ Analyze the transfer characteristics of NMOS inverters, including the determination of noise margins.
- ✓ Design an NMOS logic circuit to perform a specific logic function.
- ✓ Analyze the transfer characteristics of the CMOS inverter, including the determination of switching power and noise margins.
- ✓ Design a CMOS logic circuit to perform a specific logic function.
- ✓ Design a clocked CMOS logic circuit to perform a specific logic function.
- ✓ Design an NMOS or CMOS pass network to perform a specific logic function.
- ✓ Design an NMOS or CMOS RAM cell and design a simple sense amplifier.
- ✓ Analyze the $R-2R$ ladder network used in a D/A converter circuit.
- ✓ Describe the characteristics of a 3-bit flash A/D converter and describe the operation of the dual-slope A/D converter.

REVIEW QUESTIONS

1. Sketch the quasi-static voltage transfer characteristics of an NMOS inverter with depletion load. What effect does changing the transistor W/L ratio have on the transfer characteristics?
2. Sketch an NMOS three-input NOR logic gate. Describe its operation. Discuss the condition under which the maximum logic 0 value is obtained.
3. Discuss how more sophisticated (compared to the basic NOR and NAND) logic functions can be implemented in a single NMOS logic circuit.
4. Sketch the quasi-static voltage transfer characteristics of a CMOS inverter. Discuss the various intervals in terms of transistor bias. What is the effect on the transfer curve of changing the transistor W/L ratios? What is the advantage of the CMOS inverter compared to an NMOS inverter?
5. Discuss the parameters that affect the switching power dissipation in a CMOS inverter.

6. Define the noise margin in a CMOS inverter.
7. Sketch a CMOS three-input NAND logic gate. Describe its operation. Determine the relative transistor W/L ratios to obtain equal pull-up and pull-down switching times.
8. Discuss how more sophisticated (compared to the basic NOR and NAND) logic functions can be implemented in a single CMOS logic circuit.
9. Discuss the basic principles of a clocked CMOS logic circuit.
10. Sketch an NMOS transmission gate and describe its operation. What is the maximum output voltage?
11. Sketch a CMOS transmission gate and describe its operation. Why is the quasi-static output voltage always equal to the quasi-static input voltage?
12. Discuss what is meant by pass transistor logic.
13. If an NMOS or CMOS transmission gate is turned off (an open switch), discuss why the output voltage is, in general, not stable.
14. Sketch an NMOS dynamic shift register and describe its operation.
15. Sketch a CMOS R-S flip flop and describe its operation. Why must the input condition $R = S = 1$ be avoided?
16. Describe the basic architecture of a semiconductor random-access memory.
17. Sketch a CMOS SRAM cell and describe its operation. Discuss any advantages and disadvantages of this design. Describe how the cell is addressed.
18. Sketch a one-transistor DRAM cell and describe its operation. What makes this circuit dynamic?
19. Describe a mask-programmed MOSFET ROM memory.
20. Describe the basic operation of a floating gate MOSFET and how this can be used in an erasable ROM.

PROBLEMS

[Note: In the following problems, unless otherwise stated, assume transistor parameters of $V_{TNO} = 0.5$ V, $V_{TPO} = -0.5$ V, $k'_n = 100 \mu\text{A/V}^2$, and $k'_p = 40 \mu\text{A/V}^2$. Neglect the body effect unless otherwise stated and assume $T = 300$ K.]

Section 16.1 NMOS Inverters

- 16.1 The load resistor in the NMOS inverter in Figure 16.3(a) is $R_D = 40 \text{ k}\Omega$. The circuit is biased at $V_{DD} = 3.3$ V. (a) Design the transistor width-to-length ratio such that $v_O = 0.1$ V when $v_I = 3.3$ V. (b) Using the results of part (a), determine the transition point for the transistor. (c) Using the results of part (a), find the maximum current and maximum power dissipation in the inverter.
- 16.2 The inverter circuit in Figure 16.3(a) is biased at $V_{DD} = 3.3$ V. Assume the transistor conduction parameter is $K_n = 50 \mu\text{A/V}^2$. (a) Let $R_D = 100 \text{ k}\Omega$. (i) Determine the transition point. (ii) Determine v_O for $v_I = 3.3$ V. (b) Repeat part (a) for $R_D = 30 \text{ k}\Omega$. (c) Repeat part (a) for $R_D = 5 \text{ k}\Omega$.
- D16.3 (a) Redesign the resistive load inverter in Figure 16.3(a) so that the maximum power dissipation is 0.25 mW with $V_{DD} = 3.3$ V and

- $v_O = 0.15$ V when the input is a logic 1. (b) Using the results of part (a), what is the input voltage range when the transistor is biased in the saturation region?
- D16.4 (a) Design the saturated load inverter circuit in Figure 16.5(a) such that the power dissipation is 0.30 mW and the output voltage is 0.08 V for $v_I = 1.4$ V. The circuit is biased at $V_{DD} = 1.8$ V and the transistor threshold voltage of each transistor is $V_{TNO} = 0.4$ V. (b) Using the results of part (a), find the range of input voltage such that the driver transistor is biased in the saturation region.
- 16.5 An NMOS inverter with saturated load is shown in Figure 16.5(a). The bias is $V_{DD} = 3$ V and the transistor threshold voltages are 0.5 V. (a) Find the ratio K_D/K_L such that $v_O = 0.25$ V when $v_I = 3$ V. (b) Repeat part (a) for $v_I = 2.5$ V. (c) If $W/L = 1$ for the load transistor, determine the power dissipation in the inverter for parts (a) and (b).
- D16.6 Consider the NMOS inverter with saturated load in Figure 16.5(a). Let $V_{DD} = 3$ V. (a) Design the circuit such that the power dissipation in the circuit is $400 \mu\text{W}$ and the output voltage is 0.10 V when the input voltage is a logic 1. (b) Determine the transition point of the driver transistor.
- 16.7 The NMOS inverter with saturated load in Figure 16.5(a) operates with a supply voltage of V_{DD} . The MOSFETs have threshold voltages of $V_{TN} = 0.2 V_{DD}$. Determine $(W/L)_D/(W/L)_L$ such that $v_O = 0.08 V_{DD}$. Neglect the body effect.
- 16.8 The enhancement-load transistor in the NMOS inverter in Figure P16.8 has a separate bias applied to the gate. Assume transistor parameters of $K_n = 1 \text{ mA/V}^2$ for M_D , $K_n = 0.4 \text{ mA/V}^2$ for M_L , and $V_{TN} = 1$ V for both transistors. Using the appropriate logic 0 and logic 1 input voltages, determine V_{OH} and V_{OL} for: (a) $v_B = 4$ V, (b) $v_B = 5$ V, (c) $v_B = 6$ V, and (d) $v_B = 7$ V.
- 16.9 For the depletion load inverter shown in Figure 16.7(a), assume parameters of $V_{DD} = 3.3$ V, $V_{TND} = 0.5$ V, $V_{TNL} = -0.8$ V, $K_D = 500 \mu\text{A/V}^2$, and $K_L = 100 \mu\text{A/V}^2$. (a) Find the transition points of the driver and load transistors. (b) Determine v_O for $v_I = 3.3$ V. (c) Determine the maximum current and maximum power dissipation in the circuit.
- 16.10 In the depletion-load NMOS inverter circuit in Figure 16.7(a), let $V_{TND} = 0.5$ V and $V_{DD} = 3$ V, $K_L = 50 \mu\text{A/V}^2$, and $K_D = 500 \mu\text{A/V}^2$. Calculate the value of V_{TNL} such that $v_O = 0.10$ V when $v_I = 3$ V.
- D16.11 Consider the NMOS inverter with depletion load in Figure 16.7(a). Let $V_{DD} = 1.8$ V, and assume $V_{TND} = 0.3$ V and $V_{TNL} = -0.6$ V. (a) Design the circuit such that the power dissipation is $80 \mu\text{W}$ and the output voltage is $v_O = 0.06$ V when v_I is a logic 1. (b) Using the results of part (a), determine the transition points for the driver and load transistors. (c) If $(W/L)_D$ found in part (a) is doubled, what is the maximum power dissipation in the inverter and what is v_O when v_I is a logic 1?
- D16.12 The NMOS inverter with depletion load is shown in Figure 16.7(a). The bias is $V_{DD} = 2.5$ V. The transistor parameters are $V_{TND} = 0.5$ V and $V_{TNL} = -1$ V. The width-to-length ratio of the load device is $W/L = 1$. (a) Design the driver transistor such that $v_O = 0.05$ V when the input is a logic 1. (b) What is the power dissipated in the circuit when $v_I = 2.5$ V?

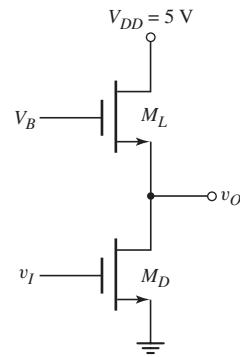


Figure P16.8

- 16.13 Calculate the power dissipated in each inverter circuit in Figure P16.13 for the following input conditions: (a) Inverter a: (i) $v_I = 0.5 \text{ V}$, (ii) $v_I = 5 \text{ V}$; (b) Inverter b: (i) $v_I = 0.25 \text{ V}$, (ii) $v_I = 4.3 \text{ V}$; (c) Inverter c: (i) $v_I = 0.03 \text{ V}$, (ii) $v_I = 5 \text{ V}$.

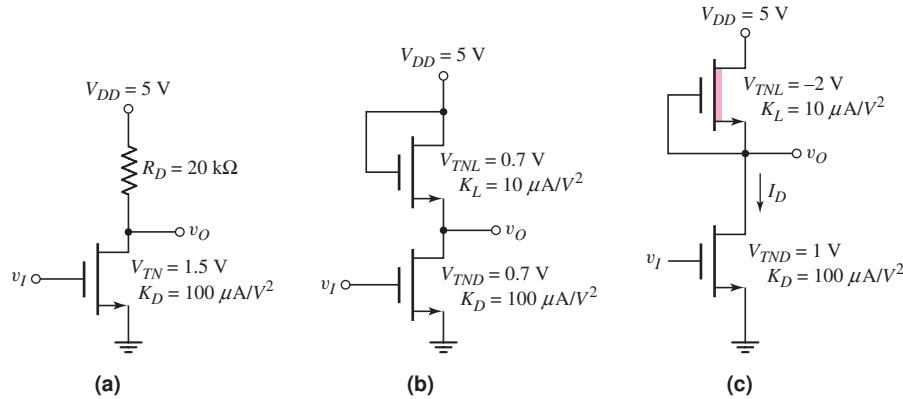


Figure P16.13

- 16.14 For the two inverters in Figure P16.14, assume $(W/L)_L = 1$ for the load devices and $(W/L)_D = 10$ for the driver devices. (a) If v_I is a logic 1, determine the values of v_{O1} and v_{O2} . (b) Repeat part (a) if v_I is a logic 0.

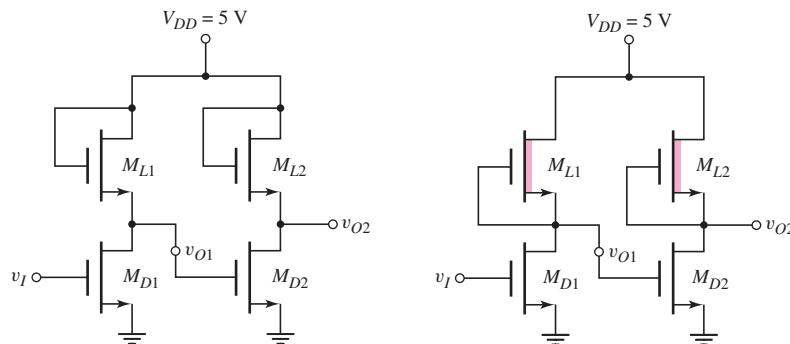


Figure P16.14

Figure P16.15

- 16.15 Consider the circuit in Figure P16.15. The parameters of the driver transistors are $V_{TND} = 0.8 \text{ V}$ and $(W/L)_D = 4$, and those of the load transistors are $V_{TNL} = -1.2 \text{ V}$ and $(W/L)_L = 1$. (a) If v_I is a logic 1, determine the values of v_{O1} and v_{O2} . (b) Repeat part (a) if v_I is a logic 0.
- 16.16 For the saturated load inverter shown in Figure 16.9(a), assume transistor parameters of $V_{TNDO} = V_{TNLO} = 0.5 \text{ V}$, $K_D = 200 \mu\text{A}/\text{V}^2$, $K_L = 20 \mu\text{A}/\text{V}^2$, $\gamma = 0.25 \text{ V}^{1/2}$, and $\phi_{fp} = 0.35 \text{ V}$. Determine v_O for $v_I = 0.12 \text{ V}$ for the case when (a) the body effect is neglected and (b) the body effect is taken into account.
- 16.17 Consider the NMOS inverter with depletion load in Figure 16.9(b). The transistor parameters are $V_{TNDO} = 0.4 \text{ V}$, $V_{TNLO} = -0.6 \text{ V}$, $K_D = 100 \mu\text{A}/\text{V}^2$, $K_L = 20 \mu\text{A}/\text{V}^2$, $\gamma = 0.25 \text{ V}^{1/2}$, and $\phi_{fp} = 0.35 \text{ V}$. Determine v_I when $v_O = 1.25 \text{ V}$ for the case when (a) the body effect is neglected and (b) the body effect is taken into account.

Section 16.2 NMOS Logic Circuits

- 16.18 Consider the circuit with a depletion load device shown in Figure P16.18.
- (a) For $v_X = 1.8 \text{ V}$ and $v_Y = 0.1 \text{ V}$, determine K_D/K_L such that $v_O = 0.1 \text{ V}$. (b) Using the results of part (a), determine v_O when $v_X = v_Y = 1.8 \text{ V}$. (c) If the width-to-length ratio of the depletion load is $(W/L)_L = 1$, determine the power dissipation in the logic circuit for the input conditions listed in parts (a) and (b).

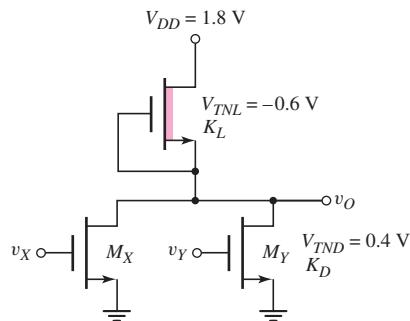


Figure P16.18

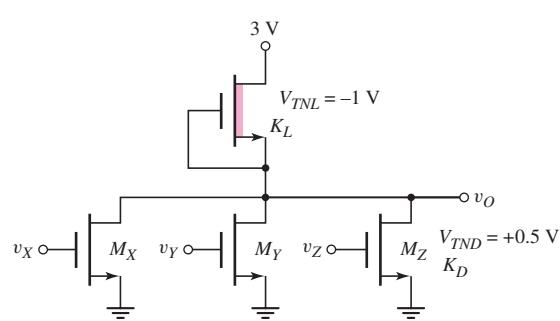


Figure P16.19

- D16.19 Consider the three-input NOR logic gate in Figure P16.19. The transistor parameters are $V_{T_{NL}} = -1 \text{ V}$ and $V_{T_{ND}} = 0.5 \text{ V}$. The maximum value of v_O in its low state is to be 0.1 V. (a) Determine K_D/K_L . (b) The maximum power dissipation in the NOR logic gate is to be 0.1 mW. Determine the width-to-length ratios of the transistors. (c) Determine v_O when $v_X = v_Y = v_Z = 3 \text{ V}$.
- 16.20 Consider a four-input NMOS NOR logic gate with a depletion load similar to the circuit in Figure P16.19. Assume $V_{DD} = 2.5 \text{ V}$, $V_{T_{ND}} = 0.4 \text{ V}$, and $V_{T_{NL}} = -0.6 \text{ V}$. The maximum value of v_O in its low state is to be 50 mV. (a) Determine K_D/K_L . (b) The maximum power dissipation in this NOR logic gate is to be $50 \mu\text{W}$. Determine the width-to-length ratio of each transistor. (c) Determine v_O when (i) two inputs are a logic 1, (ii) three inputs are a logic 1, and (iii) all inputs are a logic 1.
- 16.21 The transistor parameters for the circuit in Figure P16.21 are: $V_{TN} = 0.8 \text{ V}$ for all enhancement-mode devices, $V_{TN} = -2 \text{ V}$ for the depletion-mode

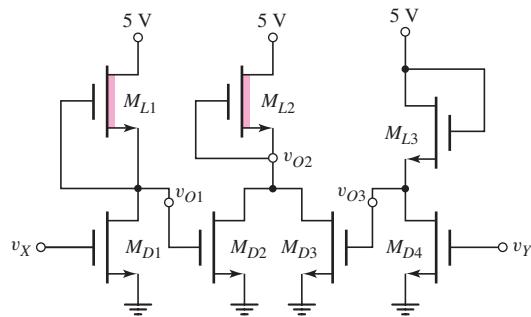


Figure P16.21

devices, and $k'_n = 60 \mu\text{A/V}^2$ for all devices. The width-to-length ratios of M_{L2} and M_{L3} are 1, and those for M_{D2} , M_{D3} , and M_{D4} are 8. (a) For $v_X = 5 \text{ V}$, output v_{O1} is 0.15 V, and the power dissipation in this inverter is to be no more than $250 \mu\text{W}$. Determine $(W/L)_{ML1}$ and $(W/L)_{MD1}$. (b) For $v_X = v_Y = 0$, determine v_{O2} .

- 16.22 Consider the NMOS circuit in Figure P16.22. The transistor parameters are $(W/L)_X = (W/L)_Y = 12$, $(W/L)_L = 1$, and $V_{TN} = 0.4 \text{ V}$. Neglect the body effect. (a) Determine v_O when $v_X = v_Y = 2.9 \text{ V}$. (b) What are the values of v_{GSX} , v_{GSY} , v_{DSX} , and v_{DSY} ? [Hint: Set the drain current in each device equal to each other. Also, neglect the terms v_O^2 , v_{DSX}^2 , and v_{DSY}^2].

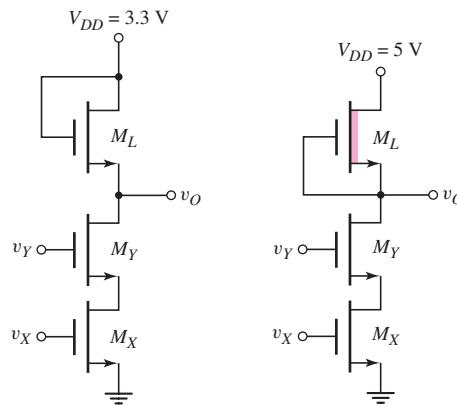


Figure P16.22

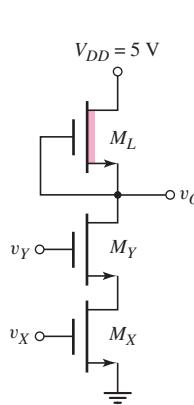


Figure P16.23

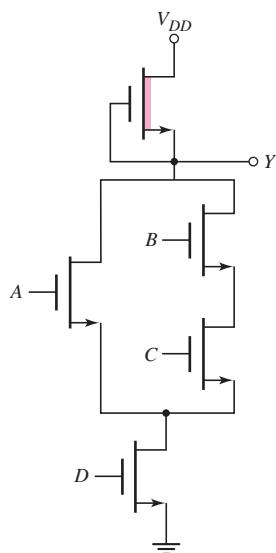


Figure P16.25

- 16.23 In the NMOS circuit in Figure P16.23, the transistor parameters are: $(W/L)_X = (W/L)_Y = 4$, $(W/L)_L = 1$, $V_{TNX} = V_{TNY} = 0.8 \text{ V}$, and $V_{TNL} = -1.5 \text{ V}$. (a) Determine v_O when $v_X = v_Y = 5 \text{ V}$. (b) What are the values of v_{GSX} , v_{GSY} , v_{DSX} , and v_{DSY} ? Repeat part (a) for $\gamma = 0.5$.
- 16.24 Consider a four-input NMOS NAND logic gate with a depletion load similar to the circuit in Figure P16.23. The bias voltage is $V_{DD} = 3.3 \text{ V}$, and the threshold voltages are $V_{TND} = 0.4 \text{ V}$ and $V_{TNL} = -0.6 \text{ V}$. The logic 0 output voltage is to be 0.10 V. (a) Using approximation methods, determine K_D/K_L . (b) The maximum power dissipation in the circuit is to be $100 \mu\text{W}$. Determine $(W/L)_L$ and $(W/L)_D$.
- 16.25 Determine the logic function implemented by the circuit in Figure P16.25.
- 16.26 Find the logic function implemented by the circuit in Figure P16.26.
- 16.27 What is the logic function implemented by the circuit in Figure P16.27?
- D16.28 The Boolean function for a carry-out signal of a one-bit full adder is given by

$$\text{Carry-out} = A \cdot B + A \cdot C + B \cdot C$$

- (a) Design an NMOS logic circuit with depletion load to perform this function. Signals A , B , and C are available. (b) Assume $(W/L)_L = 1$, $V_{DD} = 5 \text{ V}$, $V_{TNL} = -1.5 \text{ V}$, and $V_{TND} = 0.8 \text{ V}$. Determine the W/L ratio of the other transistors such that the maximum logic 0 value in any part of the circuit is 0.2 V.

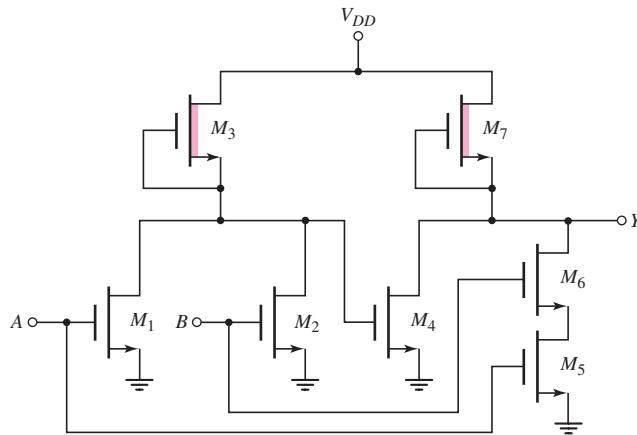


Figure P16.26

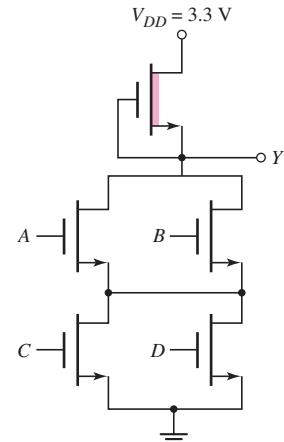


Figure P16.27

- D16.29 (a) Design an NMOS depletion-load logic gate that implements the function $\bar{Y} = [A + B \cdot (C + D)]$. (b) Assume $V_{DD} = 2.5$ V, $(W/L)_L = 1$, $V_{TND} = 0.4$ V, and $V_{TNL} = -0.6$ V. Determine $(W/L)_D$ of each transistor such that $V_{OL}(\max) = 50$ mV.
- D16.30 Design an NMOS logic circuit with a depletion load that will sound an alarm in an automobile if the ignition is turned off while the headlights are still on and/or the parking brake has not been set. Separate indicator lights are also to be included showing whether the headlights are on or the parking brake needs to be set. State any assumptions that are made.

Section 16.3 CMOS Inverter

- 16.31 Consider the CMOS inverter in Figure 16.21 biased at $V_{DD} = 2.5$ V. The transistor parameters are $V_{TN} = 0.4$ V, $V_{TP} = -0.4$ V, and $K_n = K_p = 100 \mu\text{A/V}^2$. (a) Find the transition points for the p-channel and n-channel transistors. (b) Sketch the voltage transfer characteristics, including the appropriate voltage values at the transition points. (c) Determine v_O for $v_I = 1.1$ V and $v_I = 1.4$ V.
- 16.32 For the CMOS inverter in Figure 16.21, let $V_{DD} = 3.3$ V, $k'_n = 100 \mu\text{A/V}^2$, $k'_p = 40 \mu\text{A/V}^2$, $V_{TN} = 0.4$ V, and $V_{TP} = -0.4$ V. (a) Let $(W/L)_n = 2$ and $(W/L)_p = 5$. (i) Find the transition points for the p-channel and n-channel transistors. (ii) Sketch the voltage transfer characteristics including the appropriate voltage values at the transition points. (iii) Find v_I for $v_O = 0.25$ V and $v_O = 3.05$ V. (b) Repeat part (a) for $(W/L)_n = 4$ and $(W/L)_p = 5$.
- 16.33 (a) For the CMOS inverter in Figure 16.21 in the text, let $V_{DD} = 3.3$ V, $V_{TN} = +0.4$ V, and $V_{TP} = -0.4$ V. Assume $(W/L)_n = 4$ and $(W/L)_p = 12$. Determine (i) the input switching voltage, (ii) the input voltage when $v_O = 3.1$ V, and (iii) the input voltage when $v_O = 0.2$ V. (b) Repeat part (a) for $(W/L)_n = 6$ and $(W/L)_p = 4$.

- 16.34 Consider the CMOS inverter pair in Figure P16.34. Let $V_{TN} = 0.8$ V, $V_{TP} = -0.8$ V, and $K_n = K_p$. (a) If $v_{O1} = 0.6$ V, determine v_I and v_{O2} . (b) Determine the range of v_{O2} for which both N_2 and P_2 are biased in the saturation region.

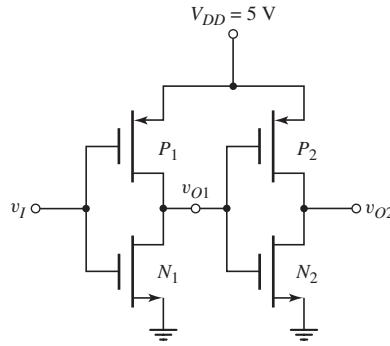


Figure P16.34

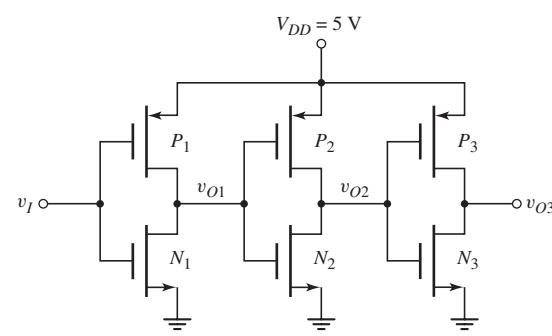


Figure P16.35

- 16.35 Consider the series of CMOS inverters in Figure P16.35. The threshold voltages of the n-channel transistors are $V_{TN} = 0.8$ V, and the threshold voltages of the p-channel transistors are $V_{TP} = -0.8$ V. The conduction parameters are all equal. (a) Determine the range of v_{O1} for which both N_1 and P_1 are biased in the saturation region. (b) If $v_{O2} = 0.6$ V, determine the values of v_{O3} , v_{O1} , and v_I .
- 16.36 (a) A CMOS inverter is biased at $V_{DD} = 2.5$ V. The transistor parameters are $K_n = K_p = 120 \mu\text{A/V}^2$, $V_{TN} = 0.4$ V, and $V_{TP} = -0.4$ V. Calculate and plot the current in the transistors as a function of the input voltage for $0 \leq v_I \leq 2.5$ V. (b) Repeat part (a) for $V_{DD} = 1.8$ V and $0 \leq v_I \leq 1.8$ V.
- 16.37 The transistor parameters in the CMOS inverter are $V_{TN} = 0.35$ V, $V_{TP} = -0.35$ V, $k'_n = 80 \mu\text{A/V}^2$, and $k'_p = 40 \mu\text{A/V}^2$. Let $V_{DD} = 1.8$ V. (a) Determine the peak current in the inverter during a switching cycle for $(W/L)_n = 2$ and $(W/L)_p = 4$. (b) Repeat part (a) for $(W/L)_n = 2$ and $(W/L)_p = 6$. (c) Repeat part (a) for $(W/L)_n = (W/L)_p = 4$.
- 16.38 A CMOS inverter is biased at $V_{DD} = 3.3$ V. The transistor threshold voltages are $V_{TN} = +0.4$ V and $V_{TP} = -0.4$ V. Determine the peak current in the inverter and the input voltage at which it occurs for (a) $(W/L)_n = 3$, $(W/L)_p = 7.5$; (b) $(W/L)_n = (W/L)_p = 4$; (c) $(W/L)_n = 3$, $(W/L)_p = 12$.
- 16.39 A load capacitor of 0.2 pF is connected to the output of a CMOS inverter. Determine the power dissipated in the CMOS inverter for a switching frequency of 10 MHz, for inverter parameters described in (a) Problem 16.36 and (b) Problem 16.37.
- 16.40 (a) A CMOS digital logic circuit contains the equivalent of 4 million CMOS inverters and is biased at $V_{DD} = 1.8$ V. The equivalent load capacitance of each inverter is 0.12 pF and each inverter is switching at 150 MHz. Determine the total average power dissipated in the circuit. (b) If the switching frequency is doubled, but the total power dissipation is to remain the same with the same load capacitance, determine the required bias voltage.
- 16.41 A particular IC chip can dissipate 3 W and contains 10 million CMOS inverters. Each inverter is being switched at a frequency f . (a) Determine the average power that each inverter can dissipate without exceeding the total

- allowed power. (b) If the switching frequency is $f = 5$ MHz, what is the maximum capacitive load on each inverter if (i) $V_{DD} = 5$ V, (ii) $V_{DD} = 3.3$ V, and (iii) $V_{DD} = 1.5$ V.
- 16.42 Repeat Problem 16.41 for the case when the chip contains 5 million CMOS inverters being switched at $f = 8$ MHz and the total power dissipated can be 10 W.
- 16.43 Consider a CMOS inverter. (a) Show that when $v_I \cong V_{DD}$, the resistance of the NMOS device is approximately $1/[k'_n(W/L)_n(V_{DD} - V_{TN})]$, and when $v_I \cong 0$, the resistance of the PMOS device is approximately $1/[k'_p(W/L)_p(V_{DD} + V_{TP})]$. (b) Using the results of part (a), determine the maximum current that the NMOS device can sink such that the output voltage stays below 0.5 V, and determine the maximum current that the PMOS device can source such that the output voltage does not drop more than 0.5 V below V_{DD} .
- 16.44 The CMOS inverter in Figure 16.21 is biased at $V_{DD} = 3.3$ V. Let $K_n = K_p$, $V_{TN} = 0.5$ V, and $V_{TP} = -0.5$ V. (a) Determine the two values of v_I and the corresponding values of v_O for which $(dv_O/dv_I) = -1$ on the voltage transfer characteristics. (b) Find the noise margins.
- 16.45 Repeat Problem 16.44 if the circuit and transistor parameters are $V_{DD} = 2.5$ V, $V_{TN} = 0.35$ V, $V_{TP} = -0.35$ V, $K_n = 100 \mu\text{A/V}^2$, and $K_p = 50 \mu\text{A/V}^2$.
- 16.46 (a) Determine the noise margins of a CMOS inverter biased at $V_{DD} = 3.3$ V with $(W/L)_n = 2$ and $(W/L)_p = 5$. Assume $V_{TN} = 0.4$ V and $V_{TP} = -0.4$ V. (b) Repeat part (a) for $(W/L)_n = 4$ and $(W/L)_p = 12$.

Section 16.4 CMOS Logic Circuits

- 16.47 Consider the three-input CMOS NAND circuit in Figure P16.47. Assume $k'_n = 2k'_p$ and $V_{TN} = |V_{TP}| = 0.8$ V. (a) If $v_A = v_B = 5$ V, determine v_C such that both N_3 and P_3 are biased in the saturation region when $(W/L)_p = 2(W/L)_n$. (State any assumptions you make.) (b) If $v_A = v_B = v_C = v_I$, determine the relationship between $(W/L)_p$ and $(W/L)_n$ such that $v_I = 2.5$ V when all transistors are biased in the saturation region. (c) Using the results

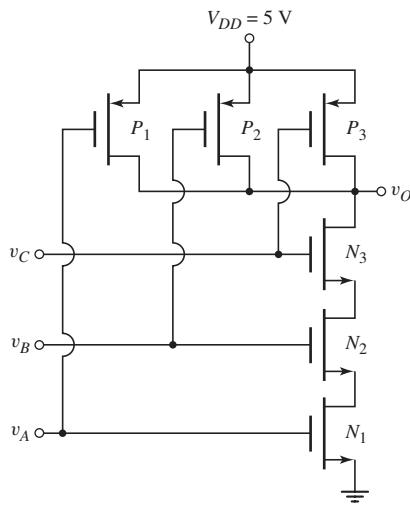


Figure P16.47

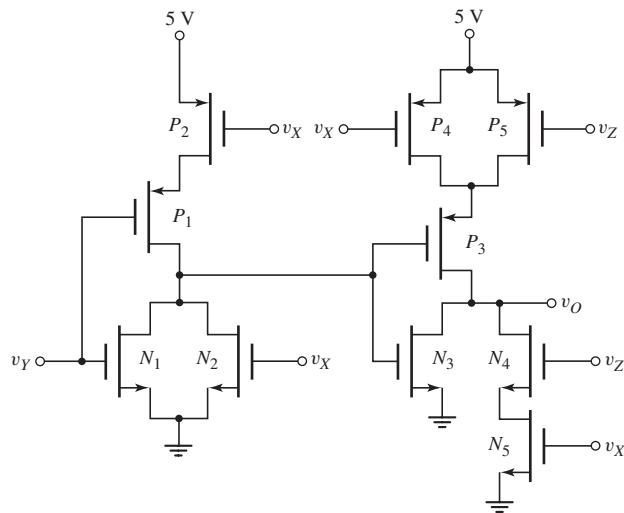


Figure P16.48

of part (b) and assuming $v_A = v_B = 5\text{ V}$, determine v_C such that both N_3 and P_3 are biased in the saturation region. (State any assumptions you make.)

- 16.48 Consider the circuit in Figure P16.48. (a) The inputs v_X , v_Y , and v_Z listed in the following table are either a logic 0 or a logic 1. These inputs are the outputs from similar-type CMOS logic circuits. The input logic conditions listed are sequential in time. State whether the transistors listed are “on” or “off,” and determine the output voltage. (b) What logic function does this circuit implement?

v_X	v_Y	v_Z	N_1	N_2	N_3	N_4	N_5	v_O
1	0	1						
0	0	1						
1	1	0						
1	1	1						

- 16.49 Consider a four-input CMOS NOR logic gate. Determine the W/L ratios of the transistors to provide for symmetrical switching based on the CMOS inverter design with $(W/L)_n = 2$ and $(W/L)_p = 4$. (b) If the load capacitance of the NOR gate doubles, determine the required W/L ratios to provide the same switching speed as the logic gate in part (a).
- 16.50 Repeat Problem 16.49 for a four-input CMOS NAND logic gate.
- 16.51 Repeat Problem 16.49 for a three-input CMOS NOR logic gate.
- 16.52 Repeat Problem 16.49 for a three-input CMOS NAND logic gate.
- D16.53 Figure P16.53 shows a classic CMOS logic circuit. (a) What is the logic function performed by the circuit? (b) Design the NMOS network. (c) Determine the transistor W/L ratios to provide symmetrical switching times at twice the switching speed of the basic CMOS inverter with $(W/L)_n = 2$ and $(W/L)_p = 4$.
- D16.54 Figure P16.54 is a classic CMOS logic gate. (a) What is the logic function performed by the circuit? (b) Design the PMOS network. (c) Determine the transistor W/L ratios to provide symmetrical switching times at twice

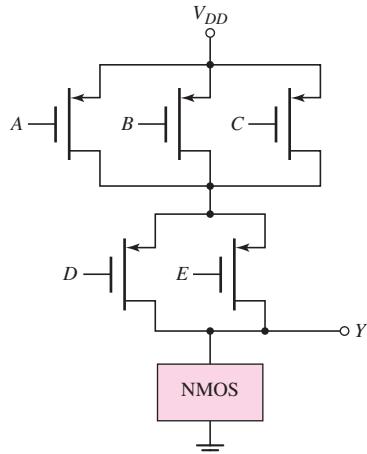


Figure P16.53

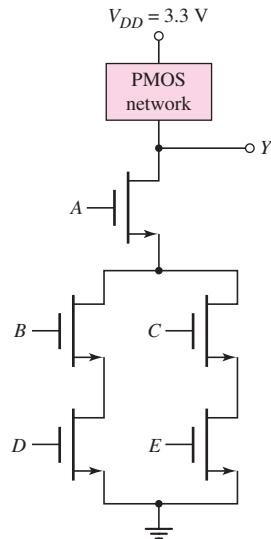


Figure P16.54

the switching speed as the basic CMOS inverter with $(W/L)_n = 2$ and $(W/L)_p = 4$.

- D16.55 Figure P16.55 is a classic CMOS logic gate. (a) What is the logic function performed by the circuit? (b) Design the NMOS network. (c) Determine the transistor W/L ratios to provide symmetrical switching times equal to the basic CMOS inverter with $(W/L)_n = 2$ and $(W/L)_p = 4$.
 D16.56 Consider the classic CMOS logic circuit in Figure P16.56. (a) What is the logic function performed by the circuit? (b) Design the PMOS network. (c) Determine the transistor W/L ratios to provide symmetrical switching times equal to the basic CMOS inverter with $(W/L)_n = 2$ and $(W/L)_p = 4$.

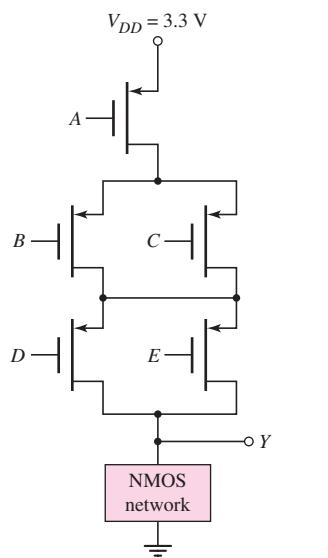


Figure P16.55

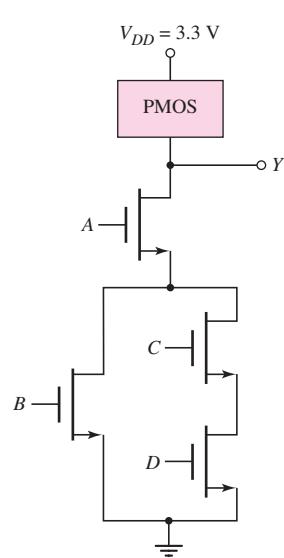


Figure P16.56

D16.57 (a) Given inputs $A, B, C, \bar{A}, \bar{B}$, and \bar{C} , design a CMOS circuit to implement the logic function $Y = A\bar{B}\bar{C} + \bar{A}\bar{B}C + A\bar{B}\bar{C}$. The design should not include a CMOS inverter at the output. (b) For $k'_n = 2k'_p$, size the transistors in the design to provide equal switching times equal to the basic CMOS inverter with $(W/L)_n = 1$ and $(W/L)_p = 2$.

D16.58 (a) Given inputs A, B, C, D , and E , design a CMOS circuit to implement the logic function $\bar{Y} = A(B + C) + D + E$. (b) Repeat part (b) of Problem 16.57 for this circuit.

16.59 (a) Determine the logic function performed by the circuit in Figure P16.59. (b) Determine the W/L ratios to provide symmetrical switching times equal to the basic CMOS inverter with $(W/L)_n = 2$ and $(W/L)_p = 4$.

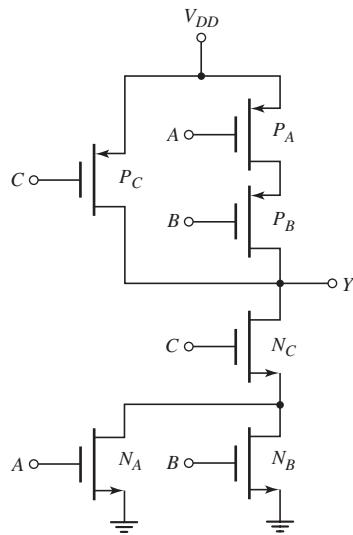


Figure P16.59

D16.60 (a) Consider a five-input CMOS NOR logic gate. Design the W/L ratios of the transistors to provide symmetrical switching times equal to the basic CMOS inverter with $(W/L)_n = 2$ and $(W/L)_p = 4$. (b) Repeat part (a) for a five-input CMOS NAND logic gate.

Section 16.5 Clocked CMOS Logic Circuits

16.61 (a) Figure P16.61 shows a clocked CMOS logic circuit. Make a table showing the state of each transistor (“on” or “off”), and determine the output voltages v_{O1} and v_{O2} for the input logic states listed in the following table. Assume the input conditions are sequential in time from state 1 to state 6. (b) What logic function does the circuit implement?

State	CLK	v_A	v_B	v_C
1	0	0	0	0
2	1	1	0	0
3	0	0	0	0
4	1	0	0	1
5	0	0	0	0
6	1	0	1	1

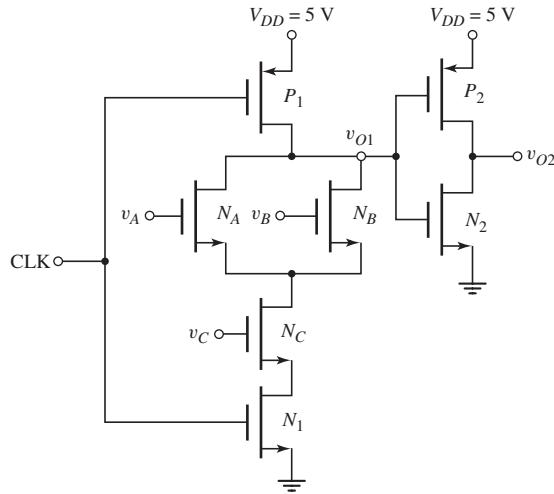


Figure P16.61

- 16.62 (a) For the circuit in Figure P16.62, make a table showing the state of each transistor (“on” or “off”), and determine the output voltages v_{O1} , v_{O2} , and v_{O3} for the input logic states listed in the following table. Assume the input conditions are sequential in time from state 1 to state 6. (b) What logic function does the circuit implement?

State	CLK	v_X	v_Y	v_Z
1	0	0	0	0
2	1	1	1	1
3	0	0	0	0
4	1	0	1	1
5	0	0	0	0
6	1	1	0	1

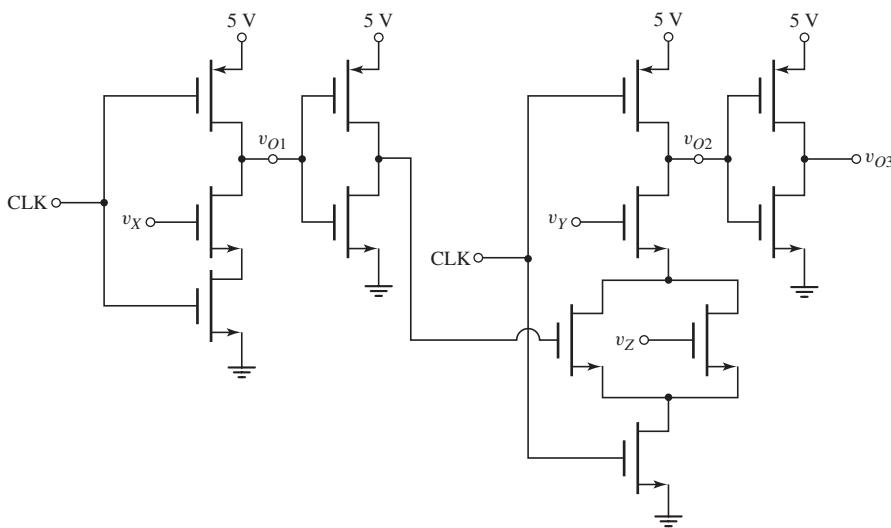


Figure P16.62

- D16.63 Sketch a clocked CMOS domino logic circuit that realizes the function $Y = A\bar{B} + \bar{A}B$. Assume that both the variable and its complement are available as input signals.
- D16.64 Sketch a clocked CMOS domino logic circuit that realizes the function $Y = AB + C(D + E)$.
- D16.65 Sketch a clocked CMOS domino logic circuit that realizes the function $Y = A(B + C)(D + E)$.
- 16.66 Consider the CMOS clocked circuit in Figure 16.44(b). Assume the effective capacitance at the v_{O1} terminal is 25 fF. If the leakage current through the M_{NA} and M_{NB} transistors is $I_{\text{Leakage}} = 2 \text{ pA}$ when these transistors and M_{P1} are cutoff, determine the time for which v_{O1} will decay by 0.5 V.

Section 16.6 Transmission Gates

- 16.67 The parameters of an NMOS transmission gate are $V_{TN} = 0.4 \text{ V}$, $K_n = 0.15 \text{ mA/V}^2$, and $C_L = 0.2 \text{ pF}$. (a) For a gate voltage of $\phi = 3.3 \text{ V}$, determine the quasi-steady-state output voltage for (i) $v_I = 0$, (ii) $v_I = 3.3 \text{ V}$, and (iii) $v_I = 2.5 \text{ V}$. (b) Repeat part (a) for a gate voltage of $\phi = 1.8 \text{ V}$.
- 16.68 The NMOS transistors in the circuit shown in Figure P16.68 have parameters $K_n = 0.2 \text{ mA/V}^2$, $V_{TN} = 0.5 \text{ V}$, $\lambda = 0$, and $\gamma = 0$. (a) For gate voltages of $\phi = 2.5 \text{ V}$, determine the quasi-steady-state output voltage for (i) $v_I = 0$, (ii) $v_I = 2.5 \text{ V}$, and (iii) $v_I = 1.8 \text{ V}$. (b) Repeat part (a) for gate voltages of $\phi = 2.0 \text{ V}$.

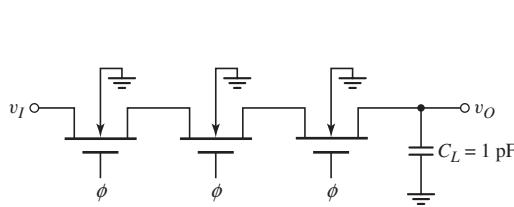


Figure P16.68

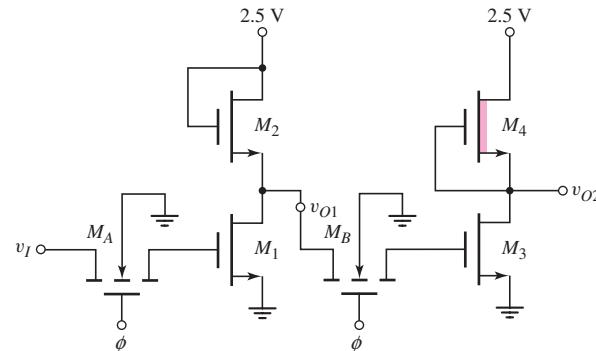


Figure P16.69

- 16.69 Consider the circuit shown in Figure P16.69. The input voltage v_I is either 0.1 V or 2.5 V. Assume gate voltages of $\phi = 2.5 \text{ V}$. The threshold voltages are $V_{TN} = -0.6 \text{ V}$ for M_4 and $V_{TN} = 0.4 \text{ V}$ for all other transistors. The width-to-length ratios are 1 for M_2 and M_4 , and 5 for M_A and M_B . (a) What are the logic 1 values for v_{O1} and v_{O2} ? (b) Design the width-to-length ratios of M_1 and M_3 such that the logic 0 values of v_{O1} and v_{O2} are 0.1 V.
- 16.70 Consider the circuit in Figure P16.70. What logic function is implemented by this circuit? Are there any potential problems with this circuit?

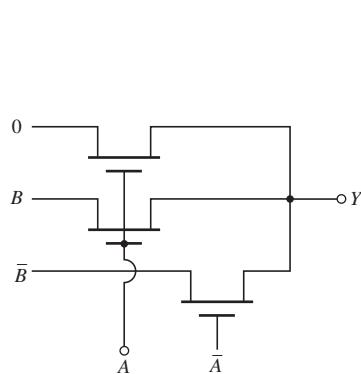


Figure P16.70

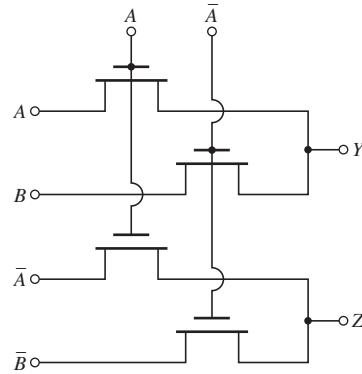


Figure P16.71

- 16.71 What is the logic function implemented by the circuit in Figure P16.71?
- D16.72 (a) Design an NMOS pass transistor logic circuit to perform the function $Y = A + B(C + D)$. Assume that both the variable and its complement are available as input signals. (b) Repeat part (a) for the function $Y = ABC + \bar{A}\bar{B}\bar{C}$.
- 16.73 Consider the circuit in Figure P16.73. (a) Determine the value of Y for $\phi = 2.5$ V and (i) $A = B = 0$; (ii) $A = 0$, $B = 2.5$ V; (iii) $A = 2.5$ V, $B = 0$; and (iv) $A = B = 2.5$ V. (b) Repeat part (a) for $\phi = 0$. (c) What is the logic function implemented by the circuit?
- 16.74 What is the logic function implemented by the circuit in Figure P16.74?
- 16.75 Consider the circuit in Figure P16.75. (a) Determine the value of Y for (i) $A = B = 0$; (ii) $A = 2.5$ V, $B = 0$; (iii) $A = 0$, $B = 2.5$ V; and (iv) $A = B = 2.5$ V. (b) What is the logic function implemented by the circuit?

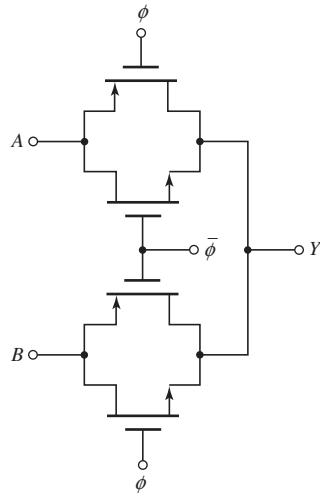


Figure P16.73

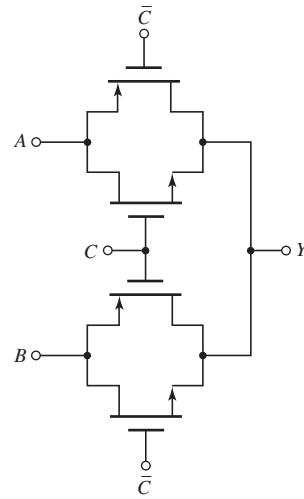


Figure P16.74

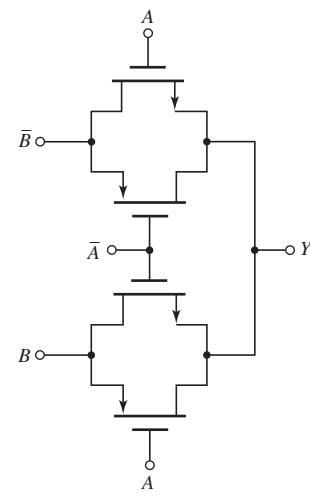


Figure P16.75

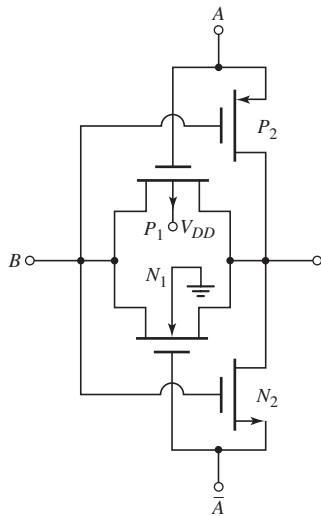


Figure P16.76

- 16.76 What is the logic function implemented by the circuit in Figure P16.76?
- 16.77 The circuit in Figure P16.77 is a form of clocked shift register. Signals ϕ_1 and ϕ_2 are nonoverlapping clock signals. Describe the operation of the circuit. Discuss any possible relationship between the width-to-length ratios of the load and driver transistors for “proper” circuit operation.

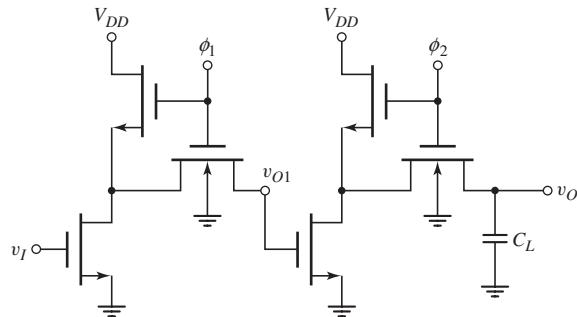


Figure P16.77

Section 16.7 Sequential Logic Circuits

- 16.78 Consider the NMOS R-S flip-flop in Figure 16.63 biased at $V_{DD} = 2.5$ V. The threshold voltages are 0.4 V (enhancement-mode devices) and -0.6 V (depletion-mode devices). The conduction parameters are $K_3 = K_6 = 40 \mu\text{A/V}^2$, $K_2 = K_5 = 100 \mu\text{A/V}^2$, and $K_1 = K_4 = 150 \mu\text{A/V}^2$. If $Q = \text{logic 0}$ and $\bar{Q} = \text{logic 1}$ initially, determine the voltage at S that will cause the flip-flop to change states.
- 16.79 Figure P16.79 shows two CMOS inverters in cascade. This circuit can be thought of as an uncoupled CMOS R/S flip flop. The transistor parameters are $K_n = K_p = 0.2 \text{ mA/V}^2$, $V_{TN} = 0.5$ V, $V_{TP} = -0.4$ V, and $\lambda_n = \lambda_p = 0$. Plot v_{O1} and v_O versus v_I . In particular, calculate the values of v_{O1} and v_O at $v_I = 1.5, 1.6, 1.7$, and 1.8 V.
- 16.80 Consider the circuit in Figure P16.80. Determine the state of the outputs for various input signals. What is the purpose of the input signal ϕ ?

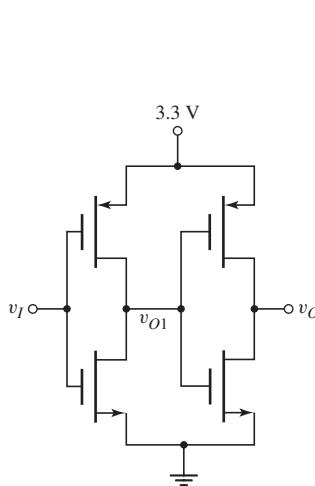


Figure P16.79

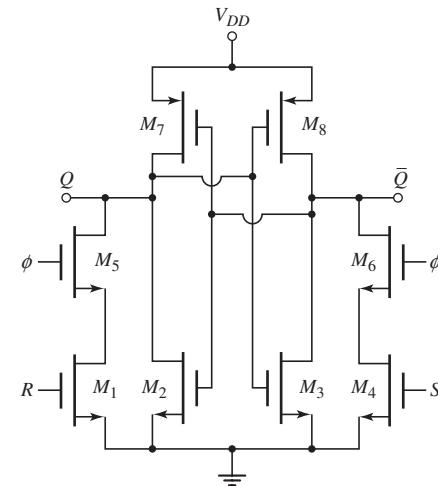


Figure P16.80

- D16.81 The circuit in Figure P16.81 is an example of a D flip-flop. (a) Explain the operation of the circuit. Is this a positive- or negative-edge-triggered flip-flop? (b) Redesign the circuit to make this a static flip-flop.

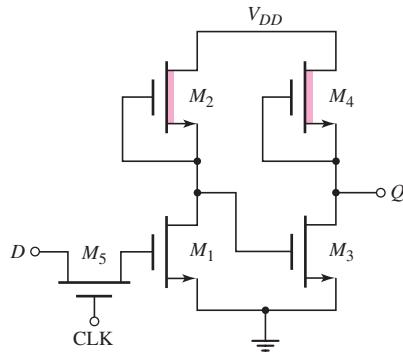


Figure P16.81

- 16.82 Show that the circuit in Figure P16.82 is a J-K flip-flop.

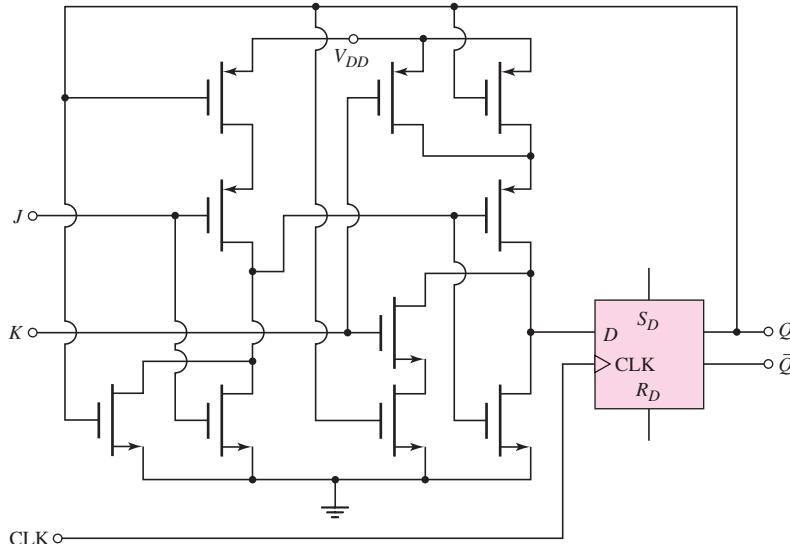


Figure P16.82

- 16.83 Reconsider the circuit shown in Figure P16.48. Show that this circuit is a J-K flip-flop with $J = v_X$, $K = v_Y$, and $\text{CLK} = v_Z$.

Section 16.8 Memories: Classifications and Architectures

- 16.84 A 256-K memory is organized in a square array and uses the NMOS NOR decoder in Figure 16.73(b) for the row- and column-decoders. (a) How many inputs does each decoder require? (b) What input to the row decoder is required to address row (i) 52, (ii) 129, and (iii) 241? (c) What input to the column decoder is required to address column (i) 24, (ii) 165, and (iii) 203?

- 16.85 (a) A 1 megabit memory is organized in a square with each memory cell being individually addressed. Determine the number of input address lines required for the row and column decoders. (b) If the 1 megabit memory is organized as 250K words \times 4 bits, determine the minimum number of input address lines required for the row and column decoders.
- 16.86 A 4096-bit RAM consists of 512 words of 8 bits each. Design the memory array to minimize the number of row and column address decoder transistors required. How many row and column address lines are required?
- 16.87 Assume that an NMOS address decoder can source $250 \mu\text{A}$ when the output goes high. If the effective capacitance of each memory cell is $C_L = 0.8 \text{ pF}$ and the effective capacitance of the address line is $C_{LA} = 5 \text{ pF}$, determine the rise time of the address line voltage if $V_{IH} = 2.7 \text{ V}$.

Section 16.9 RAM Memory Cells

- 16.88 Consider the NMOS RAM cell with resistor load in Figure 16.74(b). Assume parameters values of $k'_n = 80 \mu\text{A/V}^2$, $V_{TN} = 0.4 \text{ V}$, $V_{DD} = 2.5 \text{ V}$, and $R = 1 \text{ M}\Omega$. (a) Design the width-to-length ratio of the driver transistor such that $V_{DS} = 20 \text{ mV}$ for the on transistor. (b) Consider a 16-K memory with the cell described in part (a). Determine the standby cell current and the total memory power dissipation for a standby voltage of $V_{DD} = 1.2 \text{ V}$.
- D16.89 A 16-K NMOS RAM, with the cell design shown in Figure 16.74(b), is to dissipate no more than 200 mW in standby when biased at $V_{DD} = 2.5 \text{ V}$. Design the width-to-length ratios of the transistors and the resistance value. Assume $V_{TN} = 0.7 \text{ V}$ and $k'_n = 35 \mu\text{A/V}^2$.
- 16.90 Consider the CMOS RAM cell and data lines in Figure 16.76 biased at $V_{DD} = 2.5 \text{ V}$. Assume transistor parameters $k'_n = 80 \mu\text{A/V}^2$, $k'_p = 35 \mu\text{A/V}^2$, $V_{TN} = 0.4 \text{ V}$, $V_{TP} = -0.4 \text{ V}$, $W/L = 2$ (M_{N1} and M_{N2}), $W/L = 4$ (M_{P1} and M_{P2}), and $W/L = 1$ (all other transistors). If $Q = 0$ and $\bar{Q} = 1$, determine the steady-state values of D and \bar{D} after the row has been selected. Neglect the body effect.
- 16.91 Consider the CMOS RAM cell and data lines in Figure 16.76 with circuit and transistor parameters described in Problem 16.90. Assume initially that $Q = 0$ and $\bar{Q} = 1$. Assume the row is selected with $X = 2.5 \text{ V}$ and assume the data lines, through a write cycle, are $\bar{D} = 0$ and $D = 2.5 \text{ V}$. Determine the values of Q and \bar{Q} just after the row select has been applied.
- *16.92 Consider a general sense amplifier configuration shown in Figure 16.82 for a dynamic RAM. Assume that each bit line has a capacitance of 1 pF and is precharged to 4 V . The storage capacitance is 0.05 pF , the reference capacitance is 0.025 pF , and each is charged to 5 V for a logic 1 and to 0 V for a logic 0. The M_S and M_R gate voltages are 5 V when each cell is addressed and the transistor threshold voltages are 0.5 V . Determine the bit line voltages v_1 and v_2 after the cells are addressed for the case when (a) a logic 1 is stored and (b) a logic 0 is stored.

Section 16.10 Read-Only Memory

- D16.93 Design a 4-word \times 4-bit NMOS mask-programmed ROM to produce outputs of 1011, 1111, 0110, and 1001 when rows 1, 2, 3, and 4, respectively, are addressed.

- D16.94 Design an NMOS 16×4 mask-programmed ROM that provides the 4-bit product of two 2-bit variables.
- D16.95 Design an NMOS mask-programmed ROM that decodes a binary input and produces the output for a seven-segment array. (See Figure 2.45, Chapter 2.) The output is to be high when a particular LED is to be turned on.

Section 16.11 Data Converters

- 16.96 An analog signal in the range 0 to 5 V is to be converted to a digital signal with a quantization error of less than one percent. (a) What is the required number of bits? (b) What input voltage value represents 1 LSB? (c) What digital output represents an input voltage of 3.5424 V?
- 16.97 An analog signal in the range 0 to 3.3 V is to be converted to a digital signal with a quantization error of less than 0.5 percent. (a) What is the required number of bits? (b) What input voltage value represents 1 LSB? (c) What digital output represents an input voltage of 2.5321 V.
- 16.98 (a) What is the output voltage of the 4-bit weighted-resistor D/A in Figure 16.90 if the input is 0110? Assume $R_F = 10 \text{ k}\Omega$. (b) The input signal changes to 1001. What is the output voltage?
- 16.99 Consider the 4-bit weighted-resistor D/A converter in Figure 16.90. Let $R_F = 10 \text{ k}\Omega$. (a) What is the maximum allowed tolerance (\pm percent) in the value of R_1 so that the maximum error in the output is limited to $\pm \frac{1}{2}$ LSB? (b) Repeat part (a) for the resistor R_4 .
- 16.100 The weighted-resistor D/A converter in Figure 16.90 is to be expanded to an 8-bit device. (a) What are the required resistance values of the additional four input resistors? (b) What is the output voltage if the input is 00000001?
- 16.101 The N -bit D/A converter with an $R-2R$ ladder network in Figure 16.92 is to be designed as a 6-bit D/A device. Let $V_{\text{REF}} = -5.0 \text{ V}$ and $R = R_F = 5.0 \text{ k}\Omega$. (a) What are currents I_1, I_2, I_3, I_4, I_5 , and I_6 ? (b) The input changes by 1 LSB. What is the change in the output voltage? (c) What is the output voltage if the input is 010011? (d) What is the change in output voltage if the input changes from 101010 to 010101?
- 16.102 The 3-bit flash A/D converter in Figure 16.93 has a reference voltage of $V_{\text{REF}} = 3.3 \text{ V}$. The 3-bit output is 101. What is the range of v_A that produces this output?
- 16.103 A 6-bit flash A/D converter, similar to the one in Figure 16.93, is to be fabricated. How many resistors and comparators are required?
- 16.104 A 10-bit counting A/D converter has an analog input in the range $0 \leq v_A \leq 5 \text{ V}$ and has a clock frequency of 1 MHz. (a) What is the maximum conversion time? (b) If the output is 0010010010, what is the range of the input signal v_A (assume a quantization error of $\pm \frac{1}{2}$ LSB). (c) How many clock pulses are required to produce an output of 0100100100?
- 16.105 Consider the 10-bit counting A/D converter described in Problem 16.104. (a) What is the output if the analog input is $v_A = 3.125 \text{ V}$? (b) Repeat part (a) if $v_A = 1.8613 \text{ V}$.

COMPUTER SIMULATION PROBLEMS

- 16.106 Consider the three types of NMOS inverters shown in Figures 16.3(a), 16.5(a), and 16.7(a). Using a computer simulation, investigate the voltage transfer characteristics and the current versus input voltage characteristics

of the three types of inverters as a function of various width-to-length ratios and as a function of the body effect.

- 16.107 Using a computer simulation, investigate the propagation delay time and switching characteristics of a CMOS inverter by setting up a series of CMOS inverters in cascade. Use standard transistors and assume effective C_T load capacitances of 0.05 pF. Determine the propagation delay time as a function of various transistor width-to-length ratios.
- 16.108 Consider a three-input CMOS NAND logic circuit similar to the two-input circuit shown in Figure 16.34(a). Using a computer simulation, investigate the voltage transfer characteristics and switching characteristics for various NMOS and PMOS width-to-length ratios. What is the optimum relation between the PMOS and NMOS width-to-length ratios for symmetrical switching speeds?
- 16.109 Using a computer simulation, investigate the Q and \bar{Q} values in the CMOS RAM cell shown in Figure 16.76 during read and write cycles for various transistor width-to-length ratios. In particular, consider the relations given by Equations (16.82) and (16.84).



DESIGN PROBLEMS

- *D16.110 Design a classic CMOS logic circuit that will implement the logic function $Y = A \cdot (B + C) + D \cdot E$.
- *D16.111 Design clocked CMOS logic circuits that will implement the logic functions (a) $Y = [A \cdot B + C \cdot D]$ and (b) $Y = [A \cdot (B + C) + D]$.
- *D16.112 Design an NMOS pass logic network that implements the logic functions described in Problem 16.111.
- *D16.113 Design a clocked CMOS dynamic shift register in which the output becomes valid on the positive-going edge of a clock signal.

Physical Constants and Conversion Factors

APPENDIX

A

GENERAL CONSTANTS AND CONVERSION FACTORS

Angstrom	\AA	$1 \text{ \AA} = 10^{-4} \mu\text{m} = 10^{-8} \text{ cm} = 10^{-10} \text{ m}$
Boltzmann's constant	k	$k = 1.38 \times 10^{-23} \text{ J/K} = 8.6 \times 10^{-5} \text{ eV/K}$
Electron–volt	eV	$1 \text{ eV} = 1.6 \times 10^{-19} \text{ J}$
Electronic charge	e or q	$q = 1.6 \times 10^{-19} \text{ C}$
Micron	μm	$1 \mu\text{m} = 10^{-4} \text{ cm} = 10^{-6} \text{ m}$
Mil		$1 \text{ mil} = 0.001 \text{ in.} = 25.4 \mu\text{m}$
Nanometer	nm	$1 \text{ nm} = 10^{-9} \text{ m} = 10^{-3} \mu\text{m} = 10 \text{ \AA}$
Permittivity of free space	ϵ_o	$\epsilon_o = 8.85 \times 10^{-14} \text{ F/cm}$
Permeability of free space	μ_o	$\mu_o = 4\pi \times 10^{-9} \text{ H/cm}$
Planck's constant	h	$h = 6.625 \times 10^{-34} \text{ J-s}$
Thermal voltage	V_T	$V_T = kT/q \cong 0.026 \text{ V at } 300 \text{ K}$
Velocity of light in free space	c	$c = 2.998 \times 10^{10} \text{ cm/s}$

SEMICONDUCTOR CONSTANTS

	Si	Ge	GaAs	SiO_2
Relative dielectric constant	11.7	16.0	13.1	3.9
Bandgap energy, $E_g(\text{eV})$	1.1	0.66	1.4	
Intrinsic carrier concentration, n_i (cm^{-3} at 300 K)	1.5×10^{10}	2.4×10^{13}	1.8×10^6	

Selected Manufacturers' Data Sheets

APPENDIX

B

This appendix contains data sheets representative of transistors and op-amps. This appendix is not meant as a substitute for the appropriate data books. In some cases, therefore, only selected information is presented. These data sheets are provided courtesy of National Semiconductor.

CONTENTS

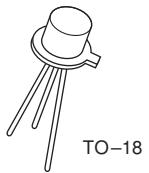
1. 2N2222 npn Bipolar transistor
2. 2N2907 pnp Bipolar transistor
3. NDS9410 n-Channel enhancement-mode MOSFET
4. LM741 Operational amplifier

2N2222/PN2222/MMBT2222/MPQ2222/2N2222A/PN2222A/MMBT2222A NPN General Purpose Amplifier



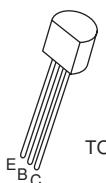
**National
Semiconductor**

**2N2222
2N2222A**



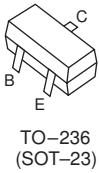
TO-18

**PN2222
PN2222A**

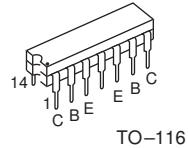


TO-92

**MMBT2222
MMBT2222A**

TO-236
(SOT-23)

MPQ2222



TO-116

NPN General Purpose Amplifier

Electrical Characteristics $T_A = 25^\circ\text{C}$ unless otherwise noted

Symbol	Parameter	Min	Max	Units
OFF CHARACTERISTICS				
$V_{(\text{BR})\text{CEO}}$	Collector-Emitter Breakdown Voltage (Note 1) ($I_C = 10 \text{ mA}$, $I_B = 0$)	2222 2222A	30 40	V
$V_{(\text{BR})\text{CBO}}$	Collector-Base Breakdown Voltage ($I_C = 10 \mu\text{A}$, $I_E = 0$)	2222 2222A	60 75	V
$V_{(\text{BR})\text{EBO}}$	Emitter Base Breakdown Voltage ($I_E = 10 \mu\text{A}$, $I_C = 0$)	2222 2222A	5.0 6.0	V
I_{CEX}	Collector Cutoff Current ($V_{\text{CE}} = 60 \text{ V}$, $V_{\text{EB}(\text{OFF})} = 3.0 \text{ V}$)	2222A		10 nA
I_{CBO}	Collector Cutoff Current ($V_{\text{CB}} = 50 \text{ V}$, $I_E = 0$) ($V_{\text{CB}} = 60 \text{ V}$, $I_E = 0$) ($V_{\text{CB}} = 50 \text{ V}$, $I_E = 0$, $T_A = 150^\circ\text{C}$) ($V_{\text{CB}} = 60 \text{ V}$, $I_E = 0$, $T_A = 150^\circ\text{C}$)	2222 2222A 22 2222A		0.01 0.01 10 10 μA
I_{EBO}	Emitter Cutoff Current ($V_{\text{EB}} = 3.0 \text{ V}$, $I_C = 0$)	2222A		10 nA
I_{BL}	Base Cutoff Current ($V_{\text{CE}} = 60 \text{ V}$, $V_{\text{EB}(\text{OFF})} = 3.0$)	2222A		20 nA
ON CHARACTERISTICS				
h_{FE}	DC Current Gain ($I_C = 0.1 \text{ mA}$, $V_{\text{CE}} = 10 \text{ V}$) ($I_C = 1.0 \text{ mA}$, $V_{\text{CE}} = 10 \text{ V}$) ($I_C = 10 \text{ mA}$, $V_{\text{CE}} = 10 \text{ V}$) ($I_C = 10 \text{ mA}$, $V_{\text{CE}} = 10 \text{ V}$, $T_A = -55^\circ\text{C}$) ($I_C = 150 \text{ mA}$, $V_{\text{CE}} = 10 \text{ V}$) (Note 1) ($I_C = 150 \text{ mA}$, $V_{\text{CE}} = 1.0 \text{ V}$) (Note 1) ($I_C = 500 \text{ mA}$, $V_{\text{CE}} = 10 \text{ V}$) (Note 1)	2222 2222A	35 50 75 35 100 50 30 40	300

Note 1: Pulse Test: Pulse Width $\leq 300 \mu\text{s}$, Duty Cycle $\leq 2.0\%$.

2222/PN2222/MMBT2222/MMQ2222/NPN General Purpose Amplifier

NPN General Purpose Amplifier (Continued)

Electrical Characteristics $T_A = 25^\circ\text{C}$ unless otherwise noted (Continued)

Symbol	Parameter		Min	Max	Units
ON CHARACTERISTICS (Continued)					
V_{CE} (sat)	Collector-Emitter Saturation Voltage (Note 1) ($I_C = 150 \text{ mA}$, $I_B = 15 \text{ mA}$) ($I_C = 500 \text{ mA}$, $I_B = 50 \text{ mA}$)	2222 2222A 2222 2222A		0.4 0.3 1.6 1.0	V
V_{BE} (sat)	Base-Emitter Saturation Voltage (Note 1) ($I_C = 150 \text{ mA}$, $I_B = 15 \text{ mA}$) ($I_C = 500 \text{ mA}$, $I_B = 50 \text{ mA}$)	2222 2222A 2222 2222A	0.6 0.6	1.3 1.2 2.6 2.0	V
SMALL-SIGNAL CHARACTERISTICS					
f_T	Current Gain—Bandwidth Product (Note 3) ($I_C = 20 \text{ mA}$, $V_{CE} = 20 \text{ V}$, $f = 100 \text{ MHz}$)	2222 2222A	250 300		MHz
C_{obo}	Output Capacitance (Note 3) ($V_{CB} = 10 \text{ V}$, $I_E = 0$, $f = 100 \text{ kHz}$)			8.0	pF
C_{ibo}	Input Capacitance (Note 3) ($V_{EB} = 0.5 \text{ V}$, $I_C = 0$, $f = 100 \text{ kHz}$)	2222 2222A		30 25	pF
$r_b' C_C$	Collector Base Time Constant ($I_E = 20 \text{ mA}$, $V_{CB} = 20 \text{ V}$, $f = 31.8 \text{ MHz}$)	2222A		150	ps
NF	Noise Figure ($I_C = 100 \mu\text{A}$, $V_{CE} = 10 \text{ V}$, $R_S = 1.0 \text{ k}\Omega$, $f = 1.0 \text{ kHz}$)	2222A		4.0	dB
$\text{Re}(h_{ie})$	Real Part of Common-Emitter High Frequency Input Impedance ($I_C = 20 \text{ mA}$, $V_{CE} = 20 \text{ V}$, $f = 300 \text{ MHz}$)			60	Ω
SWITCHING CHARACTERISTICS					
t_D	Delay Time	$(V_{CC} = 30 \text{ V}, V_{BE(OFF)} = 0.5 \text{ V}, I_C = 150 \text{ mA}, I_{B1} = 15 \text{ mA})$	except MPQ2222	10	ns
t_R	Rise Time			25	ns
t_S	Storage Time	$(V_{CC} = 30 \text{ V}, I_C = 150 \text{ mA}, I_{B1} = I_{B2} = 15 \text{ mA})$	except MPQ2222	225	ns
t_F	Fall Time			60	ns

Note 1: Pulse Test: Pulse Width <300 μs , Duty Cycle $\leq 2.0\%$.

Note 2: For characteristics curves, see Process 19.

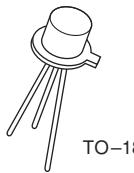
Note 3: f_T is defined as the frequency at which h_{fe} extrapolates to unity.

2N2907/PN2907/MMBT2907/MPQ2907/2N2907A/PN2907A/MMBT2907A PNP General Purpose Amplifier



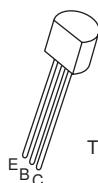
**National
Semiconductor**

**2N2907
2N2907A**



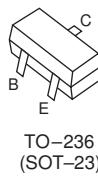
TO-18

**PN2907
PN2907A**

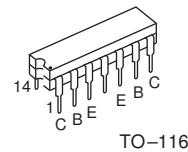


TO-92

**MMBT2907
MMBT2907A**

TO-236
(SOT-23)

MPQ2907



TO-116

PNP General Purpose Amplifier

Electrical Characteristics $T_A = 25^\circ\text{C}$ unless otherwise noted

Symbol	Parameter	Min	Max	Units
OFF CHARACTERISTICS				
$V_{(BR)CEO}$	Collector-Emitter Breakdown Voltage (Note 1) ($I_C = 10 \text{ mA}_\text{dc}$, $I_B = 0$)	2907 2907A	40 60	Vdc
$V_{(BR)CBO}$	Collector-Base Breakdown Voltage ($I_C = 10 \mu\text{A}_\text{dc}$, $I_E = 0$)		60	Vdc
$V_{(BR)EBO}$	Emitter Base Breakdown Voltage ($I_E = 10 \mu\text{A}_\text{dc}$, $I_C = 0$)		5.0	Vdc
I_{CEX}	Collector Cutoff Current ($V_{CE} = 30 \text{ Vdc}$, $V_{BE} = 0.5 \text{ Vdc}$)		50	nAdc
I_{CBO}	Collector Cutoff Current ($V_{CB} = 50 \text{ Vdc}$, $I_E = 0$) ($V_{CB} = 50 \text{ Vdc}$, $I_E = 0$, $T_A = 150^\circ\text{C}$)	2907 2907A 2907 2907A	0.020 0.010 20 10	μAdc
I_B	Base Cutoff Current ($V_{CE} = 30 \text{ Vdc}$, $V_{EB} = 0.5 \text{ Vdc}$)		50	nAdc

2N2907/PN2907/MMBT2907/MMBT2907A/PN2907A/MMBT2907A PNP General Purpose Amplifier

PNP General Purpose Amplifier (Continued)

Electrical Characteristics $T_A = 25^\circ\text{C}$ unless otherwise noted (Continued)

Symbol	Parameter	Min	Max	Units	
ON CHARACTERISTICS					
h_{FE}	DC Current Gain ($I_C = 0.1 \text{ mAdc}$, $V_{CE} = 10 \text{ Vdc}$) ($I_C = 1.0 \text{ mAdc}$, $V_{CE} = 10 \text{ Vdc}$) ($I_C = 10 \text{ mAdc}$, $V_{CE} = 10 \text{ Vdc}$) ($I_C = 150 \text{ mAdc}$, $V_{CE} = 10 \text{ Vdc}$) (Note 1) ($I_C = 500 \text{ mAdc}$, $V_{CE} = 10 \text{ Vdc}$) (Note 1)	2907 2907A 2907 2907A 2907 2907A 2907 2907A	35 75 50 100 75 100 100 30	300	
$V_{CE(\text{sat})}$	Collector-Emitter Saturation Voltage (Note 1) ($I_C = 150 \text{ mAdc}$, $I_B = 15 \text{ mAdc}$) ($I_C = 500 \text{ mAdc}$, $I_B = 50 \text{ mAdc}$)			0.4 1.6	Vdc
$V_{BE(\text{sat})}$	Base-Emitter Saturation Voltage ($I_C = 150 \text{ mAdc}$, $I_B = 15 \text{ mAdc}$) (Note 1) ($I_C = 150 \text{ mAdc}$, $I_B = 50 \text{ mAdc}$)			1.3 2.6	Vdc
SMALL-SIGNAL CHARACTERISTICS					
f_T	Current Gain—Bandwidth Product ($I_C = 50 \text{ mAdc}$, $V_{CE} = 20 \text{ Vdc}$, $f = 100 \text{ MHz}$)		200		MHz
C_{obo}	Output Capacitance ($V_{CB} = 10 \text{ Vdc}$, $I_E = 0$, $f = 100 \text{ kHz}$)			8.0	pF
C_{ibo}	Input Capacitance ($V_{EB} = 2.0 \text{ Vdc}$, $I_C = 0$, $f = 100 \text{ kHz}$)			30	pF
SWITCHING CHARACTERISTICS					
t_{on}	Turn-On Time	$(V_{CC} = 30 \text{ Vdc}$, $I_C = 150 \text{ mAdc}$, $I_{B1} = 15 \text{ mAdc}$)		45	ns
t_d	Delay Time			10	ns
t_r	Rise Time			40	ns
t_{off}	Turn-Off Time	$(V_{CC} = 6.0 \text{ Vdc}$, $I_C = 150 \text{ mAdc}$, $I_{B1} = I_{B2} = 15 \text{ mAdc}$)		100	ns
t_s	Storage Time			80	ns
t_f	Fall Time			30	ns

Note 1: Pulse Test: Pulse Width $\leq 300 \mu\text{s}$, Duty Cycle $\leq 2.0\%$.



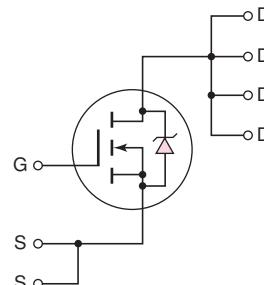
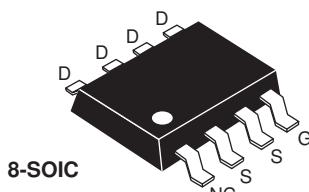
NDS9410 Single N-Channel Enhancement Mode Field Effect Transistor

General Description

These N-channel enhancement mode power field effect transistors are produced using National's proprietary, high cell density, DMOS technology. This very high density process has been especially tailored to minimize on-state resistance, provide superior switching performance, and withstand high energy pulses in the avalanche and commutation modes. These devices are particularly suited to low voltage applications such as laptop computer power management and other battery powered circuits where fast switching, low in-line power loss, and resistance to transients are needed.

Features

- 7.0 A, 30 V, $R_{DS(ON)} = 0.03 \Omega$
- Rugged internal source-drain diode can eliminate the need for an external Zener diode transient suppressor
- High density cell design (3.8 million/in²) for extremely low $R_{DS(ON)}$
- High power and current handling capability in a widely used surface mount package
- Critical DC electrical parameters specified at elevated temperature



ABSOLUTE MAXIMUM RATINGS

$T_C = 25^\circ\text{C}$ unless otherwise noted

Symbol	Parameter	NDS9410	Units
V_{DSS}	Drain-Source Voltage	30	V
V_{DGR}	Drain-Gate Voltage ($R_{GS} \leq 1 \text{ M}\Omega$)	30	V
V_{GSS}	Gate-Source Voltage	± 20	V
I_D	Drain Current – Continuous @ $T_A = 25^\circ\text{C}$	± 7.0	A
	– Continuous @ $T_A = 70^\circ\text{C}$	± 5.8	A
	– Pulsed	± 20	A
P_D	Maximum Power Dissipation @ $T_A = 25^\circ\text{C}$	2.5 (Note 1)	W
T_JT_{STG}	Operating and Storage Temperature Range	-55 to 150	$^\circ\text{C}$

THERMAL CHARACTERISTICS

$R_{\theta JA(t)}$	Thermal Resistance, Junction-to-Ambient (Pulse = 10 seconds)	50 (Note 1)	$^\circ\text{C/W}$
$R_{\theta JA}$	Thermal Resistance, Junction-to-Ambient (Steady-State)	100 (Note 2)	$^\circ\text{C/W}$

TYPICAL ELECTRICAL CHARACTERISTICS

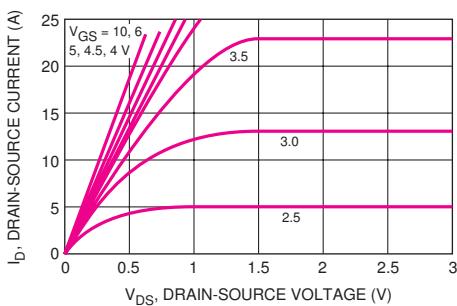


Figure 1. On-Region Characteristics

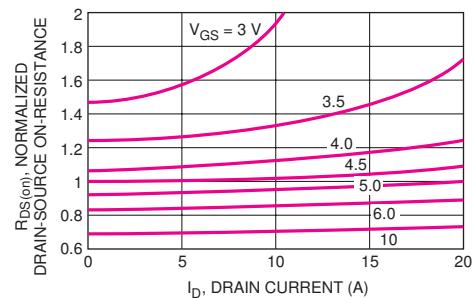


Figure 2. On-Resistance Variation with Drain Current and Gate Voltage

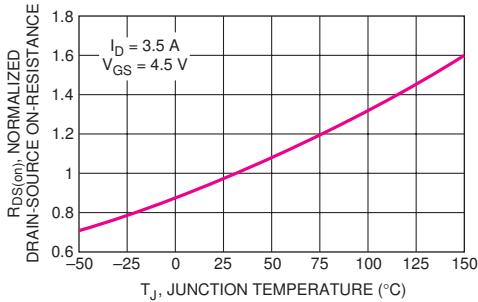


Figure 3. On-Resistance Variation with Temperature

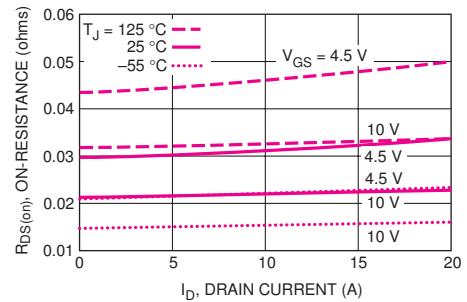


Figure 4. On-Resistance Variation with Drain Current and Temperature

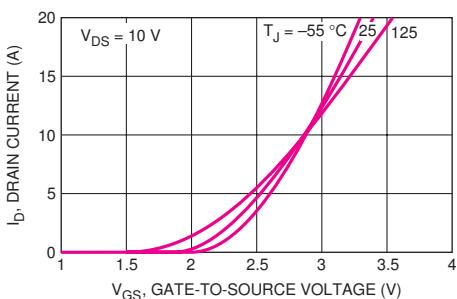


Figure 5. Transfer Characteristics

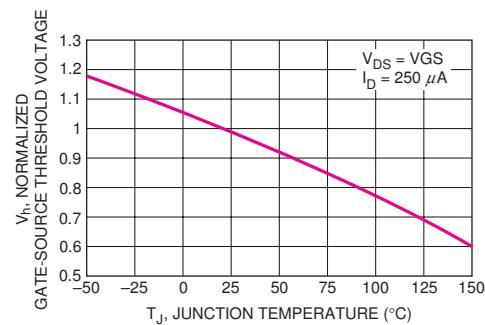


Figure 6. Gate Threshold Variation with Temperature



National Semiconductor

LM741 Operation Amplifier

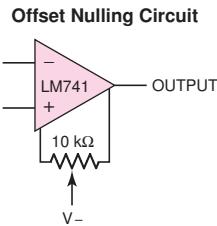
General Description

The LM741 series are general purpose operational amplifiers which feature improved performance over industry standards like the LM709. They are direct, plug-in replacements for the 709C, LM201, MC1439 and 748 in most applications. The amplifiers offer many features which make their application nearly foolproof: overload protection on the input and output,

no latch-up when the common mode range is exceeded as well as freedom from oscillations. The LM741C/LM741E are identical to the LM741/LM741A except that the LM741C/LM741E have their performance guaranteed over a 0 °C to +70 °C temperature range, instead of -55 °C to +125 °C.

LM741 Operational Amplifier

Schematic Diagram – (See Figure 13.3 in text)



Absolute Maximum Ratings

	LM741A	LM741E	LM741	LM741C
Supply Voltage	±22 V	±22 V	±22 V	±18 V
Power Dissipation	500 mW	500 mW	500 mW	500 mW
Differential Input Voltage	±30 V	±30 V	±30 V	±30 V
Input Voltage (Note 2)	±15 V	±15 V	±15 V	±15 V
Output Short Circuit Duration	Continuous	Continuous	Continuous	Continuous
Operating Temperature Range	−55 °C to +125 °C	0 °C to +70 °C	−55 °C to +125 °C	0 °C to +70 °C
Storage Temperature Range	−65 °C to +150 °C			
Junction Temperature	150 °C	100 °C	150 °C	100 °C

Electrical Characteristics

Parameter	Conditions	LM741A/LM741E			LM741			LM741C			Units
		Min	Typ	Max	Min	Typ	Max	Min	Typ	Max	
Input Offset Voltage	$T_A = 25^\circ\text{C}$ $R_S \leq 10 \text{ k}\Omega$ $R_S \leq 50 \Omega$		0.8	3.0		1.0	5.0		2.0	6.0	mV mV
	$T_{A\text{MIN}} \leq T_A \leq T_{A\text{MAX}}$ $R_S \leq 50 \Omega$ $R_S \leq 10 \text{ k}\Omega$			4.0			6.0			7.5	mV mV
Average Input Offset Voltage Drift				15							$\mu\text{V}/^\circ\text{C}$
Input Offset Voltage Adjustment Range	$T_A = 25^\circ\text{C}$, $V_S = \pm 20 \text{ V}$	±10			±15			±15			mV
Input Offset Current	$T_A = 25^\circ\text{C}$		3.0	30		20	200		20	200	nA
	$T_{A\text{MIN}} \leq T_A \leq T_{A\text{MAX}}$			70		85	500			300	nA
Average Input Offset Current Drift				0.5							$\text{nA}/^\circ\text{C}$
Input Bias Current	$T_A = 25^\circ\text{C}$		30	80		80	500		80	500	nA
	$T_{A\text{MIN}} \leq T_A \leq T_{A\text{MAX}}$			0.210			1.5			0.8	μA
Input Resistance	$T_A = 25^\circ\text{C}$, $V_S = \pm 20 \text{ V}$	1.0	6.0		0.3	2.0		0.3	2.0		MΩ
	$T_{A\text{MIN}} \leq T_A \leq T_{A\text{MAX}}$, $V_S = \pm 20 \text{ V}$	0.5									MΩ
Input Voltage Range	$T_A = 25^\circ\text{C}$							±12	±13		V
	$T_{A\text{MIN}} \leq T_A \leq T_{A\text{MAX}}$				±12	±13					V
Large Signal Voltage Gain	$T_A = 25^\circ\text{C}$, $R_L \geq 2 \text{ k}\Omega$ $V_S = \pm 20 \text{ V}$, $V_O = \pm 15 \text{ V}$ $V_S = \pm 15 \text{ V}$, $V_O = \pm 10 \text{ V}$	50			50	200		20	200		V/mV V/mV
	$T_{A\text{MIN}} \leq T_A \leq T_{A\text{MAX}}$, $R_L \geq 2 \text{ k}\Omega$ $V_S = \pm 20 \text{ V}$, $V_O = \pm 15 \text{ V}$ $V_S = \pm 15 \text{ V}$, $V_O = \pm 10 \text{ V}$ $V_S = \pm 5 \text{ V}$, $V_O = \pm 2 \text{ V}$	32			25			15			V/mV V/mV V/mV

Electrical Characteristics (Continued)											
Parameter	Conditions	LM741A/LM741E			LM741			LM741C			Units
		Min	Typ	Max	Min	Typ	Max	Min	Typ	Max	
Output Voltage Swing	$V_S = \pm 20 V$ $R_L \geq 10 k\Omega$ $R_L \geq 2 k\Omega$	± 16 ± 15									V V
	$V_S = \pm 15 V$ $R_L \geq 10 k\Omega$ $R_L \geq 2 k\Omega$				± 12 ± 10	± 14 ± 13		± 12 ± 10	± 14 ± 13		V V
Output Short Circuit Current	$T_A = 25 ^\circ C$ $T_{AMIN} \leq T_A \leq T_{AMAX}$	10 10	25	35 40		25			25		mA mA
Common-Mode Rejection Ratio	$T_{AMIN} \leq T_A \leq T_{AMAX}$ $R_S \leq 10 k\Omega$, $V_{CM} = \pm 12 V$ $R_S \leq 50 \Omega$, $V_{CM} = \pm 12 V$	80	95		70	90		70	90		dB dB
Supply Voltage Rejection Ratio	$T_{AMIN} \leq T_A \leq T_{AMAX}$ $V_S = \pm 20 V$ to $V_S = \pm 5 V$ $R_S \leq 50 \Omega$ $R_S \leq 10 \Omega$	86	96		77	96		77	96		dB dB
Transient Response Rise Time Overshoot	$T_A = 25 ^\circ C$, Unity Gain		0.25 6.0	0.8 20		0.3 5			0.3 5		μs %
Bandwidth	$T_A = 25 ^\circ C$	0.437	1.5								MHz
Slew Rate	$T_A = 25 ^\circ C$, Unity Gain	0.3	0.7			0.5			0.5		V/ μs
Supply Current	$T_A = 25 ^\circ C$				1.7	2.8		1.7	2.8		mA
Power Consumption	$T_A = 25 ^\circ C$ $V_S = \pm 20 V$ $V_S = \pm 15 V$		80	150		50	85		50	85	mW mW
LM741A	$V_S = \pm 20 V$ $T_A = T_{AMIN}$ $T_A = T_{AMAX}$			165 135							mW mW
LM741E	$V_S = \pm 20 V$ $T_A = T_{AMIN}$ $T_A = T_{AMAX}$			150 150							mW mW
LM741	$V_S = \pm 15 V$ $T_A = T_{AMIN}$ $T_A = T_{AMAX}$				60 45	100 75					mW mW

Note 2: For supply voltages less than $\pm 15 V$, the absolute maximum input voltage is equal to the supply voltage.

Standard Resistor and Capacitor Values

APPENDIX C

In this appendix, we list standard component values, which are used for selecting resistor and capacitor values in designing discrete electronic circuits and systems. Low-power carbon and film resistors with 2 percent to 20 percent tolerances have a standard set of values and a standard color-band marking scheme. These tabulated values may vary from one manufacturer to another, so the tables should be considered as typical.

C.1 CARBON RESISTORS

Standard resistor values are listed in Table C.1. The lightface type indicates 2 percent and 5 percent tolerance values; the boldface type indicates 10% tolerance resistor values.

Table C.1 Standard resistance values ($\times 10^3$)

10	16	27	43	68
11	18	30	47	75
12	20	33	51	82
13	22	36	56	91
15	24	39	62	100

Discrete carbon resistors have a standard color-band marking scheme, which makes it easy to recognize resistor values in a circuit or a parts bin, without having to search for a printed legend. The color bands start at one end of the resistor, as shown in Figure C.1. Two digits and a multiplier digit determine the resistor value. The additional color bands indicate the tolerance and reliability. The digit and multiplier color-code is given in Table C.2.

For example, the first three color bands of a $4.7\text{ k}\Omega$ resistor are yellow, violet, and red. The first two digits are 47 and the multiplier is 100. The first three color bands on a $150\text{ k}\Omega$ resistor are brown, green, and yellow.

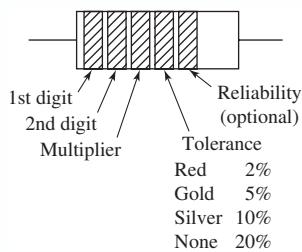


Figure C.1 Color-band notation of low-power carbon-composition resistors

Table C.2 Digit and multiple color code

Digit	Color	Multiplier	Number of zeros
	Silver	0.01	-2
	Gold	0.1	-1
0	Black	1	0
1	Brown	10	1
2	Red	100	2
3	Orange	1 k	3
4	Yellow	10 k	4
5	Green	100 k	5
6	Blue	1 M	6
7	Violet	10 M	7
8	Gray		
9	White		

Ten percent tolerance carbon resistors are available in the following power ratings: $\frac{1}{4}$, $\frac{1}{2}$, 1, and 2 W.

C.2 PRECISION RESISTORS (ONE PERCENT TOLERANCE)

Metal-film precision resistors can have tolerance levels in the $\frac{1}{2}$ percent to 1 percent range. These resistors use a four-digit code printed on the resistor body, rather than the color-band scheme. The first three digits denote a value, and the last digit is the multiplier for the number of zeros. For example, 2503 denotes a $250\text{ k}\Omega$ resistor, and 2000 denotes a $200\text{ }\Omega$ resistor. If the resistor's value is too small to be described in this way, an *R* is used to indicate the decimal point; for example, 37R5 is a $37.5\text{ }\Omega$ resistor, and 10R0 is a $10.0\text{ }\Omega$ resistor.

The standard values typically range from $10\text{ }\Omega$ to $301\text{ k}\Omega$. Standard values in each decade are given in Table C.3.

Table C.3 Standard precision resistance values

100	140	196	274	383	536	750
102	143	200	280	392	549	768
105	147	205	287	402	562	787
107	150	210	294	412	576	806
110	154	215	301	422	590	825
113	158	221	309	432	604	845
115	162	226	316	442	619	866
118	165	232	324	453	634	887
121	169	237	332	464	649	909
124	174	243	340	475	665	931
127	178	249	348	487	681	953
130	182	255	357	499	698	976
133	187	261	365	511	715	
137	191	267	374	523	732	

One percent resistors are often used in applications that require excellent stability and accuracy; a small adjustable trimmer resistor may be connected in series to the 1 percent resistor to set a precise resistance value. It is important to realize that 1 percent resistors are only guaranteed to be within 1 percent of their rated value under a specified set of conditions. Resistance variations due to temperature or humidity changes, and operation at full rated power can exceed the 1 percent tolerance.

C.3 CAPACITORS

Typical capacitor values for 10 percent tolerance capacitors from one manufacturer are listed in Table C.4. The range of capacitance values for the ceramic-disk capacitor is approximately 10 pF to 1 μF .

Table C.4 Ceramic-disk capacitors

3.3	30	200	600	2700
5	39	220	680	3000
6	47	240	750	3300
6.8	50	250	800	3900
7.5	51	270	820	4000
8	56	300	910	4300
10	68	330	1000	4700
12	75	350	1200	5000
15	82	360	1300	5600
18	91	390	1500	6800
20	100	400	1600	7500
22	120	470	1800	8200
24	130	500	2000	
25	150	510	2200	
27	180	560	2500	

Tantalum capacitors ($\times 10^n$) (to 330 μF)

0.0047	0.010	0.022
0.0056	0.012	0.027
0.0068	0.015	0.033
0.0082	0.018	0.039

Reading List

APPENDIX

D

GENERAL ELECTRONICS TEXTS

- Burns, S. G.; and P. R. Bond. *Principles of Electronic Circuits*. 2nd ed. Boston: PWS Publishing Co., 1997.
- Hambley, A. R. *Electronics*, 2nd ed. Upper Saddle River, NJ: Prentice-Hall, Inc., 2003.
- Hayt, W. H., Jr.; and G. W. Neudeck. *Electronic Circuit Analysis and Design*. 2nd ed. Boston: Houghton Mifflin Co., 1984.
- Horenstein, M. N. *Microelectronic Circuits and Devices*. 2nd ed. Englewood Cliffs, NJ: Prentice Hall, Inc., 1995.
- Horowitz, P.; and W. Hill. *The Art of Electronics*. 2nd ed. New York: Cambridge University Press, 1989.
- Howe, R. T.; and C. G. Sodini. *Microelectronics: An Integrated Approach*. Upper Saddle River, NJ: Prentice-Hall, Inc., 1997.
- Jaeger, R. C.; and T. N. Blalock. *Microelectronic Circuit Design*, 3rd ed. New York: McGraw-Hill, 2008.
- Malik, N. R. *Electronic Circuits: Analysis, Simulation, and Design*. Englewood Cliffs, NJ: Prentice Hall, Inc., 1995.
- Mauro, R. *Engineering Electronics*. Englewood Cliffs, NJ: Prentice Hall, Inc., 1989.
- Millman, J.; and A. Graybel. *Microelectronics*. 2nd ed. New York: McGraw-Hill Book Co., 1987.
- Mitchell, F. H., Jr.; and F. H. Mitchell, Sr. *Introduction to Electronics Design*. 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, Inc., 1992.
- Rashid, M. H. *Microelectronic Circuits: Analysis and Design*. Boston: PWS Publishing Co., 1999.
- Razaavi, B. *Fundamentals of Microelectronics*. New York: John Wiley and Sons, Inc., 2008.
- Roden, M. S.; and G. L. Carpenter. *Electronic Design: From Concept to Reality*. 3rd ed. Burbank, CA: Discovery Press, 1997.
- Schubert, T., Jr; and E. Kim. *Active and Non-Linear Electronics*. New York: John Wiley and Sons, Inc., 1996.
- Spencer, R. R.; and M. S. Ghausi. *Introduction to Electronic Circuit Design*. Upper Saddle River, NJ: Prentice-Hall, Inc., 2003.
- Sedra, A. S.; and K. C. Smith. *Microelectronic Circuits*, 5th ed. New York: Oxford University Press, 2004.

LINEAR CIRCUIT THEORY

- Alexander, C. K.; and M. N. O. Sadiku. *Fundamentals of Electric Circuits*, 3rd ed. New York: McGraw-Hill, 2007.
- Bode, H. W. *Network Analysis and Feedback Amplifier Design*. Princeton, NJ; D. Van Nostrand Co., 1945.
- Hayt, W. H., Jr.; J. E. Kemmerly; and S. M. Durbin. *Engineering Circuit Analysis*, 7th ed. New York: McGraw-Hill, 2007.

- Irwin, J. D.; and R. M. Nelms. *Basic Engineering Circuit Analysis*, 9th ed. New York: John Wiley and Sons, Inc., 2008.
- Johnson, D. E.; J. L. Hillburn; J. R. Johnson; and P. D. Scott. *Basic Electric Circuit Analysis*. 5th ed. Englewood Cliffs, NJ: Prentice Hall, Inc., 1995.
- Nilsson, J. W.; and S. A. Riedel. *Electric Circuits*, 8th ed. Upper Saddle River, NJ: Prentice-Hall, Inc., 2007.
- Thomas, R. E.; A. J. Rosa; and G. J. Toussaint. *The Analysis and Design of Linear Circuits*, 6th ed. New York: John Wiley and Sons, Inc., 2008.

SEMICONDUCTOR DEVICES

- Neamen, D. A. *An Introduction to Semiconductor Devices*. Boston: McGraw-Hill, 2006.
- Neamen, D. A. *Semiconductor Physics and Devices: Basic Principles*, 3rd ed. Boston: McGraw-Hill, 2003.
- Streetman, B. G.; and S. Banerjee. *Solid State Electronic Devices*, 6th ed. Upper Saddle River, NJ: Prentice Hall, 2006.
- Sze, S. M.; and K. K. Ng. *Physics of Semiconductor Devices*, 3rd ed. New York: John Wiley and Sons, Inc., 2007.
- Taur, Y.; and T. H. Ning. *Fundamentals of Modern VLSI Devices*. Cambridge, United Kingdom: Cambridge University Press, 1998.

ANALOG INTEGRATED CIRCUITS

- Allen, P. E.; and D. R. Holberg. *CMOS Analog Circuit Design*, 2nd ed. New York: Oxford University Press, 2002.
- Geiger, R. L.; P. E. Allen; and N. R. Strader. *VLSI Design Techniques for Analog and Digital Circuits*. New York: McGraw-Hill Publishing Co., 1990.
- Gray, P. R.; P. J. Hurst; S. H. Lewis; and R. G. Meyer. *Analysis and Design of Analog Integrated Circuits*, 5th ed. New York: John Wiley and Sons, Inc., 2009.
- Johns, D. A.; and K. Martin. *Analog Integrated Circuit Design*. New York: John Wiley and Sons, Inc., 1997.
- Laker, K. R.; and W. M. C. Sansen. *Design of Analog Integrated Circuits and Systems*. New York: McGraw-Hill, Inc., 1994.
- Northrop, R. B. *Analog Electronic Circuits*. Reading, MA: Addison-Wesley Publishing Co., 1990.
- Razavi, B. *Design of Analog CMOS Integrated Circuits*. Boston: McGraw-Hill, 2001.
- Soclof, S. *Design and Applications of Analog Integrated Circuits*. Englewood Cliffs, NJ: Prentice Hall, Inc., 1991.
- Solomon, J. E. “The Monolithic Op-Amp: A Tutorial Study,” *IEEE Journal of Solid-State Circuits* SC-9, No. 6 (December 1974), pp. 314–32.
- Widlar, R. J. “Design Techniques for Monolithic Operational Amplifiers,” *IEEE Journal of Solid-State Circuits* SC-4 (August 1969), pp. 184–91.

OPERATIONAL AMPLIFIER CIRCUITS

- Barna, A.; and D. I. Porat. *Operational Amplifiers*. 2nd ed. New York: John Wiley and Sons, Inc., 1989.

- Berlin, H. M. *Op-Amp Circuits and Principles*. Carmel, IN: SAMS, A division of Macmillan Computer Publishing, 1991.
- Clayton, G.; and B. Newby. *Operational Amplifiers*. London: Butterworth-Heinemann, Ltd., 1992.
- Coughlin, R. F.; and F. F. Driscoll. *Operational Amplifiers and Linear Integrated Circuits*. Englewood Cliffs, NJ: Prentice Hall, Inc., 1977.
- Fiore, J. M. *Operational Amplifiers and Linear Integrated Circuits: Theory and Application*. New York: West Publishing Co., 1992.
- Franco, S. *Design with Operational Amplifiers and Analog Integrated Circuits*, 2nd ed. New York: McGraw-Hill, 1998.
- Graeme, J. G.; G. E. Tobey; and L. P. Huelsman. *Operational Amplifiers: Design and Applications*. New York: McGraw-Hill Book Co., 1971.
- Helms, H. *Operational Amplifiers 1987 Source Book*. Englewood Cliffs, NJ: Prentice Hall, Inc., 1987.

DIGITAL CIRCUITS AND DEVICES

- Ayers, J. E. *Digital Integrated Circuits: Analysis and Design*. New York: CRC Press, 2004.
- Baker, R. J.; H. W. Li, and D. E. Boyce. *CMOS Circuit Design, Layout, and Simulation*. New York: IEEE Press, 1998.
- CMOS/NMOS Integrated Circuits*, RCA Solid State, 1980.
- DeMassa, T. A.; and Z. Ciccone. *Digital Integrated Circuits*. New York: John Wiley and Sons, Inc., 1996.
- Glasford, G. M. *Digital Electronic Circuits*. Englewood Cliffs, NJ: Prentice Hall, Inc., 1988.
- Hauser, J. R., “Noise Margin Criteria for Digital Logic Circuits.” *IEEE Transactions on Education* 36, No. 4 (November 1993), pp. 363–68.
- Hodges, D. A.; H. G. Jackson, and R. A. Saleh. *Analysis and Design of Digital Integrated Circuits*, 3rd ed. New York: McGraw-Hill, 2004.
- Kang, S. M.; and Y. Leblebici. *CMOS Digital Integrated Circuits: Analysis and Design*. 3rd ed. New York: McGraw-Hill, 2003.
- Lohstroh, J. “Static and Dynamic Noise Margins of Logic Circuits,” *IEEE Journal of Solid-State Circuits* SC-14, No. 3 (June 1979), pp. 591–98.
- Mead, C.; and L. Conway. *Introduction to VLSI Systems*. Reading, MA: Addison-Wesley Publishing Co., Inc., 1980.
- Mukherjee, A. *Introduction to nMOS and CMOS VLSI Systems Design*. Englewood Cliffs, NJ: Prentice Hall, Inc., 1986.
- Prince, B. *Semiconductor Memories: A Handbook of Design, Manufacture and Applications*. 2nd ed. New York: John Wiley and Sons, Inc., 1991.
- Rabaey, J. M.; A. Chandrakasan, and B. Nikolic. *Digital Integrated Circuits: A Design Perspective*, 2nd ed. Upper Saddle River, NJ: Prentice-Hall, 2003.
- Segura, J.; and C. F. Hawkins. *CMOS Electronics: How It Works, How It Fails*. Piscataway, NJ: IEEE Press, 2004.
- Wang, N. *Digital MOS Integrated Circuits*. Englewood Cliffs, NJ: Prentice Hall, Inc., 1989.
- Wilson, G. R. “Advances in Bipolar VLSI.” *Proceedings of the IEEE* 78, No. 11 (November 1990), pp. 1707–19.

SPICE AND PSPICE REFERENCES

- Banzhap, W. *Computer-Aided Circuit Analysis Using PSpice*. 2nd ed. Englewood Cliffs, NJ: Prentice Hall, Inc., 1992.
- Brown, W. L.; and A. Y. J. Szeto. "Verifying Spice Results with Hand Calculations: Handling Common Discrepancies." *IEEE Transactions on Education*, 37, No. 4 (November 1994), pp. 358–68.
- Goody, R. W. *MicroSim PSpice for Windows: Volume I: DC, AC, and Devices and Circuits*. 2nd ed. Upper Saddle River, NJ: Prentice-Hall, Inc., 1998.
- Goody, R. W. *MicroSim PSpice for Windows: Volume II: Operational Amplifiers and Digital Circuits*. 2nd ed. Upper Saddle River, NJ: Prentice-Hall, 1998.
- Herniter, M. E. *Schematic Capture with MicroSim PSpice*. 3rd ed. Upper Saddle River, NJ: Prentice-Hall, 1998.
- Meares, L. G.; and C. E. Hymowitz. *Simulating with Spice*. San Pedro, CA: Intusoft, 1988.
- MicroSim Staff. *PSpice User's Manual Version 4.03*. Irvine, CA: MicroSim Corporation, 1990.
- Natarajan, S. "An Effective Approach to Obtain Model Parameters for BJTs and FETs from Data Books." *IEEE Transactions on Education* 35, No. 2 (May 1992), pp. 164–69.
- Rashid, M. H. *SPICE for Circuits and Electronics Using PSpice*. Englewood Cliffs, NJ: Prentice Hall, Inc., 1990.
- Roberts, G. W.; and A. S. Sedra. *SPICE for Microelectronic Circuits*. 3rd ed. New York: Saunders College Publishing, 1992.
- Thorpe, T. W. *Computerized Circuit Analysis with SPICE*. New York: John Wiley and Sons, Inc., 1992.
- Tront, J. G. *PSpice for Basic Microelectronics with CD*. New York: McGraw-Hill, 2008.
- Tuinenga, P. W. *SPICE: A Guide to Circuit Simulation and Analysis Using PSpice*. 2nd ed. Englewood Cliffs, NJ: Prentice Hall, Inc., 1992.

Answers to Selected Problems

APPENDIX E

CHAPTER 1

- 1.1 (a) Silicon
 (i) $n_i = 1.61 \times 10^8 \text{ cm}^{-3}$
 (ii) $n_i = 3.97 \times 10^{11} \text{ cm}^{-3}$
- (b) GaAs
 (i) $n_i = 6.02 \times 10^3 \text{ cm}^{-3}$
 (ii) $n_i = 1.09 \times 10^8 \text{ cm}^{-3}$
- 1.3 Silicon
 (a) $n_i = 8.79 \times 10^{-10} \text{ cm}^{-3}$
 (b) $n_i = 1.5 \times 10^{10} \text{ cm}^{-3}$
 (c) $n_i = 1.63 \times 10^{14} \text{ cm}^{-3}$
- Germanium
 (a) $n_i = 35.9 \text{ cm}^{-3}$
 (b) $n_i = 2.40 \times 10^{13} \text{ cm}^{-3}$
 (c) $n_i = 8.62 \times 10^{15} \text{ cm}^{-3}$
- 1.5 (a) p-type, $p_o = 10^{16} \text{ cm}^{-3}$
 $n_o = 3.24 \times 10^{-4} \text{ cm}^{-3}$
 (b) p-type, $p_o = 10^{16} \text{ cm}^{-3}$
 $n_o = 5.76 \times 10^{10} \text{ cm}^{-3}$
- 1.7 (a) p-type, $p_o = 5 \times 10^{16} \text{ cm}^{-3}$
 $n_o = 4.5 \times 10^3 \text{ cm}^{-3}$
 (b) p-type, $p_o = 5 \times 10^{16} \text{ cm}^{-3}$
 $n_o = 6.48 \times 10^{-5} \text{ cm}^{-3}$
- 1.9 (a) $n_o = 5 \times 10^{15} \text{ cm}^{-3}$
 $p_o = 4.5 \times 10^4 \text{ cm}^{-3}$
 (b) n-type
 $N_d = 5 \times 10^{15} \text{ cm}^{-3}$
- 1.11 (a) $I = 0.15 \text{ mA}$
 (b) $E = 2.4 \text{ V/cm}$
- 1.13 (a) $N_d = 7.69 \times 10^{15} \text{ cm}^{-3}$
 (b) $E = 104 \text{ V/cm}$
- 1.15 (a) $1.36 \leq \sigma \leq 1.36 \times 10^4 (\Omega\text{-cm})^{-1}$
 (b) $0.136 \leq J \leq 1.36 \times 10^3 \text{ A/cm}^2$
- 1.17 (a) $J_p = 2.4 \text{ A/cm}^2$
 (b) $J_p = 0.883 \text{ A/cm}^2$
 (c) $J_p = 0.119 \text{ A/cm}^2$
- 1.19 (a) (i) 0.661 V, (ii) 0.739 V, (iii) 0.937 V
 (b) (i) 1.13 V, (ii) 1.21 V, (iii) 1.41 V
- 1.21 For $N_a = 10^{15} \text{ cm}^{-3}$, $V_{bi} = 0.637 \text{ V}$
 For $N_a = 10^{18} \text{ cm}^{-3}$, $V_{bi} = 0.817 \text{ V}$
- 1.23 (a) $C_j = 0.255 \text{ pF}$
 (b) $C_j = 0.172 \text{ pF}$
 (c) $C_j = 0.139 \text{ pF}$
- 1.25 (a) $f_o = 6.57 \text{ MHz}$
 (b) $f_o = 7.95 \text{ MHz}$
 (c) $f_o = 8.86 \text{ MHz}$
- 1.27 (a) (i) $1.03 \mu\text{A}$, (ii) 2.25 mA ,
 (iii) 4.93 A , (iv) $-5.37 \times 10^{-12} \text{ A}$,
 (v) -10^{-11} A , (vi) -10^{-11} A
 (b) (i) $0.0103 \mu\text{A}$, (ii) $22.5 \mu\text{A}$,
 (iii) 49.3 mA , (iv) $-5.37 \times 10^{-14} \text{ A}$,
 (v) -10^{-13} A , (vi) -10^{-13} A
- 1.29 (a) $I_S = 2.03 \times 10^{-15} \text{ A}$
 (b) (i) At $V_D = 0.1 \text{ V}$,
 $I_D = 9.50 \times 10^{-14} \text{ A}$
 At $V_D = 0.7 \text{ V}$, $I_D = 10^{-3} \text{ A}$
 (ii) At $V_D = 0.1 \text{ V}$,
 $I_D = 1.39 \times 10^{-14} \text{ A}$
 At $V_D = 0.7 \text{ V}$, $I_D = 1.42 \times 10^{-9} \text{ A}$
- 1.31 (a) $\Delta V_D \cong 60 \text{ mV}$
 (b) $\Delta V_D \cong 120 \text{ mV}$
- 1.33 $0.5150 \leq V_D \leq 0.6347 \text{ V}$
- 1.35 (a) 2.31 nA , $5.05 \mu\text{A}$, 11.1 mA ,
 $-5.37 \times 10^{-23} \text{ A}$, -10^{-22} A , -10^{-22} A
 (b) 115 pA , $0.253 \mu\text{A}$, 0.554 mA ,
 $-2.68 \times 10^{-24} \text{ A}$, $-5 \times 10^{-24} \text{ A}$,
 $-5 \times 10^{-24} \text{ A}$
- 1.37 $\frac{I_D(100)}{I_D(-55)} = 2.83 \times 10^3$
- 1.39 $I_D = 0.4740 \text{ mA}$, $V_D = 0.5194 \text{ V}$
- 1.41 (a) $I_{D1} = I_{D2} = 1 \text{ mA}$,
 (i) $V_{D1} = V_{D2} = 0.599 \text{ V}$
 (ii) $V_{D1} = 0.617 \text{ V}$, $V_{D2} = 0.557 \text{ V}$
 (b) (i) $V_{D1} = V_{D2} = 0.581 \text{ V}$,
 $I_{D1} = I_{D2} = 0.5 \text{ mA}$
 (ii) $V_{D1} = V_{D2} = 0.554 \text{ V}$,
 $I_{D1} = 0.0909 \text{ mA}$, $I_{D2} = 0.909 \text{ mA}$

- 1.43 (a) $V_D = 0.7 \text{ V}$
 $I_D = 26.7 \mu\text{A}$
- (b) $V_D = 0.45 \text{ V}$
 $I_D = 0$
- 1.45 (a) $V_O = I_I \cdot (1), I_D = 0;$
for $0 \leq I_I \leq 0.7 \text{ mA}$
 $V_O = 0.7 \text{ V}, I_D = (I_I - 0.7) \text{ mA};$
for $I_I \geq 0.7 \text{ mA}$
- (b) $V_O = I_I \cdot (1), I_D = 0;$
for $0 \leq I_I \leq 1.7 \text{ mA}$
 $V_O = 1.7 \text{ V}, I_D = (I_I - 1.7) \text{ mA};$
for $I_I \geq 1.7 \text{ mA}$
- (c) $V_O = 0.7 \text{ V}, I_{D1} = I_I \text{ and } I_{D2} = 0;$
for $0 \leq I_I \leq 2 \text{ mA}$
- 1.47 (a) (i) $I = 0.215 \text{ mA}, V_O = 0.7 \text{ V}$
(ii) $I = 0.220 \text{ mA}, V_O = 0.6 \text{ V}$
- (b) (i) $I = 0.2325 \text{ mA}, V_O = -0.35 \text{ V}$
(ii) $I = 0.235 \text{ mA}, V_O = -0.30 \text{ V}$
- (c) (i) $I = 0.372 \text{ mA}, V_O = 0.14 \text{ V}$
(ii) $I = 0.376 \text{ mA}, V_O = 0.12 \text{ V}$
- (d) (i) $I = 0, V_O = -5 \text{ V}$
(ii) $I = 0, V_O = -5 \text{ V}$
- 1.49 (a) $V_D = 0.7 \text{ V}, V = 2.42 \text{ V}$
(b) $P = 0.28 \text{ mW}$
- 1.51 (a) $v_d = 1.30 \text{ mV}$
(b) $v_d = 1.30 \text{ mV}$
- 1.53 (b) For $I = 1 \text{ mA}, v_o/v_s = 0.0909$
For $I = 0.1 \text{ mA}, v_o/v_s = 0.50$
For $I = 0.01 \text{ mA}, v_o/v_s = 0.909$
- 1.55 $I_S = 4.87 \times 10^{-12} \text{ A}$
- 1.57 (a) $V_O = 5.685 \text{ V}$
(b) $\Delta V_O = 0.0392 \text{ V}$
(c) $V_O = 5.658 \text{ V}$
- 1.59 (a) $V_O = 6.921 \text{ V}$
(b) $\Delta V_O = -0.13 \text{ V}$
- 1.61 $V_D = 0, I_D = 0.2 \text{ A}$
 $V_D = 0.60, I_D = 0.1995 \text{ A}$
 $V_D = 0.70, I_D = 0.1754 \text{ A}$
 $V_D = 0.7545, I_D = 0$
- CHAPTER 2**
- 2.1 (a) $v_O = 0 \text{ for } v_I \leq 0.6 \text{ V}$
 $v_O = \left(\frac{1000}{1020}\right)(v_I - 0.6)$
for $v_I \geq 0.6 \text{ V}$
- (b) (ii) $v_O(\text{avg}) = 2.89 \text{ V}$
(iii) $i_d(\text{max}) = 9.216 \text{ mA}$
(iv) $\text{PIV} = 10 \text{ V}$
- 2.3 (a) $v_O(\text{peak}) = 16.27 \text{ V}$
(b) $i_D(\text{peak}) = 8.14 \text{ mA}$
(c) 48.7%
(d) $v_O(\text{avg}) = 5.06 \text{ V}$
(e) $i_d(\text{avg}) = 2.53 \text{ mA}$
- 2.5 (a) $R = 4.417 \Omega$
(b) $i_D(\text{avg}) = 0.436 \text{ A}$
(c) 27.6%
- 2.7 (a) $\frac{N_1}{N_2} = 6.06$
(b) $\frac{N_1}{N_2} = 1.58$
(c) For (a), PIV = 52.1 V
For (b), PIV = 202.1 V
- 2.9 (a) $v_S(\text{rms}) = 8.98 \text{ V}$
(b) $C = 4444 \mu\text{F}$
(c) $i_D(\text{max}) = 4.58 \text{ A}$
- 2.11 (a) $10.3 \leq v_O \leq 12.3 \text{ V}$
(b) $0.490 \leq V_r \leq 0.586 \text{ V}$
(c) $C = 513 \mu\text{F}$
- 2.13 (a) $v_S(\text{rms}) = 11.6 \text{ V}$
(b) $C = 2857 \mu\text{F}$
- 2.15 (a) $\frac{N_1}{N_2} = 13.3$
(b) $C = 0.0116 \text{ F}$
(c) $i_D(\text{peak}) = 13.3 \text{ A}$
(d) $i_D(\text{avg}) = 0.539 \text{ A}$
(e) PIV = 24.8 V
- 2.19 (a) $I_Z = 0.367 \text{ mA}, I_L = 0.975 \text{ mA},$
 $P = 1.43 \text{ mW}$
- (b) $I_Z = 0.952 \text{ mA}, I_L = 0.39 \text{ mA},$
 $P = 3.71 \text{ mW}$
- 2.21 (a) $15 \leq I_Z \leq 259.74 \text{ mA}$
(b) $54.5 \leq R_L \leq 410 \Omega$
- 2.23 (a) $R_i = 8.08 \Omega$
(b) $P_Z = 11.9 \text{ W}, P_L = 5 \text{ W}$
- 2.25 (a) $R_i = 257 \Omega$
(b) $7.450 \leq V_L \leq 7.556 \text{ V}; 4.42\%$
(c) $7.50 \leq V_L \leq 7.586 \text{ V}; 1.15\%$
- 2.27 (a) $R_i = 200 \Omega$
(b) $\Delta V_O = 0.35 \text{ V}$
(c) 3.5%
- 2.29 5.0%, $C = 0.0357 \text{ F}$
- 2.31 (a) $v_O = 0 \text{ for } -10 \leq v_I \leq 3 \text{ V}$
 $v_O = [(0.5)v_I - 1.5] \text{ V for } v_I \geq 3 \text{ V}$

- (b) $-i_1 = \frac{0 - v_I}{10}$ mA for $v_I \leq 0$
 $i_1 = 0$ for $0 \leq v_I \leq 3$ V
 $i_1 = \frac{v_I - 3}{20}$ mA for $v_I \geq 3$ V
- 2.33 (a) (i) $v_O = v_I$ for $v_I \geq 1.1$ V
 $v_O = 1.1$ V for $v_I \leq 1.1$ V
(ii) $v_O = v_I$ for $v_I \geq -2.5$ V
 $v_O = -2.5$ V for $v_I \leq -2.5$ V
(b) (i) $v_O = 2.5$ V for $v_I \geq 2.5$ V
 $v_O = v_I$ for $v_I \leq 2.5$ V
(ii) $v_O = -1.1$ V for $v_I \geq -1.1$ V
 $v_O = v_I$ for $v_I \leq -1.1$ V
- 2.35 (a) (i) $v_O = v_I - 5.7$ V for $v_I \geq 5.7$ V
 $v_O = 0$ for $v_I \leq 5.7$ V
(ii) $v_O = v_I + 4.3$ V for $v_I \geq -4.3$ V
 $v_O = 0$ for $v_I \leq -4.3$ V
(b) (i) $v_O = 0$ for $v_I \geq 4.3$ V
 $v_O = v_I - 4.3$ V for $v_I \leq 4.3$ V
(ii) $v_O = 0$ for $v_I \geq -5.7$ V
 $v_O = v_I + 5.7$ V for $v_I \leq -5.7$ V
- 2.37 $v_O(\max) = 6.7$ V, $v_O(\min) = -4.7$ V
- 2.39 (a) Square wave between $+40$ V and 0.
(b) Square wave between $+35$ V and -5 V.
(c) Square wave between $+5$ V and -35 V.
- 2.41 Circuit similar to Figure 2.31(a)
with $V_B = -10$ V.
- 2.43 (i) $v_O = (10 \sin \omega t - 5)$ V
(ii) $v_O = (10 \sin \omega t - 15)$ V
- 2.45 (a) $I = I_{D1} = I_{D2} = 0$, $V_O = 10$ V
(b) $I = I_{D2} = 0.94$ mA,
 $I_{D1} = 0$, $V_O = 1.07$ V
(c) $I = I_{D2} = 0.44$ mA,
 $I_{D1} = 0$, $V_O = 5.82$ V
(d) $I = 0.964$ mA, $V_O = 0.842$ V,
 $I_{D1} = I_{D2} = 0.482$ mA
- 2.47 (a) $R_1 = 25$ k Ω , $R_2 = 10$ k Ω , $R_3 = 4.4$ k Ω
(b) $V_1 = 4.4$ V, $V_2 = -0.6$ V, $I_{D1} = 0.5$ mA,
 $I_{D2} = 0.75$ mA, $I_{D3} = 0.75$ mA
(c) $V_1 = 6.07$ V, $V_2 = -0.6$ V,
 $I_{D1} = 1.11$ mA,
 $I_{D2} = 0$, $I_{D3} = 0.65$ mA
(d) $V_1 = 4.4$ V, $V_2 = 1.27$ V,
 $I_{D1} = 0.833$ mA,
 $I_{D2} = 0.211$ mA, $I_{D3} = 0$
- 2.49 (a) $I_{D1} = 1.0$ mA, $I_{D2} = 6.0$ mA,
 $V_A = 2.30$ V
(b) $I_{D1} = 0$, $I_{D2} = 5$ mA, $V_A = 8.2$ V
- (c) $I_{D1} = 2.15$ mA, $I_{D2} = 7.15$ mA,
 $R_2 = 0.60$ k Ω
- 2.51 $v_O = v_I$ for $-4.65 \leq v_I \leq 4.65$ V,
 $v_O = 4.65$ V for $v_I \geq 4.65$ V,
 $v_O = -4.65$ V for $v_I \leq -4.65$ V
- 2.53 (a) $V_O = 0$, $I_{D1} = 0.86$ mA
(b) $I_{D1} = 0$, $V_O = -3.57$ V
- 2.55 (a) $I_{D1} = 0.93$ mA, $V_O = -15$ V
(b) $I_{D1} = 1.86$ mA, $V_O = -15$ V
- 2.57 (a) $I_D = 0$, $V_D = 0$, $V_A = V_B = 3$ V
(b) $I_D = 0.08$ mA, $V_D = 0.7$ V,
 $V_A = 1.4$ V, $V_B = 2.1$ V
(c) $I_D = 0$, $V_D = -0.5$ V, $V_A = 2.5$ V,
 $V_B = 2.0$ V
(d) $I_D = 0.23$ mA, $V_D = 0.7$ V,
 $V_A = 2.15$ V, $V_B = 2.85$ V
- 2.59 (a) $I_{D1} = I_{D2} = I_{D3} = 0$, $v_O = 0.5$ V
(b) $I_{D1} = 0.0667$ mA, $I_{D2} = I_{D3} = 0$,
 $v_O = 1.23$ V
(c) $I_{D1} = 0.171$ mA, $I_{D2} = 0.0615$ mA,
 $I_{D3} = 0$, $v_O = 2.069$ V
(d) $I_{D1} = 0.275$ mA, $I_{D2} = 0.20$ mA,
 $I_{D3} = 0.05$ mA, $v_O = 2.90$ V
- 2.61 (a) $V_{O1} = V_{O2} = 0$
(b) $V_{O1} = 4.4$ V, $V_{O2} = 3.8$ V
(c) $V_{O1} = 4.4$ V, $V_{O2} = 3.8$ V
Logic 1 level degrades
- 2.63 $(V_1 \text{ AND } V_2) \text{ OR } (V_3 \text{ AND } V_4)$
- 2.65 $V_I = 2.3$ V
- 2.67 $V_{PS} = 2.6$ V

CHAPTER 3

- 3.1 (a) (i) $I_D = 0$
(ii) $I_D = 82.5$ μ A
(iii) $I_D = 0.2325$ mA
(iv) $I_D = 0.3825$ mA
(b) (i) $I_D = 0$
(ii) $I_D = 0.27$ mA
(iii) $I_D = 1.92$ mA
(iv) $I_D = 5.07$ mA
- 3.3 (a) Enhancement-mode
(b) $V_{TN} = 1.5$ V, $K_n \cong 0.064$ mA/V 2
(c) $i_D(\text{sat}) = 0.256$ mA, for $v_{GS} = 3.5$ V
 $i_D(\text{sat}) = 0.576$ mA, for $v_{GS} = 4.5$ V
- 3.5 (a) Saturation
(b) Nonsaturation
(c) Cutoff
- 3.7 $\frac{W}{L} = 5.79$

- 3.9 (a) $K_n = 1.40 \text{ mA/V}^2$
 (b) $I_D = 3.58 \text{ mA}$
 (c) $V_{DS}(\text{sat}) = 1.6 \text{ V}$
- 3.11 $W = 3.22 \mu\text{m}$
- 3.13 $V_{TP} = -0.571 \text{ V}, \frac{W}{L} = 4.41$
- 3.15 (a) Enhancement-mode
 (b) $V_{TP} \approx -0.5 \text{ V}, K_p \approx 0.20 \text{ mA/V}^2$
 (c) $i_D(\text{sat}) = 1.8 \text{ mA}, i_D(\text{sat}) = 3.2 \text{ mA}$
- 3.17 (a) $I_D = 0.21 \text{ mA}$
 (b) $I_D = 0.66 \text{ mA}$
 (c) $I_D = 0.81 \text{ mA}$
 (d) $I_D = 0.844 \text{ mA}$
 (e) $I_D = 0.844 \text{ mA}$
- 3.19 $K_n = K_p = 41.3 \mu\text{A/V}^2$,
 $W_n = 7.09 \mu\text{m}, W_p = 12.8 \mu\text{m}$
- 3.21 $\lambda = 0.00926 \text{ V}^{-1}, V_A = 108 \text{ V}$
- 3.23 (a) $I_D = 0.739 \text{ mA}$
 (b) $I_D = 0.296 \text{ mA}$
- 3.25 $t_{ox} = 1200 \text{ \AA}$
- 3.27 (a) $V_{DS} = 1.88 \text{ V}, I_D = 2.12 \text{ mA}$
 (b) $V_{DS} = 0.741 \text{ V}, I_D = 1.42 \text{ mA}$
- 3.29 (a) $V_{SD} = 1.90 \text{ V}, I_D = 1.33 \text{ mA}$
 (b) $V_{SD} = 0.698 \text{ V}, I_D = 1.08 \text{ mA}$
- 3.31 $V_S = 2.21 \text{ V}, V_{SD} = 5.21 \text{ V}$
- 3.33 $\frac{W}{L} = 20.8, R_1 = 360 \text{ k}\Omega, R_2 = 450 \text{ k}\Omega$
- 3.35 $V_{GS} = 1.71 \text{ V}, I_D = 2.58 \text{ mA},$
 $V_{DS} = 5.62 \text{ V}$
- 3.37 (a) (i) $V_{GS} = 1.516 \text{ V}, V_{DS} = 6.516 \text{ V}$
 (ii) $V_{GS} = 2.61 \text{ V}, V_{DS} = 7.61 \text{ V}$
 (b) (i) $V_{GS} = V_{DS} = 1.516 \text{ V}$
 (ii) $V_{GS} = V_{DS} = 2.61 \text{ V}$
- 3.39 (a) $R_D = 8 \text{ k}\Omega, R_S = 4.38 \text{ k}\Omega$
 (b) Let $R_D = 8.2 \text{ k}\Omega, R_S = 4.3 \text{ k}\Omega$
 Then $V_{GS} = 2.82 \text{ V}$,
 $I_D = 0.504 \text{ mA}, V_{DS} = 3.70 \text{ V}$
 (c) For $R_S = 4.73 \text{ k}\Omega, I_D = 0.476 \text{ mA}$
 For $R_S = 3.87 \text{ k}\Omega, I_D = 0.548 \text{ mA}$
- 3.41 $I_{DQ} = 1.25 \text{ mA}, R_2 = 59.4 \text{ k}\Omega,$
 $R_1 = 20.6 \text{ k}\Omega$
- 3.43 $R_D = 0.8 \text{ k}\Omega, R_1 = 408 \text{ k}\Omega, R_2 = 99.5 \text{ k}\Omega$
- 3.45 $\left(\frac{W}{L}\right)_1 = 2.78$
- 3.47 (a) $\left(\frac{W}{L}\right)_1 = 2.31, \left(\frac{W}{L}\right)_2 = 1.59,$
 $\left(\frac{W}{L}\right)_3 = 3.69$
- (b) (i) $V_1 = 2.5 \text{ V}, V_2 = 6 \text{ V}$
 (ii) $V_1 = 2.5 \text{ V}, V_2 = 6 \text{ V}$
 (c) $V_1 = 2.469 \text{ V}, V_2 = 5.916 \text{ V}$
- 3.49 $\left(\frac{W}{L}\right)_D = 3.31$
- 3.51 $R_D = 271 \Omega, \frac{W}{L} = 231$
- 3.53 (a) $\frac{W}{L} = 0.623$
 (b) $V_O = 0.297 \text{ V}$
- 3.55 (a) $\left(\frac{W}{L}\right)_A = \left(\frac{W}{L}\right)_B = 12.5,$
 $\left(\frac{W}{L}\right)_C = 7.81$
 (b) $R_D = 11 \text{ k}\Omega$
- 3.57 $I_D = I_{DSS}$
- 3.59 $V_P = 3.97 \text{ V}, I_{DSS} = 5.0 \text{ mA}$
- 3.61 $V_{TN} = 0.221 \text{ V}, K_n = 1.11 \text{ mA/V}^2$
- 3.63 $V_{GS} = -1.17 \text{ V}, I_D = 5.85 \text{ mA},$
 $V_{DS} = 7.13 \text{ V}$
- 3.65 $V_{GS} = 0.838 \text{ V}, V_{SD} = 7.5 \text{ V},$
 $R_1 = 109 \text{ k}\Omega, R_2 = 1.21 \text{ M}\Omega$
- 3.67 $V_G = 6 \text{ V}, V_{GS} = -1.30 \text{ V},$
 $I_D = 3.65 \text{ mA}, V_{DS} = 2.85 \text{ V}$
- 3.69 $I_D = I_{DSS} = 4 \text{ mA},$
 $R_D = 1.75 \text{ k}\Omega$
- 3.71 $R_D = 15 \text{ k}\Omega, R_1 = 100 \text{ k}\Omega$
 $R_2 = 50 \text{ k}\Omega$

CHAPTER 4

- 4.1 (a) (i) $\frac{W}{L} = 2.5$
 (ii) $V_{GSQ} = 2.4 \text{ V}$
 (b) (i) $\frac{W}{L} = 8.33$
 (ii) $V_{GSQ} = 1.0 \text{ V}$
- 4.3 $\lambda = 0.0308 \text{ V}^{-1}, r_o = 12.5 \text{ k}\Omega$
- 4.5 (a) $\lambda = 0.01826 \text{ V}^{-1}, r_o = 225 \text{ k}\Omega$
 (b) $I_D = 0.2655 \text{ mA}$
- 4.7 $I_D = 0.25 \text{ mA}, r_o = 400 \text{ k}\Omega$
- 4.9 (a) $\frac{W}{L} = 6.02$
 (b) $\frac{W}{L} = 10.4$
- 4.13 (a) $I_{DQ} = 0.270 \text{ mA}, V_{DSQ} = 1.14 \text{ V}$
 (b) $g_m = 2.078 \text{ mA/V}, r_o = 185 \text{ k}\Omega$
 (c) $A_v = -15.3$

- 4.15 (a) $I_{DQ} = 0.608 \text{ mA}$, $V_{DSQ} = 1.96 \text{ V}$
 (b) $A_v = -2.44$
 (c) $R_L = 12 \text{ k}\Omega$
- 4.17 (a) From Problem 4.16, $R_S = 0.5 \text{ k}\Omega$,
 $R_D = 3 \text{ k}\Omega$, $R_1 = 894 \text{ k}\Omega$, $R_2 = 347 \text{ k}\Omega$
 (b) $A_v = -7.99$
- 4.19 (a) $g_m = 1 \text{ mA/V}$, $R_D = 16 \text{ k}\Omega$
 (b) $R_S = 0.6 \text{ k}\Omega$
- 4.21 (a) $R_S = 38.6 \text{ k}\Omega$, $R_D = 6.43 \text{ k}\Omega$
 (b) $g_m = 0.583 \text{ mA/V}$, $r_o = 500 \text{ k}\Omega$
 (c) $A_v = -3.19$
- 4.23 (a) $R_S = 9 \Omega$, $I_{DQ} = 1.244 \text{ mA}$
 (b) $A_v = -3.10$
- 4.25 (a) $R_S = 7.76 \text{ k}\Omega$, $R_D = 10.2 \text{ k}\Omega$
 (b) $A_v = -1.50$
- 4.27 (a) $I_Q = 1.518 \text{ mA}$
 (b) $A_v = -11.8$
 (c) $A_v = -8.29$
- 4.29 Let $R_D = 8 \text{ k}\Omega$, $R_S = 2 \text{ k}\Omega$,
 Then, $R_1 = 778 \text{ k}\Omega$, $R_2 = 269 \text{ k}\Omega$
- 4.31 $g_m = 0.98 \text{ mA/V}$, $r_o = 50 \text{ k}\Omega$
- 4.33 (a) $I_{DQ} = 0.15 \text{ mA}$, $V_{GSQ} = 0.594 \text{ V}$
 (b) $A_v = 0.938$
 (c) $R_o = 605 \Omega$
- 4.35 (a) $I_Q = 1 \text{ mA}$
 (b) $\frac{W}{L} = 19.6$
 (c) $A_v = 0.990$
 (d) $R_o = 223 \Omega$
- 4.37 (a) (i) $A_v = 0.927$
 (ii) $R_o = 219 \Omega$
 (b) (i) $A_v = 0.907$
 (ii) $R_o = 349 \Omega$
- 4.39 $R_S = 1.2 \text{ k}\Omega$, $A_v = 0.878$, $R_o = 91.9 \Omega$
- 4.41 Set $g_m = 5 \text{ mA/V}$, $R_S = 4.09 \text{ k}\Omega$,
 $A_v = 0.870$
- 4.45 $\frac{W}{L} = 80$, $R_D = 10 \text{ k}\Omega$
- 4.47 $A_v = 11.3$, $R_i = 342 \Omega$
- 4.49 (a) $R_S = 3.70 \text{ k}\Omega$, $R_D = 1.63 \text{ k}\Omega$
 (b) $R_i = 0.816 \text{ k}\Omega$, $R_o = 1.63 \text{ k}\Omega$
 (c) $i_o = 1.84 \sin \omega t (\mu\text{A})$,
 $v_o = 3.68 \sin \omega t (\text{mV})$
- 4.51 (a) $R_D = 0.554 \text{ k}\Omega$
 (b) $g_m = 5.657 \text{ mA/V}$, $R_i = 177 \Omega$
 (c) $A_v = 2.75$
- 4.53 (a) $\left(\frac{W}{L}\right)_D = 25$, $V_{GSDQ} = 0.825 \text{ V}$
 (b) $I_{DQ} = 0.0633 \text{ mA}$,
 $V_{DSDQ} = V_O = 1.575 \text{ V}$
- 4.55 (a) $r = 900 \Omega$
 (b) $I_D = 3.56 \text{ mA}$, $r = 337 \Omega$
- 4.57 (a) $V_{GG} = 5.95 \text{ V}$
 (c) $A_v = 0.691$
- 4.59 (a) $\left(\frac{W}{L}\right)_D = 14.4$
 (b) $V_B = 1.393 \text{ V}$
 (c) $V_{GSDQ} = 1.233 \text{ V}$
- 4.61 $A_v = -73.7$
- 4.63 (a) $V_O = \frac{2}{3}V_I - \frac{1}{3}(0.4)$
 (b) $I_D = 0.5 \left[\frac{2}{3}V_I - 0.533 \right]^2$
 (c) $A_v = 0.667$
- 4.65 (a) $V_{GSDQ} = 1.307 \text{ V}$, $V_{SGLQ} = 2.014 \text{ V}$,
 $V_{DSDQ} = 2.593 \text{ V}$
 (b) $A_v = \sqrt{\frac{K_n}{K_p}}$
 (c) $A_v = 2$
- 4.67 (a) $g_{m1} = g_{m2} = 0.8944 \text{ mA/V}$,
 $r_{o1} = r_{o2} = 100 \text{ k}\Omega$
 (b) $A_v = -44$
 (c) $R_o = 49.7 \text{ k}\Omega$
- 4.69 (a) $R_1 = 586 \text{ k}\Omega$, $R_2 = 436 \text{ k}\Omega$,
 $R_{D1} = 20 \text{ k}\Omega$, $R_{S2} = 3.6 \text{ k}\Omega$,
 $R_{D2} = 2.4 \text{ k}\Omega$
 (b) $A_v = 13.6$
- 4.71 (a) $I_{DQ1} = 0.757 \text{ mA}$, $I_{DQ2} = 0.757 \text{ mA}$,
 $V_{DSQ1} = 12.4 \text{ V}$, $V_{DSQ2} = 8.65 \text{ V}$
 (b) $g_{m1} = g_{m2} = 3.48 \text{ mA/V}$
 (c) $A_v = 2.42$
- 4.73 (a) $R_1 = 38.8 \text{ k}\Omega$, $R_2 = 35 \text{ k}\Omega$,
 $R_3 = 26.2 \text{ k}\Omega$, $R_D = 0.6 \text{ k}\Omega$
 (b) $A_v = -5.37$
- 4.75 $g_m = 1.569 \text{ mA/V}$, $A_v = -3.62$, $A_i = -45.2$
- 4.77 (a) $R_S = 1.2 \text{ k}\Omega$, $R_1 = 265 \text{ k}\Omega$,
 $R_2 = 161 \text{ k}\Omega$
 (b) $A_v = 0.495$, $R_o = 0.273 \text{ k}\Omega$
- 4.79 $R_S = 0.516 \text{ k}\Omega$, $R_D = 2.61 \text{ k}\Omega$,
 $R_1 = 65 \text{ k}\Omega$, $R_2 = 335 \text{ k}\Omega$

CHAPTER 5

- 5.1 (a) $\beta = 115$, $\alpha = 0.9914$, $i_C = 322 \mu\text{A}$
 (b) $\beta = 89$, $\alpha = 0.9889$, $i_C = 1.78 \text{ mA}$
- 5.3 (a) $0.99099 \leq \alpha \leq 0.99448$
 (b) $5.50 \leq I_C \leq 9.0 \text{ mA}$

- 5.5 (a) $\beta = 9, 19, 49, 99, 199, 999$
 (b) $\alpha = 0.9524, 0.9804, 0.9901,$
 $0.9934, 0.9955, 0.9975$
- 5.7 $V_{BE} = 0.6109 \text{ V}, I_C = 0.7928 \text{ mA},$
 $I_B = 7.21 \mu\text{A}, V_C = 3.41 \text{ V}$
- 5.9 (a) $I_E = 1.785 \text{ mA}, I_B = 25 \mu\text{A}$
 (b) $V_{EB} = 0.7154 \text{ V}$
- 5.11 $I_{EO1} = 6.94 \times 10^{-15} \text{ A},$
 $I_{EO2} = 1.69 \times 10^{-13} \text{ A}, \text{ Ratio} = 24.4$
- 5.13 (a) $I_C = 0.622 \text{ mA}$
 (b) $r_o = 137 \text{ k}\Omega$
- 5.15 $\beta = 60.6$
- 5.17 (a) $I_C = 1.836 \text{ mA}, R_C = 3.65 \text{ k}\Omega$
 (b) $V_B = 0.164 \text{ V}, R_C = 8.11 \text{ k}\Omega$
 (c) $I_C = 1.744 \text{ mA}, V_{CE} = 1.96 \text{ V}$
 (d) $I_B = 4.61 \mu\text{A}, V_C = 1.49 \text{ V}$
- 5.19 (a) $V_{BB} = 0.7091 \text{ V}$
 (b) $V_{BB} = 2.127 \text{ V}$
- 5.21 (a) $I_C = 0.7273 \text{ mA}, V_{EC} = 0.9 \text{ V}$
 (b) $I_C = 1.185 \text{ mA}, V_{EC} = 0.2 \text{ V}$
 (c) $I_C = 0, V_{EC} = 2 \text{ V}$
- 5.23 (a) $\beta = 134.4, \alpha = 0.9926,$
 $I_C = 0.269 \text{ mA}, V_{CE} = 4.7 \text{ V}$
 (b) $\beta = 10.63, \alpha = 0.9140$
- 5.25 (a) $I_E = 0.245 \text{ mA}, I_C = 0.242 \text{ mA},$
 $\beta = 80.67, \alpha = 0.9878, V_{EC} = 1.73 \text{ V}$
 (b) $V_E = 0.8371 \text{ V}, V_{EC} = 1.70 \text{ V}$
- 5.27 $V_C = -4.39 \text{ V}, V_E = 1.68 \text{ V}$
- 5.29 $I_{E1} = I_{E2} = 0.5 \text{ mA}, V_{C1} = V_{C2} = 3 \text{ V}$
- 5.31 (a) $R_C = 18 \text{ k}\Omega, R_B = 2.656 \text{ M}\Omega$
 (b) $I_{CQ} = 0.375 \text{ mA}, V_{CEQ} = 2.25 \text{ V}$
- 5.33 (a) (i) $V_O = 3.33 \text{ V},$ (ii) $V_O = 1.83 \text{ V},$
 (iii) $V_O = 0.2 \text{ V}$
- 5.35 (a) $I_Q = 0.2017 \text{ mA}$
 (b) $I_Q = 0.605 \text{ mA}$
 (c) $I_Q = 1.008 \text{ mA}$
- 5.37 $I_E = 2.075 \text{ mA}, I_C = 2.06 \text{ mA},$
 $V_{BC} = 4.47 \text{ V}$
- 5.39 $1.86 \leq V_1 \leq 3.96 \text{ V}$
- 5.41 $V_O = V_T \ln(10)$
- 5.43 $V_O = 0 \text{ for } 4.3 \leq V_I \leq 5 \text{ V},$
 $V_O = 4.8 \text{ V for } 0 \leq V_I \leq 1.6 \text{ V}$
- 5.45 $R_B = 3.58 \text{ k}\Omega$
- 5.47 $R_B = 4.69 \text{ k}\Omega$
- 5.49 $R_1 = 399 \text{ k}\Omega, R_2 = 401 \text{ k}\Omega,$
 $V_{CEQ} = 1.30 \text{ V}$
- 5.51 $I_{EQ} = 4.62 \text{ mA}, V_B = 4.68 \text{ V}$
- 5.53 (a) $I_{BQ} = 10.98 \mu\text{A}, I_{CQ} = 0.8782 \text{ mA},$
 $V_{CEQ} = 3.50 \text{ V}$
 (b) For $I_{CQ}: +7.21\%;$ for $V_{CEQ}: -11.2\%$
- 5.55 (a) $R_{TH} = 1.714 \text{ k}\Omega, V_{TH} = -3.571 \text{ V}$
 (b) $I_{CQ} = 1.41 \text{ mA}, V_{CEQ} = 2.32 \text{ V}$
 (d) $1.332 \leq I_{CQ} \leq 1.467 \text{ mA}$
 $1.59 \leq V_{CEQ} \leq 2.97 \text{ V}$
- 5.57 (a) $I_{CQ} = 0.0888 \text{ mA}, V_{CEQ} = 3.55 \text{ V}$
 (b) $I_{CQ} = 0.266 \text{ mA}, V_{CEQ} = 3.56 \text{ V}$
- 5.59 (a) $R_1 = 44.1 \text{ k}\Omega, R_2 = 5.70 \text{ k}\Omega,$
 $R_C = 5.75 \text{ k}\Omega$
 (b) 6.75%
 (c) $R_1 = 63.6 \text{ k}\Omega, R_2 = 12.0 \text{ k}\Omega, 6.63\%$
- 5.61 $I_{CQ} = 2.73 \text{ mA}, V_{CEQ} = 6 \text{ V},$
 $R_1 = 23.2 \text{ k}\Omega, R_2 = 2.83 \text{ k}\Omega$
- 5.63 (a) Let $R_E = 0.333 \text{ k}\Omega, R_C = 2 \text{ k}\Omega,$
 $R_1 = 22.3 \text{ k}\Omega, R_2 = 3.96 \text{ k}\Omega$
 (b) $2.90 \leq I_{CQ} \leq 3.10 \text{ mA}$
- 5.65 $R_E = 0.496 \text{ k}\Omega, R_1 = 32.3 \text{ k}\Omega,$
 $R_2 = 7.37 \text{ k}\Omega$
- 5.67 $R_E = 4.92 \text{ k}\Omega, R_1 = 72.5 \text{ k}\Omega,$
 $R_2 = 51.2 \text{ k}\Omega$
- 5.69 $R_1 = 152 \text{ k}\Omega, R_2 = 37.8 \text{ k}\Omega$
- 5.71 Let $R_E = 10 \text{ k}\Omega, R_C = 14 \text{ k}\Omega,$
 $R_1 = 70.0 \text{ k}\Omega, R_2 = 474 \text{ k}\Omega$
- 5.73 (a) $R_{TH} = 54.7 \text{ k}\Omega, V_{TH} = -3.03 \text{ V}$
 (b) $I_{CQ} = 0.227 \text{ mA}, V_{CEQ} = 7.51 \text{ V}$
- 5.75 (a) $R_E = 19.8 \text{ k}\Omega, R_C = 80 \text{ k}\Omega,$
 $R_1 = 573 \text{ k}\Omega, R_2 = 123 \text{ k}\Omega$
 (b) $I_{CQ} = 0.0514 \text{ mA}, V_{CEQ} = 4.86 \text{ V}$
- 5.77 (a) $R_E = 7 \text{ k}\Omega, R_C = 23 \text{ k}\Omega,$
 $R_1 = 103 \text{ k}\Omega, R_2 = 316 \text{ k}\Omega$
 (b) $I_{CQ} = 0.1027 \text{ mA}, V_{ECQ} = 2.914 \text{ V}$
- 5.79 $I_{B1} = 0.0144 \text{ mA}, I_{C1} = 1.73 \text{ mA},$
 $I_{E1} = 1.75 \text{ mA}, I_{B2} = 0.0232 \text{ mA},$
 $I_{C2} = 2.785 \text{ mA}, I_{E2} = 2.808 \text{ mA},$
 $V_{CEQ1} = 2.99 \text{ V}, V_{CEQ2} = 5.96 \text{ V}$
- 5.81 $R_1 = 130 \text{ k}\Omega, R_2 = 60 \text{ k}\Omega, R_3 = 60 \text{ k}\Omega,$
 $R_E = 2.5 \text{ k}\Omega, R_C = 10.5 \text{ k}\Omega$
- 5.83 $R_{E1} = 2.93 \text{ k}\Omega, R_{E2} = 4.25 \text{ k}\Omega,$
 $R_{C1} = 5.21 \text{ k}\Omega, R_{C2} = 3.21 \text{ k}\Omega$

CHAPTER 6

- 6.1 (a) (i) $g_m = 19.23 \text{ mA/V}, r_\pi = 9.36 \text{ k}\Omega,$
 $r_o = 300 \text{ k}\Omega$
 (ii) $g_m = 76.92 \text{ mA/V}, r_\pi = 2.34 \text{ k}\Omega,$
 $r_o = 75 \text{ k}\Omega$
 (b) (i) $g_m = 9.615 \text{ mA/V}, r_\pi = 8.32 \text{ k}\Omega,$
 $r_o = 400 \text{ k}\Omega$

- (ii) $g_m = 3.077 \text{ mA/V}$, $r_\pi = 26 \text{ k}\Omega$,
 $r_o = 1250 \text{ k}\Omega$
- 6.3 $30.77 \leq g_m \leq 46.15 \text{ mA/V}$,
 $1.95 \leq r_\pi \leq 5.85 \text{ k}\Omega$
- 6.5 (a) $g_m = 16.78 \text{ mA/V}$, $r_\pi = 7.15 \text{ k}\Omega$,
 $r_o = 183 \text{ k}\Omega$
(b) $A_v = -4.0$
(c) $v_s(t) = -0.125 \sin(100t) \text{ (V)}$
- 6.7 (a) $R_i = 56.24 \text{ k}\Omega$
(b) $R_i = 5.87 \text{ k}\Omega$
(c) $R_i = 6.24 \text{ k}\Omega$
- 6.9 $i_B(t) = 15 + 2.89 \sin \omega t (\mu\text{A})$
 $i_C(t) = 1.5 + 0.289 \sin \omega t (\text{mA})$
 $v_C(t) = 4 - 1.156 \sin \omega t (\text{V})$
 $A_v = -231$
- 6.11 $v_{be}(t) = -12 \sin \omega t (\text{mV})$
 $i_b(t) = -6 \sin \omega t (\mu\text{A})$
- 6.13 (a) $I_{CQ} = 0.506 \text{ mA}$, $V_{ECQ} = 1.78 \text{ V}$
(b) $A_v = -1.884$
(c) $1.871 \leq |A_v| \leq 1.895$
- 6.15 $R_E = 16.4 \text{ k}\Omega$, $R_C = 11.3 \text{ k}\Omega$,
 $A_v = -55.1$, $R_i = 14.5 \text{ k}\Omega$
- 6.17 (a) (i) $R_B = 20.25 \text{ k}\Omega$, $R_C = 2.53 \text{ k}\Omega$
(ii) $G_f = -11.63 \text{ mA/V}$
(b) (i) $R_B = 30.3 \text{ k}\Omega$, $R_C = 2.52 \text{ k}\Omega$
(ii) $G_f = -11.62 \text{ mA/V}$
- 6.19 (a) (i) $A_v = -26.97$, $G_f = -5.39 \text{ mA/V}$
(ii) $v_o(t) = -0.108 \sin \omega t (\text{V})$
 $i_o(t) = -21.6 \sin \omega t (\mu\text{A})$
(b) (i) $A_v = -30.5$, $G_f = -6.1 \text{ mA/V}$
(ii) $v_o(t) = -0.122 \sin \omega t (\text{V})$
 $i_o(t) = -24.4 \sin \omega t (\mu\text{A})$
- 6.21 (a) $V_B = -0.0347 \text{ V}$, $V_E = -0.735 \text{ V}$
(b) $R_C = 6.43 \text{ k}\Omega$
(c) $A_v = -81.7$
(d) $A_v = -74.9$
- 6.23 (a) $R_E = 11.0 \text{ k}\Omega$
(b) $R_C = 3.71 \text{ k}\Omega$
(c) $A_v = -43.9$
(d) $R_i = 4.81 \text{ k}\Omega$
- 6.27 $r_e = r_\pi \left\| \left(\frac{1}{g_m} \right) \right\| r_0$
- 6.29 For Figure 6.28,
Let $R_E = 0.5 \text{ k}\Omega$, $R_C = 12.7 \text{ k}\Omega$,
 $R_1 = 555 \text{ k}\Omega$, $R_2 = 55 \text{ k}\Omega$
- 6.31 Let $R_E = 2 \text{ k}\Omega$, $R_C = 15.1 \text{ k}\Omega$,
 $R_1 = 164 \text{ k}\Omega$, $R_2 = 28.4 \text{ k}\Omega$
- 6.33 3.24 V, peak-to-peak
- 6.35 $\Delta i_o = 0.546 \text{ mA}$, peak-to-peak
- 6.37 $R_1 = 34.9 \text{ k}\Omega$, $R_2 = 18.6 \text{ k}\Omega$
- 6.39 (a) 2.59 V, peak-to-peak
(b) $\Delta i_C = 1.294 \text{ mA}$, peak-to-peak
- 6.41 (a) $R_E = 0.586 \text{ k}\Omega$, $I_{CQ} = 0.832 \text{ mA}$
(b) $R_o = 29.3 \Omega$
- 6.43 (a) $I_Q = 2.37 \text{ mA}$
(b) $A_v = 0.9993$
- 6.45 (a) $I_{CQ} = 2.09 \text{ mA}$, $V_{CEQ} = 3.69 \text{ V}$
(c) $A_v = 0.484$
(d) $R_{ib} = 122 \text{ k}\Omega$, $R_o = 32.4 \Omega$
- 6.47 (a) $V_B = 0.0617 \text{ V}$, $V_E = 0.7617 \text{ V}$
(b) $g_m = 19 \text{ mA/V}$, $r_\pi = 4.21 \text{ k}\Omega$,
 $r_o = 304 \text{ k}\Omega$
(c) $A_v = 0.906$, $A_i = 14.8$
(d) $A_v = 0.728$, $A_i = 14.8$
- 6.49 (a) $I_{CQ} = 1.29 \text{ mA}$, $V_{ECQ} = 5.7 \text{ V}$
(b) $A_v = 0.974$, $R_{ib} = 85.7 \text{ k}\Omega$,
 $R_o = 19.8 \Omega$
(c) $i_s(t) = 32.7 \sin \omega t (\mu\text{A})$
 $i_o(t) = 2.73 \sin \omega t (\text{mA})$
 $v_o(t) = 2.73 \sin \omega t (\text{V})$
 $v_{eb}(t) = -72.5 \sin \omega t (\text{mV})$
- 6.51 $0.789 \leq A_v \leq 0.811$
 $17.0 \leq A_i \leq 19.6$
- 6.53 $0.9066 \leq A_v \leq 0.9999$
- 6.55 $R_1 = 318 \text{ k}\Omega$, $R_2 = 886 \text{ k}\Omega$
- 6.57 Let $R_E = 33.3 \text{ k}\Omega$, $R_1 = 594 \text{ k}\Omega$,
 $R_2 = 981 \text{ k}\Omega$, $v_o(\min) = 3.73 \sin \omega t (\text{V})$,
 $v_o(\max) = 3.86 \sin \omega t (\text{V})$
- 6.59 (a) $A_v = 76.9$
(b) $A_i = 0.9917$
(c) $R_i = 25.8 \Omega$
(d) $R_o = R_C = 2 \text{ k}\Omega$
- 6.61 (a) $A_v = 3.77$
(b) $A_i = 0.991$
(c) $R_i = 1.052 \text{ k}\Omega$
(d) $R_o = R_C = 4 \text{ k}\Omega$
- 6.63 (a) $I_{CQ} = 0.921 \text{ mA}$, $V_{ECQ} = 6.10 \text{ V}$
(b) $A_v = 161$
- 6.65 (a) $V_{CEQ} = 3.72 \text{ V}$
(b) $A_v = 63.6$
- 6.67 (a) $I_{CQ} = 0.4663 \text{ mA}$, $V_{CEQ} = 4.79 \text{ V}$
(b) $R_m = 2.12 \text{ V/mA}$
(c) $A_v = 6.73$
- 6.69 (a) $I_{CQ} = 0.984 \text{ mA}$, $V_{CEQ} = 2.34 \text{ V}$
(b) $A_v = 25.4$

- 6.71 (a) $I_{CQ} = 1.91 \text{ mA}$, $V_{ECQ} = 13.3 \text{ V}$
 (b) $2.70 \leq A_v \leq 2.71$
 (c) $0.0302 \leq R_i \leq 0.0305 \text{ k}\Omega$
 $R_o = R_C = 6.5 \text{ k}\Omega$
- 6.73 (a) $A_{v1} = 0.4927$
 (b) $A_{v2} = 153.8$
 (c) $A_v = 75.8$
- 6.75 (a) $g_{m1} = 42.74 \text{ mA/V}$, $r_{\pi 1} = 2.34 \text{ k}\Omega$,
 $g_{m2} = 48.65 \text{ mA/V}$, $r_{\pi 2} = 2.06 \text{ k}\Omega$,
 $r_{o1} = r_{o2} = \infty$
 (b) $A_{v1} = -85.48$, $A_{v2} = -97.3$
 (c) $A_v = 3909$, $A'_v = A_{v1} \cdot A_{v2} = 8317$
- 6.77 (a) $I_{C1} = 2.253 \text{ mA}$, $I_{C2} = 69.73$
 (b) $A_v = 0.990$
 (c) $R_{ib} = 467.6 \text{ k}\Omega$, $R_o = 0.512 \Omega$
- 6.79 (a) $R_o = 30.77 \text{ k}\Omega$
 (b) $R_o = 31.38 \text{ k}\Omega$
 (c) $R_o = 19.17 \text{ k}\Omega$
- 6.81 (a) $\bar{P}_Q = 2.564 \text{ mW}$, $\bar{P}_{RC} = 3.123 \text{ mW}$,
 $\bar{P}_{RE} = 3.175 \text{ mW}$
 (b) $\bar{P}_{RC} = 0.721 \text{ mW}$
- 6.83 (a) $\bar{P}_{RL} = 0.286 \text{ mW}$,
 (b) $\bar{P}_{RL} = 0.499 \text{ mW}$

CHAPTER 7

- 7.1 (a) $T(s) = \frac{1}{1 + sR_1C_1}$
 (b) $f_H = 159 \text{ Hz}$
 (c) $v_O(t) = 1 - \exp\left(\frac{-t}{R_1C_1}\right)$
- 7.3 (a) $T(s) = \left(\frac{R_2}{R_1 + R_2}\right) \cdot \frac{1}{1 + s(R_1 \| R_2)C_1}$
 (b) $\tau = 66.7 \text{ ms}$
 (c) $f = 2.39 \text{ Hz}$
- 7.5 (a) $\frac{V_o}{V_i} = \frac{R_2}{R_1 + R_2} = 0.667$
 (b) $\frac{V_o}{V_i} = 1$
 (c) $T(s) = \left(\frac{R_2}{R_1 + R_2}\right) \cdot \frac{\frac{[1 + sR_1C_1]}{[1 + s(R_1 \| R_2)C_1]}}{\frac{R_2}{R_1 + R_2}}$
 $K = \frac{R_2}{R_1 + R_2}$, $\tau_A = R_1C_1$,
 $\tau_B = (R_1 \| R_2)C_1$

- 7.7 (a) $|T(f_T)|_{\text{dB}} = -9.03 \text{ dB}$, $\phi = -135^\circ$
 (b) Slope = $-18 \text{ dB/octave} = -60 \text{ dB/decade}$, $\phi = -270^\circ$
- 7.9 (a) (ii) $\omega_1 = 1 \text{ rad/s}$, $\omega_2 = 10 \text{ rad/s}$,
 $\omega_3 = 100 \text{ rad/s}$, $\omega_4 = 1000 \text{ rad/s}$
 (iii) $|T(0)| = 10$
 (iv) $|T(\infty)| = 10$
 (b) (ii) $\omega = 5 \text{ rad/s}$
 (iii) $|T(0)| = 0$
 (iv) $|T(\infty)| = 200$
- 7.11 (a) $|T|_{\text{midband}} = 159$
 (b) $\tau_S = 5.31 \text{ ms}$, $\tau_P = 0.332 \mu\text{s}$
 (c) $C_C = 0.932 \mu\text{F}$, $C_L = 55.3 \text{ pF}$
- 7.15 (a) $R_C = 2.49 \text{ k}\Omega$
 (b) $I_{CQ} = 0.503 \text{ mA}$
 (c) $|A_v|_{\text{max}} = 48.1$
- 7.17 (a) $f_L = 959 \text{ Hz}$
 (b) $|A_v| = 6.69$
- 7.19 (a) $R_S = 2.59 \text{ k}\Omega$, $R_D = 4.41 \text{ k}\Omega$
 (b) $T(s) = \frac{I_o(s)}{V_i(s)}$
 $= \frac{-g_m R_D}{(1 + g_m R_S)(R_D + R_L)}$
 $\times \frac{s(R_D + R_L)C_C}{1 + s(R_D + R_L)C_C}$
 $\tau = (R_D + R_L)C_C$
 (c) $C_C = 1.89 \mu\text{F}$
- 7.21 (a) $R_1 = 79.6 \text{ k}\Omega$, $R_2 = 124 \text{ k}\Omega$
 (b) $A_v = 0.991$
 (c) $R_o = 17.25 \Omega$
 (d) $f_L = 19.8 \text{ Hz}$
- 7.23 (a) $T(s) = \frac{g_m(R_D \| R_L)}{1 + g_m R_S} \cdot \frac{1}{1 + s\left(\frac{1}{g_m} \| R_S\right)C_i}$
 $\cdot \left[\frac{s(R_D + R_L)C_C}{1 + s(R_D + R_L)C_C}\right]$
 (b) $\tau = \left(\frac{1}{g_m} \| R_S\right)C_i$
 (c) $\tau = (R_L + R_D)C_C$
- 7.25 (a) $f_H = 8.30 \text{ MHz}$
 (b) $f = 82.6 \text{ MHz}$

- 7.27 Let $I_{DQ} = 0.2 \text{ mA}$, $V_{DSQ} = 5 \text{ V}$,
 $V_{TN} = 0.5 \text{ V}$
 $R_D = 20 \text{ k}\Omega$, $\frac{W}{L} = 6.25$, $R_1 = 192.5 \text{ k}\Omega$,
 $R_2 = 32.5 \text{ k}\Omega$, $C_C = 0.0284 \mu\text{F}$,
 $C_L = 2.65 \text{ nF}$
- 7.29 (a) $R_E = 6 \text{ k}\Omega$, $R_B = 532 \text{ k}\Omega$
(b) $C_C = 0.0257 \mu\text{F}$
(c) $A_v = 0.983$
- 7.31 (a) $C_{C1} = 12 \mu\text{F}$, (b) $C_{C2} = 0.050 \text{ F}$
- 7.33 (a) $R_S = 6.4 \text{ k}\Omega$, $R_D = 5.6 \text{ k}\Omega$
(b) $f_A = 4.97 \text{ Hz}$, $f_B = 36.8 \text{ Hz}$
(c) $f_A \rightarrow 0$, $f_B = 31.8 \text{ Hz}$
- 7.35 (a) Same expression as Equation (7.59)
with $R_S = 0$.
(b) $\tau_A = R_E C_E$

$$\tau_B = \frac{r_\pi R_E C_E}{r_\pi + (1 + \beta) R_E}$$
- 7.37 (a) $A_v = 118$
(b) $f = 15.9 \text{ MHz}$
- 7.39 $C_L = 121 \text{ pF}$
- 7.41 (a) $A_v(s) = -g_m(R_D \| R_L)$
 $\cdot \left[\frac{1}{1 + s(R_D \| R_L) C_L} \right]$
(b) $\tau = (R_D \| R_L) C_L$
(c) $\tau = 6.67 \times 10^{-8} \text{ s}$, $f_H = 2.39 \text{ MHz}$,
 $A_v = -4.7$
- 7.45 $g_m = 9.615 \text{ mA/V}$, $f_\beta = 33.3 \text{ MHz}$,
 $C_\pi = 0.303 \text{ pF}$
- 7.47 (a) $f_\beta = 4.5 \text{ MHz}$, $C_\pi = 1.87 \text{ pF}$
(b) $f_T = 2.16 \text{ GHz}$, $f_\beta = 18.0 \text{ MHz}$
- 7.49 (a) $Z_i = 2700 - j3.33$
(b) $Z_i = 2700 - j33.3$
(c) $Z_i = 2656 - j327$
(d) $Z_i = 1100 - j1200$
- 7.51 $f = 307 \text{ kHz}$
- 7.53 $f_L = 159 \text{ Hz}$, $f_H = 489 \text{ kHz}$
- 7.55 $f_T = 1.84 \text{ GHz}$
- 7.57 (a) $f_T = 1.25 \text{ GHz}$
(b) $f_T = 3.05 \text{ GHz}$
(c) $I_D = 290 \mu\text{A}$
(d) $C_{gs} = 70 \text{ fF}$
- 7.59 (a) $C_M = 344.3 \text{ fF}$
(b) $f_{3-\text{dB}} = 37.5 \text{ MHz}$
- 7.61 (a) $r_s = 43.5 \Omega$
(b) $\frac{g'_m}{g_m} = 94.5\%$
- 7.63 (b) $C_M = 96.28 \text{ fF}$
(c) $f_{3-\text{dB}} = 1.095 \text{ GHz}$, $A_v = -10.96$
- 7.65 (a) $C_\pi = 2.20 \text{ pF}$, $C_M = 27.6 \text{ pF}$
(b) $f_H = 3.61 \text{ MHz}$, $A_v = -19.7$
- 7.67 (a) $C_M = 6.785 \text{ pF}$
(b) $f_H = 4.84 \text{ MHz}$, $A_v = -5.67$
- 7.69 (a) $f_H = -10.4 \text{ MHz}$
(b) $C_M = 18.2 \text{ pF}$
(c) $A_v = -4.66$
- 7.71 (a) $f_{P\pi} = 1.20 \text{ GHz}$, $f_{P\mu} = 56.3 \text{ MHz}$
(b) $A_v = 2.76$
(c) $f = 3.75 \text{ MHz}$, C_L dominates
- 7.73 $f = 119 \text{ MHz}$, $A_v = 1.19$

CHAPTER 8

- 8.1 (b) (i) $R_D = 6 \Omega$, $P_{D,\text{max}} = 24 \text{ W}$
(ii) $R_D = 13.3 \Omega$, $P_{D,\text{max}} = 30 \text{ W}$
(c) (i) $I_{D,\text{max}} = 4 \text{ A}$
(ii) $I_{D,\text{max}} = 3 \text{ A}$
- 8.3 (a) $R_L = 25 \Omega$, $R_B = 3.91 \text{ k}\Omega$,
 $P_{Q,\text{max}} = 9 \text{ W}$
(b) $I_{C,\text{max}} = 0.8944 \text{ A}$, $V_{CC} = 22.36 \text{ V}$
(c) For (a): $\bar{P}_L = 4.5 \text{ W}$
For (b): $\bar{P}_L = 2.5 \text{ W}$
- 8.5 (b) $V_{GG} = 5 \text{ V}$, $P = 9.375 \text{ W}$;
 $V_{GG} = 6 \text{ V}$, $P = 30 \text{ W}$;
 $V_{GG} = 7 \text{ V}$, $P = 39.375 \text{ W}$;
 $V_{GG} = 8 \text{ V}$, $P = 10.8 \text{ W}$;
 $V_{GG} = 9 \text{ V}$, $P = 7.16 \text{ W}$
- 8.7 (b) $T_{j,\text{max}} = 145^\circ\text{C}$
(c) $\theta_{\text{dev-amb}} = 2^\circ\text{C/W}$
- 8.9 (a) $P_T = 19.39 \text{ W}$
(b) $T_{\text{case}} = 90.9^\circ\text{C}$, $T_{\text{snk}} = 79.3^\circ\text{C}$
- 8.11 (a) $P_T = 13.4 \text{ W}$
(b) $V_{CE,\text{max}} = 8.96 \text{ V}$
- 8.13 (b) (i) $I_C = 29.25 \text{ mA}$
(ii) $I_C = 61.75 \text{ mA}$
(iii) $I_C = 108 \text{ mA}$
- 8.15 (a) $A_v = 0.9928$
(b) $A_v = 0.9872$
(c) $A_v = 0.939$
- 8.17 $v_O(\text{min}) = -3.93 \text{ V}$, $v_I(\text{min}) = -3.43 \text{ V}$;
 $v_O(\text{max}) = 4.43 \text{ V}$, $v_I(\text{max}) = 5.76 \text{ V}$
- 8.19 (a) $I_Q = 0.12 \text{ A}$, $R = 187 \Omega$
(b) $P_{Q1} = 2.88 \text{ W}$, $P = 5.88 \text{ W}$
(c) $\eta = 11.4\%$

- 8.23 (a) $V^+ = -V^- = 52 \text{ V}$
 (b) $I_P = 2.04 \text{ A}$
 (c) $\eta = 74\%$
- 8.25 (b) (i) $A_v = 0$
 (ii) $A_v = 0.472$
 (iii) $A_v = 0.739$
- 8.27 (a) (i) $V_{BB} = 4 \text{ V}$
 (ii) $P = 12 \text{ mW}$
 (b) (i) $v_O(\max) = 10.39 \text{ V}$
 $i_{Dn} = 10.39 \text{ mA}, i_{Dp} = 0,$
 $i_L = 10.39 \text{ mA}, v_I = 11.5 \text{ V}$
 (iii) $P_{RL} = 108 \text{ mW}, P_{Mn} = 16.7 \text{ mW},$
 $P_{Mp} = 0$
- 8.29 (a) $K_{n3} = K_{p4} = 200 \mu\text{A/V}^2$
 (b) (i) $A_v = 0$; (ii) $A_v = 0.876$
- 8.31 (a) $I_{CQ} = 50.57 \text{ mA}$
 (b) $R_L = 237 \Omega$
 (c) $\bar{P}_L,\max = 255 \text{ mW}$
 (d) $\eta = 42\%$
- 8.33 (a) $I_{CQ} = 98.91 \text{ mA}$
 (b) $a = 3.89$
 (c) $\bar{P}_L = 333.9 \text{ mW}$
 (d) $\eta = 28.1\%$
- 8.35 (a) $\frac{v_e}{v_i} = \frac{(1 + \beta)R'_E}{r_\pi + (1 + \beta)R'_E}$
 $\frac{v_o}{v_i} = \frac{1}{(n_1/n_2)} \cdot \frac{(1 + \beta) R'_E}{r_\pi + (1 + \beta) R'_E}$
 where $R'_E = \left(\frac{n_1}{n_2}\right)^2 R_L$
 (b) $a^2 = \frac{V_{CC}}{5}$
 (c) $R_o = 0.255 \Omega$
- 8.37 (a) $a = 2.86$
 (b) $P_Q = 4.95 \text{ W}$
- 8.39 (a) $V_{BB} = 1.473 \text{ V}$
 (b) $I_{CQ} = 14 \text{ mA}$
- 8.41 (a) $V_{BB} = 1.4845 \text{ V}$
 (b) $v_{BEn} = 0.75124 \text{ V}, v_{EBp} = 0.73322 \text{ V}$
 (c) $i_{Cn} = i_{Cp} = 2.828 \text{ mA}$
 (d) $v_I = -0.73322 \text{ V}$
- 8.43 (a) For $v_O(\max) = 4 \text{ V}$,
 $R_1 = R_2 = 32.5 \Omega$
 (b) $I_{E1} = I_{E2} = 0.286 \text{ A}$,
 $I_{E3} = I_{E4} = 2.86 \text{ A}$
 (c) $R_o = 5.34 \text{ m}\Omega$
- 8.45 (b) $R_1 = R_2 = 4 \text{ k}\Omega$
 (c) $I_{DQ1} = I_{DQ2} = 2 \text{ mA}$
- (d) $I_{Dn3} = 23.33 \text{ mA}, I_{Dp1} = 0.835 \text{ mA},$
 $I_{Dn2} = 3.188 \text{ mA}, I_{Dp4} = 0,$
 $\bar{P}_L = 81.7 \text{ mW}$
- 8.47 $R_o = 7.46 \Omega$
- 8.49 (a) $V_{BB} = 1.742 \text{ V}, I_{C1} \cong 10 \text{ mA},$
 $I_{C2} = 0.20 \text{ A}, I_{C3} \cong 0.5 \text{ mA},$
 $V_{BE1} = 0.58065 \text{ V}, V_{BE2} = 0.6585 \text{ V}$
 $V_{EB3} = 0.50276 \text{ V}$
 (b) $i_{C1} = 4.47 \text{ mA}, i_{C2} = 93.9 \text{ mA},$
 $i_{C3} = 2.267 \text{ mA}, i_{C4} = 45.3 \text{ mA},$
 $i_{C5} = 0.952 \text{ A}, P_{Q2} = 4.13 \text{ W},$
 $P_{Q5} = 13.3 \text{ W}$

CHAPTER 9

- 9.1 (a) $A_{od} = 500$, (b) $v_2 = 3 \text{ mV}$,
 (c) $v_1 = 0.990 \text{ V}$, (d) $v_o = 0$,
 (e) $v_2 = -0.506 \text{ V}$
- 9.3 (a) $v_O = -5 \text{ V}$
 (b) $v_1 = 3.00265 \text{ V}$
 (c) $A_{od} = 9 \times 10^4$
- 9.5 (a) $A_v = -10$
 (b) $A_v = -3$
 (c) $A_v = -1$
- 9.7 (a) $A_v = -10, R_i = 10 \text{ k}\Omega$
 (b) $A_v = -5, R_i = 10 \text{ k}\Omega$
 (c) $A_v = -5, R_i = 20 \text{ k}\Omega$
- 9.9 (a) -10 , (b) -1 , (c) -0.20 ,
 (d) -10 , (e) -2 , (f) -1
- 9.11 (a) $R_1 = 5 \text{ k}\Omega, R_2 = 32.5 \text{ k}\Omega$
 (b) $v_I = 0.6154 \text{ V}, i_1 = i_2 = 0.123 \text{ mA}$
- 9.13 (a) $-9.5\% \text{ to } +10.5\%$
 (b) $\pm 2\%$
- 9.15 $R_1 = 18.5 \text{ k}\Omega, R_2 = 599.6 \text{ k}\Omega$
 $0.731 \leq |v_O| \leq 0.769 \text{ V}$
- 9.17 $R_1 = 10 \text{ k}\Omega, R_2 = 100 \text{ k}\Omega,$
 $R_3 = 10 \text{ k}\Omega, R_4 = 100 \text{ k}\Omega$
- 9.19 (a) $v_O = 8.8 \text{ V}$
 (b) $v_O = 8.7597 \text{ V}$
 (c) $A_{od} = 1.1477 \times 10^4$
- 9.21 (a) $v_I = -0.7011 \text{ V}, v_1 = -1 \text{ mV}$
 (b) $A_{od} = 2.5 \times 10^4, v_I = 0.50022 \text{ V}$
- 9.23 (a) $R_1 = 500 \text{ k}\Omega, R_2 = R_3 = 500 \text{ k}\Omega,$
 $R_4 = 6.41 \text{ k}\Omega$
 (b) $i_1 = i_2 = -0.1 \mu\text{A}, i_4 = -7.80 \mu\text{A},$
 $i_3 = -7.90 \mu\text{A}$
- 9.25 For Figure P9.25, $A_v = -8$;
 For Figure 9.12, $A_v = -3$

- 9.27 (a) (i) $v_O = 0.79936 \text{ V}$
 (ii) 0.08%
 (b) (i) $v_O = 0.79208 \text{ V}$
 (ii) 0.99%
- 9.29 $R_2 = 111 \text{ k}\Omega$
- 9.31 (a) $R_1 = 83.3 \text{ k}\Omega, R_2 = 40 \text{ k}\Omega,$
 $R_3 = 400 \text{ k}\Omega, R_F = 250 \text{ k}\Omega$
 (b) $|i_F| = 0.75 \mu\text{A}$
- 9.33 $v_O = [-0.707 \sin(2\pi ft) \pm 2] (\text{V})$
- 9.35 (a) $R_1 = 250 \text{ k}\Omega, R_2 = 31.25 \text{ k}\Omega,$
 $R_3 = 62.5 \text{ k}\Omega, R_F = 125 \text{ k}\Omega$
 (b) $-3 \leq v_O \leq +3 \text{ V}, |i_F|_{\max} = 24 \mu\text{A}$
- 9.37 (b) $R_F = 10 \text{ k}\Omega$
 (c) (i) $v_O = 0.3125 \text{ V},$ (ii) $v_O = 4.6875 \text{ V}$
- 9.41 (a) $R_1 = 4.17 \text{ k}\Omega, R_2 = 58.33 \text{ k}\Omega$
 (b) $v_O = 3.75 \text{ V}, i_1 = i_2 = 60 \mu\text{A}$
- 9.43 (a) $A_{od} = 250.5$
 (b) $v_O = 49.9 \text{ V}$
- 9.45 (a) $v_O = \frac{6}{7} \cdot v_{I1} + \frac{3}{7} \cdot v_{I2}$
 (b) $v_O = 0.3 \text{ V}$
 (c) $v_O = 42.86 \text{ mV}$
- 9.47 (a) $A_v = \frac{1}{1-x}$
 (b) $1 \leq A_v \leq \infty$
 (c) $x \neq 1$
- 9.49 (a) $A_v = 1$
 (b) $A_v = 0.999993$
 (c) $A_{od} = 99$
- 9.51 (a) $A_{v1} = \left(1 + \frac{R_2}{R_1}\right), A_{v2} = -\left(1 + \frac{R_2}{R_1}\right)$
 (b) $v_{O1} = -2 \text{ V}, v_{O2} = 2 \text{ V}$
 (c) $(v_{O1} - v_{O2}) = 6.4 \text{ V}$
- 9.53 (a) (i) $v_O = 1 \text{ V}$
 (ii) $v_O = 1.67 \text{ V}$
 (b) (i) $v_O = 1 \text{ V}$
 (ii) $v_O = 1.67 \text{ V}$
 (c) (i) $v_O = 0.667 \text{ V}$
 (ii) $v_O = -1.111 \text{ V}$
- 9.55 From Figure P9.54, let $R_F = 1 \text{ k}\Omega$
- 9.57 (a) $R_2 = 1 \text{ k}\Omega, \text{Let } R_3 = 1 \text{ k}\Omega,$
 $R_1 = R_F = 10 \text{ k}\Omega$
 (b) $i_1 = i_2 = -0.6 \text{ mA}, v_O = 7 \text{ V},$
 $i_3 = 6 \text{ mA}, i_4 = 1 \text{ mA}, i_O = 6.6 \text{ mA}$
- 9.59 (b) $R_o = \infty$
- 9.61 (a) $R_2 = R_4 = 250 \text{ k}\Omega,$
 $R_1 = R_3 = 6.25 \text{ k}\Omega$
- (b) $R_2 = R_4 = 250 \text{ k}\Omega,$
 $R_1 = R_3 = 10 \text{ k}\Omega$
 (c) $R_2 = R_4 = 250 \text{ k}\Omega,$
 $R_1 = R_3 = 50 \text{ k}\Omega$
 (d) $R_2 = R_4 = 125 \text{ k}\Omega,$
 $R_1 = R_3 = 250 \text{ k}\Omega$
- 9.63 (a) $v_O = -4 \text{ V}, v_X = v_Y = 1.273 \text{ V},$
 $i_1 = i_2 = 0.0527 \text{ mA},$
 $i_3 = i_4 = 0.0127 \text{ mA}$
 (b) $v_O = 4 \text{ V}, v_X = v_Y = 3.273 \text{ V},$
 $i_1 = i_2 = -0.00727 \text{ mA},$
 $i_3 = i_4 = 0.0327 \text{ mA}$
 (c) $v_O = -1.5 \text{ V}, v_X = v_Y = -1.227 \text{ V},$
 $i_1 = i_2 = 0.00273 \text{ mA},$
 $i_3 = i_4 = -0.0123 \text{ mA}$
- 9.65 (b) $A_{cm} = 0.3$
 (c) $|A_{cm}|_{\max} = 0.0325$
- 9.67 (a) $v_{O1} = 1.2 - 0.72 \sin \omega t (\text{V})$
 $v_{O2} = 1.2 + 0.72 \sin \omega t (\text{V})$
 $v_O = 4.32 \sin \omega t (\text{V})$
 (b) $v_{O1} = -0.85 + 0.45 \sin \omega t (\text{V})$
 $v_{O2} = -0.40 - 0.45 \sin \omega t (\text{V})$
 $v_O = 1.35 - 2.7 \sin \omega t (\text{V})$
- 9.69 (a) $i_O = \frac{v_{I1} - v_{I2}}{R}$
 (b) $R = 100 \Omega$
 (c) $v_{O1} = 5.25 \text{ V}, v_{O2} = -0.25 \text{ V}$
 (d) $i_O = -1 \text{ mA}, v_{O1} = -1.75 \text{ V},$
 $v_{O2} = 1.75 \text{ V}$
- 9.71 $R_1(\text{fixed}) = 1 \text{ k}\Omega, R_2 = 39.5 \text{ k}\Omega,$
 $R_1(\text{var}) = 78 \text{ k}\Omega$
- 9.73 (a) $f = 398 \text{ Hz, phase} = 90^\circ$
 (b) (i) $f = 66.3 \text{ Hz},$ (ii) $f = 663 \text{ Hz}$
- 9.75 (b) $A_v(0) = -\frac{R_2}{R_1}$
 (c) $f = \frac{1}{2\pi R_2 C_2}$
- 9.77 (b) $A_v(\infty) = -\frac{R_2}{R_1}$
 (c) $f = \frac{1}{2\pi R_1 C_1}$
- 9.79 $i_2 = 1.52 \text{ mA}, v_O = -1.52 \text{ V},$
 $i_z = 0$
- 9.83 From Figure 9.40, $R_F = 500 \text{ k}\Omega,$
 $R_1 = 50 \text{ k}\Omega, R_2 = 500 \text{ k}\Omega,$
 $R_A = 500 \text{ k}\Omega, R_B = 333.3 \text{ k}\Omega,$
 $R_C = 142.8 \text{ k}\Omega$

9.85 $R_F = 5.33 \text{ k}\Omega$, $R_4 = 31.5 \text{ k}\Omega$,
 $R_3 = 10.6 \text{ k}\Omega$, $\frac{R_2}{R_1} = 1.143$

9.87 For the bridge circuit, $R_T = 20(1 + \delta) \text{ k}\Omega$,
 $R_1 = R_2 = R_3 = 20 \text{ k}\Omega$
For the instrumentation amplifier,

Let $\frac{R_4}{R_3} = 4$, $\frac{R_2}{R_1} = 4.5$

CHAPTER 10

10.1 (a)

$$I_C = \frac{1}{R_3} \left[2V_\gamma - (2V_\gamma + V^-) \left(\frac{R_2}{R_1 + R_2} \right) - V_{BE} \right]$$

(c) $R_3 = 2.5 \text{ k}\Omega$,
 $R_1 = R_2 = 2.15 \text{ k}\Omega$

10.3 $V_{BE1} = V_{BE2} = 0.6341 \text{ V}$, $I_O = 78.05 \mu\text{A}$

10.5 $V_{BE1} = 0.63345 \text{ V}$, $V_{BE2} = 0.62393 \text{ V}$,
 $I_O = 132.07 \mu\text{A}$

10.7 $I_{REF} = 0.5082 \text{ mA}$,
 $I_{C1} = I_{C2} = 0.4958 \text{ mA}$,
 $I_{B1} = I_{B2} = 6.198 \mu\text{A}$

10.9 (a) $R_1 = 18.3 \text{ k}\Omega$
(b) 6.3%

10.11 $I_{REF} = 0.210 \text{ mA}$, $R_1 = 20.5 \text{ k}\Omega$

10.13 (a) $R_1 = 9.3 \text{ k}\Omega$
(b) $I_O = 2 \text{ mA}$
(c) $R_{C2} = 2.15 \text{ k}\Omega$

10.15 Similar to Figure P10.14, biased at V^+ and V^- . $R_1 = 21.5 \text{ k}\Omega$

$$10.17 \text{ (a)} I_O = \frac{I_{REF} - \frac{V_{BE}}{(1 + \beta) R_2}}{\left(1 + \frac{2}{\beta(1 + \beta)} \right)}$$

(b) $R_1 = 12.27 \text{ k}\Omega$, $I_{REF} = 0.7011 \text{ mA}$

10.19 $R_1 = 30.63 \text{ k}\Omega$

10.21 $I_O = \frac{I_{REF}}{\left(1 + \frac{2}{\beta(2 + \beta)} \right)}$

10.23 (a) $R_o = 20 \text{ M}\Omega$
(b) $\Delta I_O = 0.25 \mu\text{A}$

10.25 (a) $V_{BE1} = 0.6347 \text{ V}$, $V_{BE2} = 0.6040 \text{ V}$,
 $I_O \cong 61.4 \mu\text{A}$

(b) $V_{BE1} = 0.6347 \text{ V}$, $V_{BE2} = 0.5991 \text{ V}$,
 $I_O \cong 71.2 \mu\text{A}$

10.27 (a) $I_{REF} = 0.186 \text{ mA}$, $I_O \cong 19.53 \mu\text{A}$,
 $V_{BE2} = 0.6414 \text{ V}$
(b) $R_o = 13.16 \text{ M}\Omega$

10.29 (a) $R_1 = 18.6 \text{ k}\Omega$, $R_E = 1.20 \text{ k}\Omega$
(b) $R_o = 6.477 \text{ M}\Omega$
(c) 1.54%

10.31 $V_{BE1} = 0.718 \text{ V}$ at 2 mA; $R_1 = 7.14 \text{ k}\Omega$,
 $R_E = 1.92 \text{ k}\Omega$

10.33 $V_{BE1} = 0.681 \text{ V}$, $I_{REF} = 0.483 \text{ mA}$,
 $I_O \cong 8.7 \mu\text{A}$, $V_{BE2} = 0.5766 \text{ V}$

10.35 (a) $R_1 = 10.1 \text{ k}\Omega$, $R_{E2} = 1.37 \text{ k}\Omega$,
 $V_{BE1} = 0.70038 \text{ V}$, $V_{BE2} = 0.67656 \text{ V}$
(b) $R_1 = 10.1 \text{ k}\Omega$, $R_{E2} = 1.46 \text{ k}\Omega$,
 $V_{BE2} = 0.65854 \text{ V}$

10.37 (a) $I_{REF} = 0.93 \text{ mA}$, $I_{O2} \cong 68 \mu\text{A}$,
 $I_{O3} \cong 40.7 \mu\text{A}$

(b) $R_{E2} = 4.99 \text{ k}\Omega$, $R_{E3} = 0.797 \text{ k}\Omega$

10.39 (a) $I_{O1} = 3.1 \text{ mA}$, $I_{O2} = 1.55 \text{ mA}$,
 $I_{O3} = 4.65 \text{ mA}$
(b) $V_{CE1} = 3.8 \text{ V}$, $V_{EC2} = 5.35 \text{ V}$,
 $V_{EC3} = 5.35 \text{ V}$

10.41 $I_{C1} = I_{C2} = 1.86 \text{ mA}$,
 $I_{C3} = I_{C4} = 1.86 \text{ mA}$,
 $I_{C5} = I_{C6} = I_{C7} = 0.136 \text{ mA}$,
 $V_{CE3} = 7.02 \text{ V}$, $V_{CE5} = 14.2 \text{ V}$,
 $V_{EC7} = 4.89 \text{ V}$

10.43 $R_{E1} = 3 \text{ k}\Omega$, $R_{E2} = 1.5 \text{ k}\Omega$,
 $R_{E3} = 0.75 \text{ k}\Omega$, $I_{O1} = 1 \text{ mA}$,
 $I_{O2} = 2 \text{ mA}$, $I_{O3} = 4 \text{ mA}$

10.45 (a) For $V_{GS} = 0.75 \text{ V}$, $R = 25 \text{ k}\Omega$,

$$\left(\frac{W}{L} \right)_1 = 20, \left(\frac{W}{L} \right)_2 = 40$$

(b) $R_o = 667 \text{ k}\Omega$

(c) 1.5%

10.47 (a) $0.19 \leq I_O \leq 0.21 \text{ mA}$
(b) $0.1921 \leq I_O \leq 0.2081 \text{ mA}$

10.49 $R_o = \frac{2 + g_m r_o}{g_m(1 + g_m r_o)} \cong \frac{1}{g_m}$

10.51 (a) $I_{REF} = 606 \mu\text{A}$
(b) $I_O = 362.5 \mu\text{A}$
(c) $I_O = 370.7 \mu\text{A}$

10.53 $\left(\frac{W}{L} \right)_1 = 5.56, \left(\frac{W}{L} \right)_2 = 2.78,$

$$\left(\frac{W}{L} \right)_3 = 1.81$$

10.55 $\left(\frac{W}{L}\right)_1 = 8.33, \left(\frac{W}{L}\right)_2 = 2.67,$

$$\left(\frac{W}{L}\right)_3 = \left(\frac{W}{L}\right)_4 = 2.72$$

10.57 $\left(\frac{W}{L}\right)_1 = 59.9, \left(\frac{W}{L}\right)_2 = 21.8,$

$$\left(\frac{W}{L}\right)_3 = 2.14$$

10.59 $R_o = 2.58 \times 10^9 \Omega$

10.61 $\left(\frac{W}{L}\right)_1 = 40, \left(\frac{W}{L}\right)_2 = 4,$

$$\left(\frac{W}{L}\right)_3 = 2.16, \left(\frac{W}{L}\right)_4 = 3.61$$

10.63 (a) $I_O = 80 \mu A$

(b) $I_O = 80.05 \mu A$

10.65 (a) $R = 6.12 \text{ k}\Omega$

(b) $(V^- - V^-)_{\min} = 1.66 \text{ V}$

$$(c) \left(\frac{W}{L}\right)_5 = 2.5, \left(\frac{W}{L}\right)_6 = 7.5$$

10.67 $\left(\frac{W}{L}\right)_1 = 86.1, \left(\frac{W}{L}\right)_2 = 12.5,$

$$\left(\frac{W}{L}\right)_3 = 25, \left(\frac{W}{L}\right)_4 = 50$$

10.69 $I_{\text{REF}} = 59.09 \mu A, I_1 = 11.82 \mu A,$
 $I_2 = 73.87 \mu A, I_3 = 47.27 \mu A,$
 $I_4 = 236.4 \mu A$

10.71 $I_{D2} = 62.5 \mu A, I_O = 41.67 \mu A,$
 $V_{GS1} = V_{GS2} = 1.107 \text{ V},$
 $V_{SG3} = V_{SG4} = 0.9893 \text{ V}$

10.73 (a) $I_O = 2.5 \text{ mA}$

(b) $I_O = 3 \text{ mA}$

(c) $I_O = 3.5 \text{ mA}$

10.75 (a) $A_v = -1846$

(b) $V_I = 0.6943 \text{ V}$

(c) $V_B = 1.824 \text{ V}$

10.77 (a) $V_{EB} = 0.5568 \text{ V}$

(b) $R_I = 4.44 \text{ k}\Omega$

(c) $V_I = 0.5389 \text{ V}$

(d) $A_v = -1846$

10.79 (a) $\left(\frac{W}{L}\right)_0 = 3.96,$

$$\left(\frac{W}{L}\right)_1 = \left(\frac{W}{L}\right)_2 = 3.24,$$

$$\left(\frac{W}{L}\right)_3 = 3.24$$

(b) $A_v = -70.35$

10.81 (a) $A_v = -1978$

(b) $A_v = -454$

(c) $A_v = -256$

10.83 (a) To a good approximation,

$$R_o = r_o [1 + g_{m2}(r_{\pi2} \| R_E)]$$

(b) $A_v = -g_{m0}(r_o \| R_L \| R_o)$

10.85 (a) $g_{m0} = 0.80 \text{ mA/V}, r_o = 333 \text{ k}\Omega$

$$g_{m2} = 0.748 \text{ mA/V}, r_{o2} = 250 \text{ k}\Omega$$

(b) $A_v = -114.3$

(c) $R_L = 143 \text{ k}\Omega$

10.87 $A_v = -316$

10.89 $A_v = -366, 165$

CHAPTER 11

11.1 (a) $v_O = 0.375 \sin \omega t (\text{V})$

(b) $v_O = 0.45 \sin \omega t (\text{V})$

(c) $v_O = 2.75 \sin \omega t (\text{V})$

11.3 (a) $R_E = 11.5 \text{ k}\Omega, R_C = 18 \text{ k}\Omega$

(c) $-2.3 \leq v_{cm} \leq 0.673 \text{ V}$

11.5 (a) $R_1 = 73.25 \text{ k}\Omega, R_C = 28.5 \text{ k}\Omega$

(b) $A_d = 62, A_{cm} = -0.113,$

$$\text{CMRR}_{\text{dB}} = 54.8 \text{ dB}$$

(c) $R_{id} = 46 \text{ k}\Omega, R_{icm} = 12.6 \text{ M}\Omega$

11.7 $R_1 = 53 \text{ k}\Omega, R_C = 40 \text{ k}\Omega, A_d = 38.5$

11.9 (a) $R_E = 62 \text{ k}\Omega$

(b) $A_d = 71.0, A_{cm} = -0.398,$

$$\text{CMRR}_{\text{dB}} = 45.0 \text{ dB}$$

(c) $R_{id} = 70.4 \text{ k}\Omega, R_{icm} = 6.28 \text{ M}\Omega$

11.11 (a) $I_{C1} = I_{C2} = 5.90 \mu A,$

$$V_{EC1} = V_{EC2} = 1.475 \text{ V}$$

(b) (i) $A_d = 81.7, A_{cm} = 0$

(ii) $A_d = 40.8, A_{cm} = -0.0442$

11.13 (a) $R_E = 17.9 \text{ k}\Omega, R_C = 16.7 \text{ k}\Omega$

(b) $|A_d| = 76.9$

(c) $A_d = 76.9, A_{cm} = 0.0279,$

$$\text{CMRR}_{\text{dB}} = 68.8 \text{ dB}$$

11.15 (a) $v_E = +0.7 \text{ V}, v_{C1} = v_{C2} = -2.87 \text{ V}$

(b) $v_E = +0.7 \text{ V}, v_{C1} = -5 \text{ V},$

$$v_{C2} = -0.736 \text{ V}$$

(c) $v_E \cong +0.7 \text{ V}, v_{C1} = -2.255 \text{ V},$

$$v_{C2} = -3.485 \text{ V}$$

11.17 (a) (i) $v_{O1} - v_{O2} = 0,$

(ii) $v_{O1} - v_{O2} = 0.08 \text{ V}$

(b) (i) $v_{O1} - v_{O2} = 0.1995 \text{ V},$

(ii) $v_{O1} - v_{O2} = 0.2795 \text{ V}$

11.19 (a) $v_d(\text{max}) \cong 14 \text{ mV}$

(b) $v_d(\text{max}) \cong 21.2 \text{ mV}$

- 11.21 (a) $R_E = 37.2 \text{ k}\Omega$
 (b) (i) $A_d = 119$, (ii) $A_{cm} = -0.673$
- 11.23 $A_{v1} = \frac{-\frac{1}{2}g_m R_L}{\left(2 + \frac{R_L}{R_C}\right)}, A_{v2} = \frac{+\frac{1}{2}g_m R_L}{\left(2 + \frac{R_L}{R_C}\right)}$
 $A_v = \frac{g_m R_L}{\left(2 + \frac{R_L}{R_C}\right)}$
- 11.25 $A_d = g_m R_C, A_{cm} = 0$
- 11.27 $-0.22 \leq v_{O2} \leq 0.22 \text{ V}$
- 11.29 (a) $R_1 = 38.6 \text{ k}\Omega, R_2 = 236 \Omega$
 (b) $R_{icm} = 292 \text{ M}\Omega$
 (c) $A_{cm} = -0.0123$
- 11.31 (a) $I_1 = I_Q = 0.337 \text{ mA}, I_{D1} = 0.168 \text{ mA}, V_{DS1} = 8.79 \text{ V}, V_{DS4} = 7.18 \text{ V}$
 (c) $v_{CM}(\text{max}) = 7.97 \text{ V}, v_{CM}(\text{min}) = -6.02 \text{ V}$
- 11.33 (a) $I_Q = 160 \mu\text{A}, R_D = 37.5 \text{ k}\Omega$
 (c) $v_{CM} = 2.50 \text{ V}$
- 11.35 (a) $R_1 = 33.9 \text{ k}\Omega$
 (b) (i) $v_{O1} - v_{O2} = 0$,
 (ii) $v_{O1} - v_{O2} = 0.125 \text{ V}$
 (c) (i) $v_{O1} - v_{O2} = -0.156 \text{ V}$
 (ii) $v_{O1} - v_{O2} = -0.031 \text{ V}$
- 11.37 (a) $v_d(\text{max}) \cong 0.19 \text{ V}$
 (b) $v_d(\text{max}) \cong 0.285 \text{ V}$
- 11.39 (a) $I_D = 6 \mu\text{A}, V_{SD} = 1.69 \text{ V}$
 (b) (i) $A_d = 9.66, A_{cm} = 0$
 (ii) $A_d = 4.83, A_{cm} = -0.0448$
- 11.41 $A_d = -9.16, A_{cm} = -0.003216$,
 $\text{CMRR}_{\text{dB}} = 69.1 \text{ dB}$
- 11.43 (a) $v_S = 1.459 \text{ V}, v_{D1} = v_{D2} = -4.115 \text{ V}$
 (b) $v_S = 2.344 \text{ V}, v_{D1} = v_{D2} = -4.336 \text{ V}$
 (c) $v_S \cong 1.459 \text{ V}, v_{D1} = -3.909 \text{ V}, v_{D2} = -4.321 \text{ V}$
 (d) $v_S \cong 2.344 \text{ V}, v_{D1} = -4.158 \text{ V}, v_{D2} = -4.515 \text{ V}$
- 11.45 Let $I_Q = 0.50 \text{ mA}$; For $V_{D1} = V_{D2} = 3 \text{ V}$,
 then $R_D = 28 \text{ k}\Omega; \left(\frac{W}{L}\right)_1 = \left(\frac{W}{L}\right)_2 = 1274$
 Need $R_o = 200 \text{ k}\Omega$
- 11.47 $A_{d1} = \frac{-\frac{1}{2}g_m R_L}{\left(2 + \frac{R_L}{R_D}\right)}, A_{d2} = \frac{+\frac{1}{2}g_m R_L}{\left(2 + \frac{R_L}{R_D}\right)}$
 $A_d = \frac{g_m R_L}{\left(2 + \frac{R_L}{R_D}\right)}$
- 11.49 (a) $R_D = 40.8 \text{ k}\Omega$
 (b) $v_{cm}(\text{max}) = 1.32 \text{ V}$
- 11.51 (a) $R_D = 19 \text{ k}\Omega, R_o = 998 \text{ k}\Omega$
 (b) Use cascode current source similar to Figure 10.18 with $v_{DS2}(\text{sat}) = 0.3 \text{ V}$.
- 11.53 Let $I_{Q1} = I_{Q2} = 0.1 \text{ mA}$,
 then $R_1 = 100 \text{ k}\Omega, R_2 = 50 \text{ k}\Omega$
 $\left(\frac{W}{L}\right)_1 = \left(\frac{W}{L}\right)_2 = 6.67$,
 $\left(\frac{W}{L}\right)_3 = \left(\frac{W}{L}\right)_4 = 240$
- 11.55 (a) $R_D = 6 \text{ k}\Omega, I_Q = 1 \text{ mA}$
 (b) $g_f(\text{max}) = 0.25 \text{ mA/V}$
 (c) $A_d = 1.5$
- 11.57 $A_d = \frac{g_m R_L}{\left(2 + \frac{R_L}{R_D}\right)}$
- 11.59 (a) $R_1 = 36.8 \text{ k}\Omega$
 (b) $A_d = 2308$
 (c) $R_{id} = 74.9 \text{ k}\Omega, R_o = 480 \text{ k}\Omega$
 (d) $-3.6 \leq v_{cm} \leq 4.3 \text{ V}$
- 11.61 (a) $I_Q = 0.749 \text{ mA}$
 (b) $v_{cm}(\text{max}) = 3.8 \text{ V}$
- 11.63 (a) $A_v = 89.72$
 (b) $A_v = 89.32$
 (c) $A_v = 89.52$
- 11.65 (a) $A_d = 79.44$
 (b) $v_{CM}(\text{max}) = 1.25 \text{ V}$
- 11.67 $I_{Q1} = I_Q$
- 11.69 (a) $R_o = 1 \| 4.704 = 0.8247 \text{ M}\Omega$
 (b) $A_d = 3172$
- 11.73 $I_Q \cong 66 \mu\text{A}, \left(\frac{W}{L}\right)_n = 23.76$
- 11.75 (a) $R_o = 172 \text{ M}\Omega$
 (b) $A_d = 46096$
- 11.77 (a) $V^+ = -V^- = 4 \text{ V}$
 (b) $A_d = 50$
- 11.79 (a) $V^+ = -V^- = 3.4 \text{ V}$
 (b) $A_d = 1004$
- 11.81 (a) $A_d = 88.9$
- 11.83 (a) $g_{m1} = 0.506 \text{ mA/V}, g_{m2} = 27.25 \text{ mA/V}, r_{\pi 2} = 6.605 \text{ k}\Omega$
 (b) $R_o = 19.8 \Omega$
- 11.85 $A_v = -48.4$
- 11.87 $R_i \cong 1.05 \text{ M}\Omega, R_o = 0.472 \text{ k}\Omega, A_v = -438$

- 11.89 (a) $R_1 = 27.2 \text{ k}\Omega$, $R_2 = 5 \text{ k}\Omega$
 (b) $g_{m1} = 0.7071 \text{ mA/V}$, $g_{m2} = 2 \text{ mA/V}$,
 $r_{o1} = 200 \text{ k}\Omega$, $r_{o2} = 50 \text{ k}\Omega$
 (c) $A_v = -15.25$
 (d) $R_o = 0.450 \text{ k}\Omega$
- 11.91 (a) $R_{C1} = 80 \text{ k}\Omega$, $R_{C2} = 20 \text{ k}\Omega$
 (b) $A_{d1} = -69.6$, $A_d = -5352$
- 11.93 (a) $A_d = -3.73$, $A_{cm} = 0.0718$
 (b) $v_{O3} = -0.975 \sin \omega t (\text{V})$,
 $v_{O3}(\text{ideal}) = -1.12 \sin \omega t (\text{V})$
- 11.95 (a) $f_z = 39.8 \text{ kHz}$
 (b) $f_p = 11.7 \text{ GHz}$
- 11.97 (a) $A_v = 19.5$
 (b) $A_v = 11.2$

CHAPTER 12

- 12.1 (a) $\beta = 9.98 \times 10^{-3}$
 (b) $A = 2000$
- 12.3 (a) (i) $T = \infty$, $A_f = 6.667$
 (ii) $T = 1.5 \times 10^3$, $A_f = 6.662$
 (iii) $T = 15$, $A_f = 6.25$
 (b) (i) $T = \infty$, $A_f = 4.0$
 (ii) $T = 2.5 \times 10^3$, $A_f = 3.9984$
 (iii) $T = 25$, $A_f = 3.846$
- 12.5 (a) $\beta = -0.01245$
 (b) $\frac{dA}{A} = 2.5\%$
- 12.7 (a) $\beta_1 = 0.096685$, $\beta_2 = 0.0195$
 (b) For (a), $A_{vf} = 47.33$ (-5.34%),
 For (b), $A_{vf} = 49.86$ (-0.28%)
- 12.9 (a) $f_C = 20 \text{ kHz}$
 (b) $A_{vf} = 40$
- 12.11 (a) (i) $A_v = 5.25 \times 10^5$
 (ii) $\beta = 0.01333$
 (b) (i) $A_{vf} = 74.99$
 (ii) $f_C = 31.5 \text{ kHz}$
- 12.13 (a) For (a), $f_{3-\text{dB}} = 2.352 \text{ kHz}$,
 For (b), $f_{3-\text{dB}} = 5 \text{ kHz}$
 (b) Overall feedback \Rightarrow wider bandwidth
- 12.15 For (a), at low input, gain = 81; at high input, gain = 73.4. For (b), at low input, gain = 81.1, at high input, gain = 80.1
- 12.17 $V_{fb} = 24 \text{ mV}$, $V_\varepsilon = 1 \text{ mV}$,
 $A_v = 2.5 \times 10^3 \text{ V/V}$, $A_{vf} = 100 \text{ V/V}$
- 12.19 (a) $A_{vf} = \frac{A}{1 + \left[A \sqrt{\left(1 + \frac{R_2}{R_1} \right)} \right]}$

- (b) $\beta = \frac{1}{1 + \frac{R_2}{R_1}}$
 (c) $\beta = 0.04999$, $\frac{R_2}{R_1} = 19.004$
 (d) $-2.222 \times 10^{-3}\%$
- 12.21 $I_{fb} = 24.2 \mu\text{A}$, $I_o = 3.125 \text{ mA}$,
 $\beta_i = 0.007744 \text{ A/A}$, $A_i = 3906 \text{ A/A}$
- 12.23 $R_{if} = 0.5 \Omega$, $R_{of} \cong 2.42 \Omega$
- 12.25 $A_{gf} = 40 \text{ A/V}$, $I_o = 6 \text{ mA}$,
 $V_{fb} = 147 \mu\text{V}$, $V_\varepsilon = 3 \mu\text{V}$
- 12.27 $R_{if} = 10^6 \text{ k}\Omega$, $R_{of} = 5.04 \text{ M}\Omega$
- 12.29 $A_z = 4 \text{ V}/\mu\text{A}$, $A_{zf} = 0.2 \text{ V}/\mu\text{A}$,
 $\beta_g = 4.75 \mu\text{A/V}$
- 12.31 $R_{if} = 99.79 \Omega$
- 12.33 (a) $1 \Omega \leq R_i \leq 10^5 \text{ k}\Omega$
 (b) $0.1 \Omega \leq R_o \leq 10^4 \text{ k}\Omega$
- 12.35 (a) $A_{vf} = 11.0$
 (b) $R_{if} = 273 \text{ M}\Omega$
 (c) $R_{of} = 15 \text{ m}\Omega$
- 12.37 (a) $g_{m1} = 32.81 \text{ mA/V}$, $r_{\pi 1} = 3.66 \text{ k}\Omega$
 $g_{m2} = 19.12 \text{ mA/V}$, $r_{\pi 2} = 6.28 \text{ k}\Omega$
 $g_{m3} = 78.08 \text{ mA/V}$, $r_{\pi 3} = 1.54 \text{ k}\Omega$
 (b) $A_{vf} = 20.7$
 (c) $R_{if} = 62.4 \text{ k}\Omega$
 (d) $R_{of} = 1.39 \Omega$
- 12.39 $A_{vf} = 1.273$
- 12.41 (a) (i) $A_{vf} = 0.801$
 (ii) $R_{of} = 299 \Omega$
 (b) (i) $+3.78\%$
 (ii) -15.4%
- 12.45 For example, $R_2 = 3 \text{ k}\Omega$,
 $R_1 = 247 \text{ k}\Omega$
- 12.47 (a) $I_{DQ1} = 3.99 \text{ mA}$, $I_{DQ2} = 12.01 \text{ mA}$
 (b) $A_i = -0.920$
- 12.49 (a) $g_{m1} = 7.62 \text{ mA/V}$, $r_{\pi 1} = 13.1 \text{ k}\Omega$
 $g_{m2} = 52.7 \text{ mA/V}$, $r_{\pi 2} = 1.90 \text{ k}\Omega$
 (b) $A_{if} = 8.60$
 (c) $R_{if} = 47.4 \Omega$
- 12.51 (a) $g_{m1} = 67.31 \text{ mA/V}$, $r_{\pi 1} = 1.78 \text{ k}\Omega$
 $g_{m2} = 19.23 \text{ mA/V}$, $r_{\pi 2} = 6.24 \text{ k}\Omega$
 (b) $A_v = 7.44$
- 12.53 $A_{if} = 5.33$
- 12.55 (a) $A_{gf} = \frac{\frac{R_F}{R_1}}{\left(\frac{R_L R_F}{R_1} - R_3 - \frac{R_L R_3}{R_2} \right)}$

- (c) $R_2 = 2 \text{ k}\Omega$, set $R_3 = 2 \text{ k}\Omega$,
 $R_1 = R_F = 10 \text{ k}\Omega$
- 12.57 $A_{gf} = 98.06 \text{ mA/V}$
- 12.59 $A_{gf} = -0.0653 \text{ mA/V}$
- 12.61 (a) $I_{CQ} = 2.448 \text{ mA}$, $V_{ECQ} = 1.31 \text{ V}$
(b) $A_{zf} = -8.666 \text{ V/mA}$
(c) $R_{if} = 0.1037 \text{ k}\Omega$
(d) $R_{of} = 0.1186 \text{ k}\Omega$
- 12.63 (a) $A_{zf} = -R_F$
(b) $g_m = 4.6 \text{ mA/V}$
- 12.65 $R_F = 0.807 \text{ M}\Omega$
- 12.71 $T = 4.22$
- 12.73 (a) $f_{180} = 7.1 \times 10^4 \text{ Hz}$
(b) $\beta = 0.0205$
(c) $A_{vf}(0) = 48.54$
(d) Smaller
- 12.75 (b) $\beta = 4.7 \times 10^{-4}$
- 12.77 (a) $f_{180} = 10^6 \text{ Hz}$
(b) $\beta = 2.83 \times 10^{-3}$
(c) $A_{vf}(0) = 351$
- 12.79 (a) $f_{180} = 3.33 \times 10^5 \text{ Hz}$
(b) (i) $|T(f_{180})| = 0.156$
(ii) $\phi = -142.3^\circ$
(c) $A_f(0) = 50$
- 12.81 $\beta = 0.01428$
- 12.83 (a) $f_{180} = 1.42 \times 10^5 \text{ Hz}$, $|T(f_{180})| = 4.89$
(b) $f_{PD} = 4.85 \text{ Hz}$
- 12.85 (a) $\beta = 0.0199$, $f_{PD} = 711 \text{ Hz}$
(b) Phase margin = 29.8°
- 12.87 (a) $f = 0.976 \text{ MHz}$, phase margin = 73.4°
(b) $f'_{p1} = 2.67 \text{ Hz}$, phase margin = 85.4°
- 12.89 $f_{PD} = 577 \text{ Hz}$
- 13.11 $I_{\text{REF}} = 0.22 \text{ mA}$, $I_{C10} \cong 14.2 \mu\text{A}$,
 $I_{C6} = 7.10 \mu\text{A}$, $I_{C17} = 0.165 \text{ mA}$,
 $I_{C13A} = 0.055 \text{ mA}$
- 13.13 $P = 48.8 \text{ mW}$, $I = 1.63 \text{ mA}$
- 13.15 $I_{C13A} = 0.125 \text{ mA}$, $I_{R10} \cong 0.012 \text{ mA}$,
 $I_{C19} = 0.113 \text{ mA}$, $I_{C18} = 12.565 \mu\text{A}$,
 $V_{BE19} = 0.60185 \text{ V}$, $V_{BE18} = 0.54474 \text{ V}$,
 $I_{C14} = 113 \mu\text{A}$
- 13.17 $R_1 = 45.11 \text{ k}\Omega$, $R_2 = 51.56 \text{ k}\Omega$
- 13.19 $A_d = -882$
- 13.21 $A_v = 964,650$
- 13.23 $R_{eq} = 170 \text{ }\Omega$
- 13.25 (a) $R_{id} = 2.095 \text{ M}\Omega$
(b) $R_{id} = 2.80 \text{ M}\Omega$
- 13.27 (a) $f_{PD} = 10 \text{ Hz}$
(b) $C_F = 13.25 \text{ pF}$
- 13.29 (a) $R_{D1} = 7.87 \text{ k}\Omega$, $R_{D2} = 24.8 \text{ k}\Omega$,
 $R_S = 15 \text{ k}\Omega$
(b) (i) $A_d = 2.49$
(ii) $A_2 = -19.21$
(iii) $A_3 = 0.950$
(c) $A = -45.4$
- 13.31 (a) Original: $g_{m1} = g_{m2} = 0.09975 \text{ mA/V}$,
New: $g_{m1} = g_{m2} = 0.1995 \text{ mA/V}$
- 13.33 (a) $I_{set} = I_Q = I_{D7} = 0.1776 \text{ mA}$
(b) $A_d = 111.9$, $A_2 = -96.86$,
 $A = -10,839$
- 13.35 $f_{PD} = 83.9 \text{ Hz}$
- 13.37 $R_o = 0.63 \text{ M}\Omega$
- 13.39 (a) $I_{Q2} = 180 \mu\text{A}$
(b) $\left(\frac{W}{L}\right)_{8P} = 360$, $\left(\frac{W}{L}\right)_{8N} = 180$
- 13.41 (a) $I_{D6} = I_{D7} = 43.7 \mu\text{A}$
(b) $A_v = -35,350$
- 13.43 $\left(\frac{W}{L}\right)_1 = \left(\frac{W}{L}\right)_P = 4.06$, $\left(\frac{W}{L}\right)_N = 1.85$
- 13.45 (a) $A_d = 136.1$
(b) $R_o = 92.6 \text{ k}\Omega$
(c) $f_{PD} = 343.7 \text{ kHz}$, GBW = 46.8 MHz
- 13.47 (a) $\left(\frac{W}{L}\right)_5 = \left(\frac{W}{L}\right)_6 = 16.4$
(c) GBW = 27.2 MHz
- 13.49 (a) $R_S = 10 \text{ k}\Omega$, $R_{C1} = 3.4 \text{ k}\Omega$,
 $R_{C2} = 12.39 \text{ k}\Omega$
(b) $A_d = 15.60$

CHAPTER 13

- 13.1 (a) $R_{D2} = 15 \text{ k}\Omega$, $R_{D1} = 9 \text{ k}\Omega$
(b) (i) $A_d = 2.846$, (ii) $A_2 = -12$
(c) $A = -34.15$
- 13.5 (a) $A_v = -1.59 \times 10^6$
(b) $R_{id} = 208 \text{ k}\Omega$
(c) GBW = 12.3 MHz
- 13.7 $v_d \cong 56.4 \text{ V}$
- 13.9 (a) $R_5 = 17.13 \text{ k}\Omega$, $R_4 = 2.438 \text{ k}\Omega$,
 $V_{EB12} = V_{BE11} = 0.7184 \text{ V}$,
 $V_{BE10} = 0.6453 \text{ V}$
(b) $I_{\text{REF}} = 0.5137 \text{ mA}$, $I_{C10} \cong 30.22 \mu\text{A}$
(c) I_{REF} : 2.74%; I_{C10} : 0.733%

- (c) $A_2 = -20.97$
 (d) $A_3 = 0.95$
 (e) $A = -310.8$
- 13.51 (a) For NMOS: $g_{mN} = 1.0 \text{ mA/V}$,
 $r_{oN} = 400 \text{ k}\Omega$
 For BJT: $r_{o2} = 800 \text{ k}\Omega$
 (b) $A_d = 266.7$
- 13.53 $\left(\frac{W}{L}\right)_N = 6.16, \left(\frac{W}{L}\right)_P = 13.6$
- 13.55 $K_p = 0.234 \text{ mA/V}^2$
- 13.57 $A_d = 10.38, |A_{v2}| = 1917, |A_v| = 19,895$
- 13.59 $I_{DSS} = 0.8 \text{ mA}$
- 13.61 (a) Set $V_P = 3 \text{ V}, V_{ZK} = 3 \text{ V}$,
 $R_3 = 24 \text{ k}\Omega$, Set $I_{DSS} = 0.2 \text{ mA}$
 (b) $R_4 = 22.8 \text{ k}\Omega$
- CHAPTER 14**
- 14.1 $v_i(\max)_{\text{rms}} = 39.77 \text{ mV}$
- 14.3 (1) $v_O = 2 \text{ V}$, (2) $v_2 = 12.5 \text{ mV}$,
 (3) $A_{OL} = 2 \times 10^4$, (4) $v_1 = 8 \mu\text{V}$,
 (5) $A_{OL} = 1000$
- 14.5 (a) (i) $A_{CL} = 7.90863$
 (ii) -0.03956%
 (b) (i) $A_{CL} = 7.84966$
 (ii) -0.785%
- 14.7 $A_{OL} = 8.9991 \times 10^5$
- 14.9 (a) $9.98 \leq |A| \leq 10.02$
 (b) $9.969 \leq |A| \leq 10.009$
- 14.11 $A_{OL} = 49,950$
- 14.13 (a) $A_{CL} = 4.927, R_{if} = 2.687 \text{ M}\Omega$,
 $R_{of} = 3.16 \Omega$
 (b) -0.05%
- 14.15 $v_O = -1.993v_{I1} - 3.986v_{I2}; -0.35\%$
- 14.17 (a) $R_{if} = 99.1 \Omega$
 (b) $R_{of} = 18.4 \Omega$
 (c) $A_{CL} = 0.650$
 (d) 0.650
- 14.19 $f_{PD} = 7.94 \text{ Hz}, \text{GBW} = 7.94 \times 10^5 \text{ Hz}$
- 14.21 $f_{PD} = 448 \text{ Hz}$
- 14.23 (a) $A_{CLO} = -9.9978$,
 $f_{3-\text{dB}} = 150.033 \text{ kHz}$
 (b) $A_{CLO} = -999.34, f_{3-\text{dB}} = 76.49 \text{ kHz}$
- 14.25 $|A_v|_{\max} = 50$
- 14.27 (a) $V_{PO} = 5.09 \text{ V}$
 (b) $V_{PO} = 8 \text{ V}$
- 14.29 (a) $V_{PO} = 5.0 \text{ V}$
 (b) $V_{PO} = 23.87 \text{ V}$
- 14.33 (a) $I_{S4} = 5 \times 10^{-15} \text{ A}$
 (b) $I_{S4} = 4.939 \times 10^{-15} \text{ A}$
 (c) $I_{S4} = 4.811 \times 10^{-15} \text{ A}$
- 14.35 (a) $-0.360 \leq v_O \leq -0.240 \text{ V}$
 (b) $-3.06 \leq v_O \leq -2.94 \text{ V}$
- 14.37 $V_o = \frac{V_i}{RC} \cdot t, t = 10^3 \text{ s}$
- 14.39 At V^+ node, $v_O = 0.8645 \text{ V}$,
 At center, $v_O = 0.9545 \text{ V}$,
 At V^- node, $v_O = 1.0445 \text{ V}$
- 14.41 $\frac{i_{C1}}{i_{C2}} = 1.0155$
- 14.43 (a) $v_O = -0.30 \text{ V}$
 (b) $v_O = -0.50 \text{ V}$
 (c) $v_O = -0.10 \text{ V}$
 (d) $v_O = -1.30 \text{ V}$
- 14.45 (a) $v_O = -0.010 \text{ V}$
- 14.47 (a) $v_{O1} = 0.10 \text{ V}, v_{O2} = -0.45 \text{ V}$
- 14.49 (a) $v_{O1} = 0.15 \text{ V}, v_{O2} = 0.15 \text{ V}$,
 $v_{O3} = -0.09 \text{ V}$
 (b) $R_A = 8.33 \text{ k}\Omega, R_B = 10 \text{ k}\Omega$
 (c) $v_{O1} = \pm 0.015 \text{ V}, v_{O2} = \pm 0.015 \text{ V}$,
 $v_{O3} = \pm 0.021 \text{ V}$
- 14.51 (a) $4 \leq v_O \leq 76 \text{ mV}$
 (b) $-39 \leq v_O \leq 39 \text{ mV}$
 (c) $2.161 \leq v_O \leq 2.239 \text{ V}$
- 14.53 (a) For offset voltage,
 $|v_{O1}| = 110 \text{ mV}, |v_{O2}| = 610 \text{ mV}$
 For bias currents,
 $v_{O1} = 0.31 \text{ V}, v_{O2} = -1.51 \text{ V}$
- 14.55 For circuit (a), $v_O = 9 \text{ mV}$;
 For circuit (b), $v_O = -1.0815 \text{ V}$
- 14.57 (a) $v_{O1} = 6 \text{ mV}, v_{O2} = 28 \text{ mV}$;
 (b) $v_{O1} = 6.495 \text{ mV},$
 $v_{O2} = 30.31 \text{ mV}$
- 14.59 (a) $v_{O1} = 0.10 \text{ V}, v_{O2} = 0.12 \text{ V}$;
 (b) $v_{O1} = 0.105 \text{ V},$
 $v_{O2} = 0.166 \text{ V}$;
 (c) Due to $I_B, v_{O1} = 0.125 \text{ V}$,
 $v_{O2} = 0.15 \text{ V}$;
 Due to $I_{OS}, v_{O1} = 0.133 \text{ V}$,
 $v_{O2} = 0.224 \text{ V}$
- 14.61 (a) $x = 0.4728\%$
 (b) $x = 0.0267\%$

CHAPTER 15

15.1 (a) Noninverting amplifier:

Set $R_1 = 30 \text{ k}\Omega$, $R_2 = 210 \text{ k}\Omega$ At noninverting terminal, let input
 $C = 0.001 \mu\text{F}$, then $R = 5.305 \text{ k}\Omega$

(b) Inverting amplifier:

Set $R_1 = 15 \text{ k}\Omega$, $R_2 = 300 \text{ k}\Omega$;
Capacitor in series with R_1 ,
 $C = 530.5 \text{ pF}$

15.3 6-pole filter

15.5 (a) Let $R = 20 \text{ k}\Omega$, then $C = 397.9 \text{ pF}$;
 $C_1 = 1411 \text{ pF}$, $C_2 = 553.9 \text{ pF}$,
 $C_3 = 80.5 \text{ pF}$

- (b) (i) $|T| = -0.0673 \text{ dB}$,
(ii) $|T| = -0.711 \text{ dB}$,
(iii) $|T| = -3.0 \text{ dB}$,
(iv) $|T| = -6.83 \text{ dB}$,
(v) $|T| = -10.9 \text{ dB}$

15.7 9-pole filter

15.9 $N = 35$, $f_{3-\text{dB}} = 12.25 \text{ kHz}$ 15.11 (a) $|T| = -9.31 \text{ dB}$
(b) $|T| = -14.8 \text{ dB}$
(c) $|T| = -20.5 \text{ dB}$

15.13 3-pole filter

15.15 (b) (i) $|A_v|_{\max} = 28.3$
(ii) $f_o = 5.305 \text{ kHz}$
(iii) $f_1 = 5.296 \text{ kHz}$, $f_2 = 5.315 \text{ kHz}$ 15.17 (a) $T(s) = \left(1 + \frac{R_2}{R_1}\right) \left[\frac{1 + s(R_1 \| R_2)C}{1 + sR_2C} \right]$,
 $f_{3-\text{dB}1} = \frac{1}{2\pi R_2 C}$,

$$f_{3-\text{dB}2} = \frac{1}{2\pi(R_1 \| R_2)C}$$

(b) $T(s) = \left(1 + \frac{R_2}{R_1}\right) [1 + s(R_1 \| R_2)C]$,
 $f_{3-\text{dB}} = \frac{1}{2\pi(R_1 \| R_2)C}$

15.19 Two noninverting amplifiers:

Let $R_2 = 250 \text{ k}\Omega$, $R_1 = 54.1 \text{ k}\Omega$

For high-pass filter: Set

 $R = 250 \text{ k}\Omega$, $C = 31.8 \text{ pF}$

For low-pass filter: Set

 $R = 250 \text{ k}\Omega$, $C = 0.00424 \mu\text{F}$ 15.21 (a) $Q = 10 \text{ pC}$ (b) $I_{eq} = 1 \mu\text{A}$ (c) $t = 4.61 \times 10^{-8} \text{ s}$ 15.23 (a) $\tau = 60 \mu\text{s}$ (b) $\Delta v_O = 0.167 \text{ V}$

(c) 78 clock pulses

15.25 (a) $f_o = \frac{1}{2\pi R \sqrt{C(C + 2C_V)}}$
(b) $480 \leq f_o \leq 919 \text{ kHz}$ 15.29 (a) $f_o = \frac{1}{2\pi RC} \cdot \sqrt{\frac{3R + 2R_V}{R_V}}$
(b) $\frac{R_F}{R} = 2$
(c) $13.5 \leq f_o \leq 16.84 \text{ kHz}$ 15.31 (a) $f_o = \frac{1}{2\pi RC} \cdot \sqrt{4\left(\frac{R}{R \| R_V}\right) + 2}$
(b) $19.45 \leq f_o \leq 22.66 \text{ kHz}$ 15.33 (a) $\frac{R_1}{R_1 + R_2} = \frac{sRL}{sRL + (R + sL)^2}$
(b) $f_o = \frac{R}{2\pi L}$
(c) $\frac{R_2}{R_1} = 2$ 15.35 $C_2 \cong 0.06 \mu\text{F}$, $L = 13.8 \mu\text{H}$ 15.37 (a) $f_o = \frac{1}{2\pi \sqrt{C(L_1 + L_2)}}$
(b) $\frac{L_1}{L_2} = g_m R$ 15.39 $T(s) = \left(1 + \frac{R_2}{R_1}\right) \left[\frac{sRC}{s^2 R^2 C^2 + 3sRC + 1} \right]$,
 $f_o = \frac{1}{2\pi RC} \cdot \frac{R_2}{R_1} = 2$ 15.41 $T(s) = \left(1 + \frac{R_2}{R_1}\right) \left[\frac{sRL}{s^2 L^2 + 3sRL + R^2} \right]$,
 $f_o = \frac{R}{2\pi L} \cdot \frac{R_2}{R_1} = 2$ 15.43 $R_1 = 1.47 \text{ k}\Omega$, $R_2 = 73.53 \text{ k}\Omega$ 15.45 (a) $V_{TH} = 0.4 \text{ V}$, $V_{TL} = -0.4 \text{ V}$ (b) For $33.333 \leq t \leq 33.439 \text{ ms}$,
 $v_O = +10 \text{ V}$ For $33.439 \leq t \leq 66.667 \text{ ms}$, $v_O = -10 \text{ V}$ 15.47 (a) $V_{TH} = \frac{\frac{V_{\text{REF}}}{R_3} + \frac{V_P}{R_2}}{\frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3}}$,
 $V_{TL} = \frac{\frac{V_{\text{REF}}}{R_3} - \frac{V_P}{R_2}}{\frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3}}$

- (b) $R_1 = 10.17 \text{ k}\Omega$, $R_2 = 600 \text{ k}\Omega$
(c) $V_{TH} = -4.9 \text{ V}$, $V_{TL} = -5.1 \text{ V}$
- 15.49 (a) $V_S = \left(\frac{R_2}{R_1 + R_2} \right) \cdot V_{\text{REF}}$,
 $V_{TH} = V_S + \left(\frac{R_1}{R_1 + R_2} \right) \cdot V_H$
 $V_{TL} = V_S + \left(\frac{R_1}{R_1 + R_2} \right) \cdot V_L$
(b) $R_1 = 4 \text{ k}\Omega$, $R_2 = 188 \text{ k}\Omega$,
 $V_{\text{REF}} = -1.787 \text{ V}$
(c) (i) $|i| = 71.8 \mu\text{A}$, (ii) $|i| = 53.2 \mu\text{A}$
- 15.51 (a) $V_{TH} = 0.175 \text{ V}$, $V_{TL} = -0.175 \text{ V}$
(c) (i) $I_{D1} = 0$, $I_{D2} = 0.458 \text{ mA}$,
 $I_{R2} = -0.465 \text{ mA}$, $I_{R3} = 7 \mu\text{A}$
(ii) $I_{D1} = 0.458 \text{ mA}$, $I_{D2} = 0$,
 $I_{R2} = 0.465 \text{ mA}$, $I_{R3} = -7 \mu\text{A}$
- 15.53 (a) $V_{\text{REF}} = 3.6 \text{ V}$
(b) Let $R_2 = 90 \text{ k}\Omega$, $R_1 = 10 \text{ k}\Omega$
(c) $v_O = -5.1 \text{ V}$
- 15.55 (a) $v_O = (V_{\text{REF}} + 2V_\gamma)$,
 $v_O = -(V_{\text{REF}} + 2V_\gamma)$,
 $V_{TH} = \frac{R_1}{R_2}(V_{\text{REF}} + 2V_\gamma)$,
 $V_{TL} = -\frac{R_1}{R_2}(V_{\text{REF}} + 2V_\gamma)$
- 15.57 (a) $v_X = 5 - 8.333e^{-t/\tau_X}$ for $0 \leq t \leq t_1$;
 $v_X = -10 + 11.667e^{-(t-t_1)/\tau_X}$ for
 $t_1 \leq t \leq T$
(b) $f = 847 \text{ Hz}$
- 15.59 For $0 \leq t \leq t_1$, $t_1 = 1.1 \text{ ms}$,
 $v_Y = 10(1 - e^{-t/\tau_Y})$, $\tau_Y = 4 \times 10^{-5} \text{ s}$
- 15.61 (a) Set $C_X = 0.01 \mu\text{F}$, $R_X = 23.84 \text{ k}\Omega$
(b) $v_Y < v_X$
(c) $133 \mu\text{s}$
- 15.63 (a) Let $C = 50 \mu\text{F}$, $R = 1.09 \text{ M}\Omega$
(b) $t \cong 38.7 \text{ ms}$
- 15.65 $R_B = 50 \text{ k}\Omega$, $C = 144.3 \text{ pF}$
- 15.67 $1.40 \leq f_o \leq 2.72 \text{ kHz}$
 $50.5\% \leq \text{duty cycle} \leq 98.1\%$
- 15.69 (a) $P = 3.7 \text{ W}$
(b) $V^+ \cong 19 \text{ V}$
(c) $V_P = 8.6 \text{ V}$
- 15.71 (a) $A_v = 1 + \frac{R_2}{R_1} + \frac{R_3}{R_1}$
(b) $R_1 = 50 \text{ k}\Omega$, $R_2 = 200 \text{ k}\Omega$,
 $R_3 = 250 \text{ k}\Omega$
(c) $|v_{O1}| = |v_{O2}| = 10 \text{ V}$, $i_L = 1 \text{ A}$
- 15.73 (a) Set $R_2 = R_3 = R_4 = 100 \text{ k}\Omega$,
 $R_1 = 8.69 \text{ k}\Omega$
(b) (i) $\bar{P}_L = 14.4 \text{ W}$, (ii) $R_L = 20 \Omega$
(c) $\bar{P}_L = 7.2 \text{ W}$
- 15.75 (a) $R_{of} = 4 \text{ m}\Omega$
(b) $\Delta V_O = 12 \text{ mV}$
- 15.77 (a) $R_4 = 31.5 \Omega$
(b) $R_{12} = 5.57 \text{ k}\Omega$
- 15.79 (a) $7.659 \leq V_O \leq 22.976 \text{ V}$
(b) $4.17 \times 10^{-4}\%$

CHAPTER 16

- 16.1 (a) $\left(\frac{W}{L} \right) = 2.91$
(b) $V_{I_t} = 1.172 \text{ V}$, $V_{O_t} = 0.672 \text{ V}$
(c) $i_{D,\max} = 80 \mu\text{A}$, $P_{D,\max} = 264 \mu\text{W}$
- 16.3 (a) $R = 41.6 \text{ k}\Omega$, $\left(\frac{W}{L} \right) = 0.193$
(b) $0.5 \leq V_{GS} \leq 2.38 \text{ V}$
- 16.5 (a) $\frac{K_D}{K_L} = 4.26$
(b) $\frac{K_D}{K_L} = 5.4$
(c) $P = 0.759 \text{ mW}$
- 16.7 $\frac{(W/L)_D}{(W/L)_L} = 5.4$
- 16.9 (a) Driver: $V_{I_t} = 0.8578 \text{ V}$,
 $V_{O_t} = 0.3578 \text{ V}$
Load: $V_{I_t} = 0.8578 \text{ V}$, $V_{O_t} = 2.5 \text{ V}$
(b) $v_O = 0.0230 \text{ V}$
(c) $i_{D,\max} = 64 \mu\text{A}$, $P_{D,\max} = 211 \mu\text{W}$
- 16.11 (a) $\left(\frac{W}{L} \right)_L = 2.47$, $\left(\frac{W}{L} \right)_D = 5.04$
(b) Driver: $V_{I_t} = 0.720 \text{ V}$, $V_{O_t} = 0.420 \text{ V}$
Load: $V_{I_t} = 0.720 \text{ V}$, $V_{O_t} = 1.20 \text{ V}$
(c) $P_{D,\max} = 80 \mu\text{W}$, $v_O = 0.0297 \text{ V}$
- 16.13 (a) (i) $P = 0$, (ii) $P = 1.16 \text{ mW}$
(b) (i) $P = 0$, (ii) $P = 0.825 \text{ mW}$
(c) (i) $P = 0$, (ii) $P = 0.20 \text{ mW}$
- 16.15 (a) $v_{O1} = 0.0431 \text{ V}$, $v_{O2} = 5 \text{ V}$
(b) $v_{O1} = 5 \text{ V}$, $v_{O2} = 0.0431 \text{ V}$
- 16.17 (a) $v_I = 0.6683 \text{ V}$
(b) $v_I = 0.6057 \text{ V}$
- 16.19 (a) $\frac{K_D}{K_L} = 2.04$

- (b) $\left(\frac{W}{L}\right)_L = 0.667, \left(\frac{W}{L}\right)_D = 1.36$
(c) $v_O = 0.0329 \text{ V}$
- 16.21 (a) $\left(\frac{W}{L}\right)_{ML1} = 0.417, \left(\frac{W}{L}\right)_{MD1} = 1.35$
(b) $v_{O2} = 0.0330 \text{ V}$
- 16.23 (a) $v_O \cong 0.135 \text{ V}$
(b) $v_{DSX} \cong 0.067 \text{ V}, v_{DSY} \cong 0.068 \text{ V}, v_{GSX} = 5 \text{ V}, v_{GSY} \cong 4.933 \text{ V}$
- 16.25 $\bar{Y} = [A \text{ OR } (B \text{ AND } C)] \text{ AND } D$
- 16.27 $Y = \overline{(A+B)(C+D)}$
- 16.29 (a) For NMOS circuit,
 M_{NA} in parallel with $[M_{NB}$ in series
with $(M_{NC}$ in parallel with $M_{ND})]$
(b) $\left(\frac{W}{L}\right)_A = 1.735, \left(\frac{W}{L}\right)_{B,C,D} = 3.47$
- 16.31 (a) $V_{It} = 1.25 \text{ V}, V_{OPT} = 1.65 \text{ V}, V_{ONt} = 0.85 \text{ V}$
(c) For $v_I = 1.1 \text{ V}, v_O = 2.214 \text{ V}$
For $v_I = 1.4 \text{ V}, v_O = 0.286 \text{ V}$
- 16.33 (a) (i) $V_{It} = 1.707 \text{ V}$
(ii) $v_I = 1.26 \text{ V}$
(iii) $v_I = 2.157 \text{ V}$
(b) (i) $V_{It} = 1.25 \text{ V}$
(ii) $v_I = 0.855 \text{ V}$
(iii) $v_I = 1.61 \text{ V}$
- 16.35 (a) $1.7 \leq v_{O1} \leq 3.3 \text{ V}$
(b) $v_{O3} = 5 \text{ V}, v_{O1} = 2.78 \text{ V}, v_I = 2.5 \text{ V}$
- 16.37 (a) $i_{D,\text{peak}} = 24.2 \mu\text{A}$
(b) $i_{D,\text{peak}} = 29.34 \mu\text{A}$
(c) $i_{D,\text{peak}} = 33.23 \mu\text{A}$
- 16.39 (a) $P = 12.5 \mu\text{W}$
(b) $P = 6.48 \mu\text{W}$
- 16.41 (a) $P = 3 \times 10^{-7} \text{ W}$
(b) (i) $C_L = 0.0024 \text{ pF}$
(ii) $C_L = 0.00551 \text{ pF}$
(iii) $C_L = 0.0267 \text{ pF}$
- 16.43 (b) Let $\left(\frac{W}{L}\right)_N = 2, \left(\frac{W}{L}\right)_P = 4, V_{DD} = 5 \text{ V}$
For NMOS: $i_d = 0.45 \text{ mA}$
For PMOS: $i_d = 0.36 \text{ mA}$
- 16.45 (a) $V_{IL} = 0.8268 \text{ V}, V_{IH} = 1.2713 \text{ V}$
(b) $NM_L = 0.5858 \text{ V}, NM_H = 1.0439 \text{ V}$
- 16.47 (a) $v_C = 2.5 \text{ V}$
(b) $\left(\frac{W}{L}\right)_N = \frac{9}{2} \cdot \left(\frac{W}{L}\right)_P$
(c) $V_{It} = 1.65 \text{ V}$
- 16.49 (a) $\left(\frac{W}{L}\right)_N = 2, \left(\frac{W}{L}\right)_P = 16$
(b) $\left(\frac{W}{L}\right)_N = 4, \left(\frac{W}{L}\right)_P = 32$
- 16.51 (a) $\left(\frac{W}{L}\right)_N = 2, \left(\frac{W}{L}\right)_P = 12$
(b) $\left(\frac{W}{L}\right)_N = 4, \left(\frac{W}{L}\right)_P = 24$
- 16.53 (a) $Y = \overline{ABC + DE}$
(b) M_{NA}, M_{NB}, M_{NC} in series and
are in parallel with M_{ND}, M_{NE} in series
(c) $\left(\frac{W}{L}\right)_{AN,BN,CN} = 6, \left(\frac{W}{L}\right)_{DN,EN} = 4, \left(\frac{W}{L}\right)_P = 8$
- 16.55 (a) $Y = \overline{A + BC + DE}$
(b) M_{NA} in parallel with $(M_{NB}$ in series
with $M_{NC})$ in parallel with $(M_{ND}$ in
series with $M_{NE})$
(c) NMOS: $\left(\frac{W}{L}\right)_A = 2, \left(\frac{W}{L}\right)_{B,C,D,E} = 4$
PMOS: $\left(\frac{W}{L}\right)_{A,B,C,D,E} = 12$
- 16.57 (b) All $\left(\frac{W}{L}\right)_N = 3$, All $\left(\frac{W}{L}\right)_P = 6$
- 16.59 (a) $\bar{Y} = C(A + B)$
(b) All $\left(\frac{W}{L}\right)_N = 4, \left(\frac{W}{L}\right)_{CP} = 4, \left(\frac{W}{L}\right)_{AP,BP} = 8$
- 16.61 (a) State v_{O1} v_{O2}
1 5 0
2 5 0
3 5 0
4 5 0
5 5 0
6 0 5
(b) $v_{O2} = (v_A \text{ OR } v_B) \text{ AND } v_C$

- 16.63 (M_{NA} in series with $M_{M\bar{B}}$) and in parallel with ($M_{N\bar{A}}$ in series with M_{NB})
- 16.65 M_{NA} in series with (M_{NB} in parallel with M_{NC}) in series with (M_{ND} in parallel with M_{NE})
- 16.67 (a) (i) $v_O = 0$, (ii) $v_O = 2.9 \text{ V}$,
 (iii) $v_O = 2.5 \text{ V}$
 (b) (i) $v_O = 0$, (ii) $v_O = 1.4 \text{ V}$,
 (iii) $v_O = 1.4 \text{ V}$
- 16.69 (a) $v_{O1} = 2.1 \text{ V}, v_{O2} = 2.5 \text{ V}$
 (b) $\left(\frac{W}{L}\right)_1 = 12.1, \left(\frac{W}{L}\right)_3 = 1.09$
- 16.71 $Y = A + \bar{A}B$
 $Z = \bar{A} \bar{B}$
- 16.73 (a) (i) $Y = 0$, (ii) $Y = 2.5 \text{ V}$,
 (iii) $Y = 0$, (iv) $Y = 2.5 \text{ V}$
 (b) (i) $Y = 0$, (ii) $Y = 0$,
 (iii) $Y = 2.5 \text{ V}$, (iv) $Y = 2.5 \text{ V}$
 (c) Multiplexer
- 16.75 (a) (i) $Y = 0$, (ii) $Y = 2.5 \text{ V}$,
 (iii) $Y = 2.5 \text{ V}$, (iv) $Y = 0$
 (b) Exclusive OR
- 16.79 For $v_I = 1.5 \text{ V}, v_{O1} = 2.88 \text{ V}$,
 $v_O \cong 0$
 $v_I = 1.6 \text{ V}, v_{O1} = 2.693 \text{ V}$,
 $v_O = 0.00979 \text{ V}$
 $v_I = 1.7 \text{ V}$, Switching point,
 $v_I = 1.8 \text{ V}, v_{O1} = 0.607 \text{ V}$,
 $v_O = 3.298 \text{ V}$
- 16.81 (a) Positive edge when $CLK = 1$,
 $Q = \bar{\bar{D}} = D$
- 16.83 Not actually a J-K flip-flop
- 16.85 (a) 1 Megabit memory
 $= 1,048,576 \text{ cells}$
 $\Rightarrow 1024 \times 1024 \text{ memory}$
 Then 10 input row and
 column decoder lines necessary.
- (b) 250 K \times 4 bits \Rightarrow
 $262,144 \times 4 \text{ bits} \Rightarrow$
 $512 \times 512 \text{ memory}$
 Then 9 input row and
 column decoder lines
 necessary.
- 16.87 $t = 62.6 \text{ ns}$
- 16.89 $R = 0.512 \text{ M}\Omega, \left(\frac{W}{L}\right) = 0.797$

- 16.91 $\bar{Q} = 1.794 \text{ V}, Q = 0.370 \text{ V}$
- 16.97 (a) 7 bits
 (b) 0.02578125 V
 (c) 1100010
- 16.99 (a) $\Delta R_1 = 5.88\%$
 (b) $\Delta R_4 = 33.3\%$
- 16.101 (a) $I_1 = -0.50 \text{ mA},$
 $I_2 = -0.25 \text{ mA},$
 $I_3 = -0.125 \text{ mA},$
 $I_4 = -0.0625 \text{ mA},$
 $I_5 = -0.03125 \text{ mA},$
 $I_6 = -0.015625 \text{ mA}$
 (b) $\Delta v_O = 0.078125 \text{ V}$
 (c) $v_O = 1.484375 \text{ V}$
 (f) $\Delta v_O = 1.640625 \text{ V}$
- 16.103 64 resistors, 63 comparators
- 16.105 (a) 1010000000
 (b) 0101111101

CHAPTER 17

- 17.1 (a) $R_C = 2 \text{ k}\Omega$
 (b) (i) $v_{O1} = 0, v_{O2} = -0.4 \text{ V}$
 (ii) $v_{O1} = -0.4 \text{ V}, v_{O2} = 0$
 (c) $P = 0.36 \text{ mW}$
- 17.3 (a) $R_{C2} = 6 \text{ k}\Omega$
 (b) $R_{C1} = 4 \text{ k}\Omega$
 (c) $v_I = -0.0360 \text{ V}$
- 17.5 (a) $R_{C2} = 0.758 \text{ k}\Omega$
 (b) $R_{C1} = 0.658 \text{ k}\Omega$
 (c) For $v_{in} = -0.7 \text{ V}, v_{O1} = -0.7 \text{ V}$,
 $v_{O2} = -1.7 \text{ V}$
 For $v_{in} = -1.7 \text{ V}, v_{O1} = -1.7 \text{ V}$,
 $v_{O2} = -0.7 \text{ V}$
 (d) (i) $P = 21.8 \text{ mW}$, (ii) $P = 20.7 \text{ mW}$
- 17.7 (a) $R_1 = 10.5 \text{ k}\Omega$
 (b) $R_5 = R_6 = 17.5 \text{ k}\Omega$
 (c) $I_Q = 0.20 \text{ mA}, R_{C1} = 3.5 \text{ k}\Omega$
 (d) $I_Q = 0.20 \text{ mA}, R_{C2} = 3.5 \text{ k}\Omega$
- 17.9 $R_E = 1.375 \text{ k}\Omega, R_{C1} = 0.75 \text{ k}\Omega,$
 $R_{C2} = 1.031 \text{ k}\Omega, R_2 = R_3 = 2.25 \text{ k}\Omega$
- 17.11 (a) $V_R = -1.7 \text{ V}$
 (b) Logic 1 = -0.7 V , logic 0 = -2.7 V
 (c) $V_E = -2.4 \text{ V}$ for logic 0,
 $V_E = -1.4 \text{ V}$ for logic 1
 (d) $P = 39.9 \text{ mW}$ for logic 0 and logic 1
- 17.13 (a) $V_R = 2.6 \text{ V}$
 (b) $R_{C1} = 4.57 \text{ k}\Omega$

- (c) $R_{C2} = 5.05 \Omega$
 (d) $P = 1.60 \text{ mW}$
- 17.15 (a) $i_1 = 1.4 \text{ mA}, i_2 = 0.8 \text{ mA},$
 $i_3 = 0.14 \text{ mA}, i_4 = 0.14 \text{ mA},$
 $i_D = 0.74 \text{ mA}, v_O = -0.4 \text{ V}$
 (b) $i_1 = 1.4 \text{ mA}, i_2 = 0.153 \text{ mA},$
 $i_3 = 0.153 \text{ mA}, i_4 = 0.153 \text{ mA},$
 $i_D = 0, v_O = -0.0765 \text{ V}$
 (c) $i_1 = 1.6 \text{ mA}, i_2 = 0.14 \text{ mA},$
 $i_3 = 0.14 \text{ mA}, i_4 = 0.14 \text{ mA},$
 $i_D = 0, v_O = -0.07 \text{ V}$
 (d) $i_1 = 1.6 \text{ mA}, i_2 = 0.8 \text{ mA},$
 $i_3 = 0.153 \text{ mA}, i_4 = 0.153 \text{ mA},$
 $i_D = 0.953 \text{ mA}, v_O = -0.4 \text{ V}$
- 17.17 (a) Logic 0 = -0.4 V
 Logic 1 = 0
 (b) $v_{O1} = \overline{A + B}, v_{O2} = \overline{C + D},$
 $v_{O3} = (A + B) \cdot (C + D)$
- 17.19 (a) (i) $v_1 = 0.8 \text{ V}, v_O = 2.5 \text{ V},$
 $i_1 = 0.1417 \text{ mA}, i_2 = i_3 = 0$
 (ii) $v_1 = 1.5 \text{ V}, v_O = 0.1 \text{ V},$
 $i_1 = i_2 = 0.0833 \text{ mA}, i_3 = 0.2 \text{ mA}$
 (b) (i) $v_1 = 1.4 \text{ V}, v_I = 0.7 \text{ V}$
 (ii) $v_1 = 1.5 \text{ V}, v_I = 0.8 \text{ V}$
- 17.21 (i) $v' = 0.8 \text{ V}, i_1 = 0.525 \text{ mA},$
 $i_3 = i_4 = 0$
 (ii) $v' = 2.2 \text{ V}, i_1 = 0.35 \text{ mA},$
 $i_3 = 2.04 \text{ mA}, i_4 = 0.297 \text{ mA}$
- 17.23 (a) $v_1 = 2.3 \text{ V}, v_O = 0.1 \text{ V},$
 $i_1 = 0.675 \text{ mA}, i_2 = 1.7 \text{ mA},$
 $i_3 = 1.225 \text{ mA}, i_4 = 2.375 \text{ mA},$
 $i_5 = 0.08 \text{ mA}, i_{B2} = 2.295 \text{ mA}$
 (b) $N = 42$
- 17.25 $i_1 = 1.53 \text{ mA}, i_2 = 0.0589 \text{ mA},$
 $i_3 = 1.47 \text{ mA}, i_{Bo} = 1.37 \text{ mA},$
 $i_{Co} = 0.817 \text{ mA}$
- 17.27 (a) (i) $v_1 = 0.9 \text{ V}, v_O = 2.5 \text{ V},$
 $i_1 = 0.1333 \text{ mA}, i_2 = i_3 = 0$
 (ii) $v_1 = 1.5 \text{ V}, v_O = 0.1 \text{ V},$
 $i_1 = 0.0833 \text{ mA}, i_2 = 0.09167 \text{ mA},$
 $i_3 = 0.20 \text{ mA}$
 (b) (i) $v_1 = 1.4 \text{ V}, v_I = 0.6 \text{ V}$
 (ii) $v_1 = 1.5 \text{ V}, v_I = 0.7 \text{ V}$
- 17.29 (a) (i) $v' = 0.8 \text{ V}, i_1 = 0.156 \text{ mA},$
 $i_3 = i_4 = 0$
- (ii) $v' = 2.2 \text{ V}, i_1 = 0.06875 \text{ mA},$
 $i_3 = 0.5333 \text{ mA}, i_4 = 0.02875 \text{ mA}$
- (b) $N = 5$
 (c) $N = 5$
- 17.31 (a) (i) $i_{RB} = 1.025 \text{ mA}, i_{RCP} = i_{Bo} = 0,$
 $V_{out} = 5 \text{ V}$
 (ii) $i_{RB} = 0.70 \text{ mA}, i_{RCP} = 4.2 \text{ mA},$
 $i_{Bo} = 0.0837 \text{ mA}, V_{out} = 0.8 \text{ V}$
 (b) (i) $P = 5.145 \text{ mW}$
 (ii) $P = 25.4 \text{ mW}$
- 17.33 (a) (i) $v_O = 3.6 \text{ V},$ (ii) $v_O = 3.404 \text{ V},$
 (iii) $v_O = 1.11 \text{ V}$
 (b) $I_L = 34.05 \text{ mA}$
- 17.35 (a) $i_{B1} = 1.5 \text{ mA}, i_{B2} = 0,$
 $i_{B3} = 0.4 \text{ mA}, i_{C2} = 0,$
 $i_{C3} = 0.5 \text{ mA}$
 (b) $i_{B1} = 0.5 \text{ mA}, i_{B2} = 0.8 \text{ mA},$
 $i_{B3} = 0, i_{C2} = 7.5 \text{ mA},$
 $i_{C3} = 0$
- 17.37 (a) (i) $v_1 = 0.3 \text{ V}, v_O = 1.5 \text{ V},$
 $i_1 = 1.2 \text{ mA}, i_B = i_C = 0$
 (ii) $v_1 = 1.0 \text{ V}, v_O = 0.4 \text{ V},$
 $i_1 = 0.5 \text{ mA}, i_B = 0.465 \text{ mA},$
 $i_C = 0.9167 \text{ mA}$
 (b) (i) $v_1 = 1.0 \text{ V}, v_I = 0.7 \text{ V},$
 $i_B = i_C = 0$
 (ii) $v_1 = 1.0 \text{ V}, v_I = 0.7 \text{ V},$
 $i_B = 0.03667 \text{ mA}, i_C = 0.9167 \text{ mA}$
 (c) $N = 13$
- 17.39 (a) $R_{B1} = 18 \text{ k}\Omega, R_{C1} = 1.63 \text{ k}\Omega$
 (b) $v_{B1} = 0.7 \text{ V}, v_{B2} = 0, v_O \cong 1.8 \text{ V},$
 $i_B = i_C = 0$
 (c) $v_{B1} = 1.5 \text{ V}, v_{B2} = 0.7 \text{ V}, v_O = 0.4 \text{ V},$
 $i_{B1} = 0.0555 \text{ mA}, i_{C1} = 1.043 \text{ mA},$
 $i_{B2} = 0.10 \text{ mA}, i_{C2} = 0.40 \text{ mA}$
 (d) $N = 20$
- 17.41 (a) $P = 0.4875 \text{ mW}$
 (b) $P = 1.98 \text{ mW}$
 (c) $i_{SC} \cong 78 \text{ mA}$
- 17.43 (a) $i_{DN} = 0.1 \text{ mA}, i_{DP} = 0.289 \text{ mA},$
 $i_{C1} = 14.45 \text{ mA}, i_{C2} = 5 \text{ mA}$
 (b) $t = 5.19 \text{ ns}$
 (c) $t = 260 \text{ ns}$

Index

A

A/D converter, 1226, 1229–1232
ac circuit analysis, 44
ac collector current, 373
ac collector-emitter voltage, 373
ac common-emitter current gain, 380
ac equivalent circuit, 377–378
ac load line, 413–414
ac load line analysis, 413–420
ac ripple voltage, 68n
ac signal, 372
Acceptor impurity, 15
Active bandpass filter, 1122–1125
Active device, 2
Active filter, 1062–1074
 active network design, 1062–1063
 advantages, 1062
 defined, 1062
 general two-pole, 1064
 higher-order Butterworth filter, 1068–1070
 switched-capacitor filter, 1070–1074
 two-pole high-pass Butterworth filter, 1068
 two-pole low-pass Butterworth filter,
 1065–1067
Active load, 238, 411, 687. *See also* Integrated
 circuit biasing and active loads, 790
Active load device, 720
Active region, 287
Address decoder, 1211–1212
Admittance, 47
Advanced common-emitter amplifier concepts,
 411–412
Advanced low-power Schottky (ALS) inverter
 gate, 1295
Advanced Schottky TTL circuits, 1295–1296
Airplane, 616
All-CMOS folded cascode op-amp circuit, 1012
ALS inverter gate, 1295
Alterable ROM, 1209
Alternative ECL gates, 1270–1273
Aluminum arsenide, 10
Aluminum phosphide, 10
AM signal, 82
Amplification factor, 854
Amplifier
 antilog, 656
 BJT. *See* Basic BJT amplifier
 BJT application, as, 327–330
 bridge power, 1113
 CMOS common-gate, 251–253, 251–254
 CMOS common-source, 247–249
 CMOS source-follower, 249–251
 common-base, 431–435
 common-collector, 420–431
 common-drain, 227–233
 common-emitter, 398–412
 common-gate, 234–237
 common-source, 216–226

current, 397
diff-amp. *See* Differential and multistage
 amplifiers
difference, 645–649
emitter-follower, 420–431
error, 1116, 1120
exponential, 656
feedback, 854
FET. *See* Basic FET amplifier
generalized three-stage, 436
IC power, 1107–1113
inductively coupled, 587–588
instrumentation, 650–651
inverting, 627–635
JFET, 258–263
LM380, 1108–1111
log, 655
MOSFET, 206–216
MOSFET integrated circuit, 155
MOSFET small-signal, 168–169
multistage, 254–258, 436–442. *See* Differential
 and multistage amplifiers
NMOS-active loads, 247–254
NMOS depletion load, 244–247
NMOS enhancement load, 239–244
noninverting, 638–639
op-amp. *See* Ideal op-amp
OTA, 657–658
PA12, 1111–1113
power. *See* Output stages and power amplifiers
series-shunt, 873–879
single-stage integrated circuit MOSFET,
 238–254
summing, 636–638
transconductance, 397, 870
transformer-coupled common-emitter, 588–589
transformer/coupled emitter-follower, 589–591
transresistance, 397, 872
two-stage, 264–266
voltage, 397, 873–879
Amplifier frequency response, 470–472
Amplitude-modulated input signal, 82
Amplitude-modulated signal, 82
Analog circuit, 3, 206, 370
Analog electronics, 619
Analog signal, 3, 206, 370
Analog-to-digital (A/D) converter, 1226,
 1229–1232
AND logic circuit, 105
Answers to selected problems, 1337–1358
Antilog amplifier, 656
Applications and design of integrated circuits,
 1061–1140
 active filter. *See* Active filter
 design application (active bandpass filter),
 1122–1125
 555 circuit, 1100–1107
 IC power amplifiers. *See* Integrated circuit
 power amplifiers

nonsinusoidal oscillators/timing circuits. *See*
 Nonsinusoidal oscillators/timing circuits
oscillator. *See* Oscillator
Schmitt trigger circuit. *See* Schmitt trigger
 circuit
Approximate class-B circuit, 577
Aspect ratio, 708
Astable multivibrator, 1096, 1105–1107
Astable multivibrator 555 circuit, 1105
Attenuation, 640
Audio amplifier, 445–449
Avalanche breakdown, 31
Avalanche multiplication process, 31
Average power dissipation, 580

B

B-C junction, 287
B-E junction, 287
Back-gate transconductance, 215
Back-to-back Zener diodes, 1094
Band-reject filter, 1063
Bandgap energy, 12
Bandpass filter, 1063
Bandwidth, 481
Bandwidth extension, 857–859
Bardeen, John, 1
Barkhausen criterion, 1075
Base, 286
Base-collector (B-C) junction, 287
Base-collector voltage, 1259
Base-emitter (B-E) junction, 287
Base-emitter loop, 377
Basic BiCMOS differential amplifier, 805
Basic BiCMOS inverter, 1297
Basic bipolar diff-amp with active load, 1033
Basic bipolar difference amplifier, 1031
Basic BJT amplifier, 369–468
 ac load line analysis, 413–420
 bipolar linear amplifier. *See* Bipolar linear
 amplifier
 common-base amplifier, 431–435
 common-collector amplifier. *See* Common-
 collector amplifier
 common-emitter amplifier. *See* Common-
 emitter amplifier
 design application (audio amplifier), 445–449
 maximum symmetrical swing, 415–420
 multistage amplifier. *See* Multistage BJT
 amplifier
 power considerations, 442–445
Basic BJT differential pair, 754–779
 CMRR, 765–768
 common-mode input resistance, 775
 dc transfer characteristics, 758–762
 differential-/common-mode gains, 771–774
 differential-mode input resistance, 774–775

- Basic BJT differential pair—(*Cont.*)
 differential-mode voltage gain with emitter degeneration, 777–778
 g_m mismatch, 770
 one-sided output, 763–765
 R_c mismatch, 769–770
 small-signal equivalent circuit analysis, 762–765
 two-sided output, 761, 769–770
- Basic BJT differential-pair configuration, 755
- Basic common-base circuit, 431
- Basic common-emitter amplifier circuit, 399–400
- Basic common-emitter differential pair, 952
- Basic common-source configuration, 217–222
- Basic complementary push-pull output stage, 577
- Basic differential amplifier circuit, 1256
- Basic diode-transistor logic gate, 1277–1279
- Basic DTL NAND circuit operation, 1277–1279
- Basic ECL logic gate, 1258–1259
- Basic feedback concepts, 853–862
- Basic FET amplifier, 205–284
 common-drain amplifier, 227–233, 238
 common-gate amplifier, 234–237, 238
 common-source amplifier, 216–226
 design application (two-stage amplifier), 264–266
 JFET amplifier, 258–263
 MOSFET amplifier, 206–216
 multistage amplifier, 254–258
 single-stage IC amplifier. *See* Single-stage integrated circuit MOSFET amplifier
 source-follower amplifier, 227–233, 238
- Basic FET differential pair, 779–790
 dc transfer characteristics, 779–784
 differential-/common-mode input impedances, 784
 g_m mismatch, 788
 JFET differential amplifier, 788–789
 R_d mismatch, 787–788
 small-signal equivalent circuit analysis, 784–787
 two-sided output, 787–788
- Basic inverting Schmitt trigger, 1087–1090
- Basic MOSFET differential pair configuration, 779
- Basic random access memory architecture, 1210
- Basic series-pass voltage regulator, 84, 1116
- Basic three-transistor current source, 693–696
- Basic transistor applications, 323–330
- Basic TTL gate, 1280
- Basic TTL NAND circuit, 1282–1284
- Basic two-input ECL OR/NOR logic circuit, 1258
- Basic two-transistor current source, 688
- Basic two-transistor MOSFET current source
 current relationship, 707–708
 output resistance, 708–709
 reference current, 709–710
- Basic two-transistor NMOS current source, 707
- Battery charger, 70
- Beta cutoff frequency, 505
- Bias
 bipolar transistor biasing. *See* Bipolar transistor biasing
 BJT. *See* Bipolar transistor biasing
 constant-current, 170–175
 constant-current biasing, 170–175
 diode biasing, 592–594
 forward, 27
 IC. *See* Integrated circuit biasing and active loads
 integrated circuit biasing, 342–343
 linear amplifying device, 688
 positive and negative biasing, 339–341
 reverse, 25
 single base resistor biasing, 331–333
 V_{BE} multiplier, 594–596
 voltage divider biasing, 333–339
- Bias current compensation, 1043–1044
- Bias current effects, 1042–1043
- Bias-independent current source, 715–716
- Bias-independent MOSFET current mirror, 715
- Bias resistor, 590
- Bias stable, 336
- Bibliography. *See* Reading list
- BiCMOS circuits, 801–805
 basic amplifier stages, 801–803
 BiCMOS diff-amp, 804, 805
 current sources, 803–804
- BiCMOS Darlington pair configuration, 802
- BiCMOS diff-amp, 804, 805
- BiCMOS digital circuits, 1296–1298
- BiCMOS double cascode constant-current source, 803
- BiCMOS folded cascode amplifier, 982
- BiCMOS folded cascode op-amp, 982–983
- BiCMOS inverter, 1296–1297
- BiCMOS logic circuit, 1297–1298
- BiCMOS operational amplifier circuits, 981–989
 BiCMOS folded cascode op-amp, 982–983
 CA3140 BiCMOS circuit description, 983–984
 CA3140 dc analysis, 984–986
 CA3140 small-signal analysis, 986–989
- BiCMOS technology, 801
- Bipolar ac analysis, 384
- Bipolar cascode constant-current source, 803
- Bipolar cascode current mirror, 696
- Bipolar class-AB output stage, 582
- Bipolar common-emitter circuit, 719
- Bipolar Darlington pair configuration, 802
- Bipolar digital circuits, 1255–1314
 BiCMOS digital circuits, 1296–1298
 design application (static ECL gate), 1298–1299
- ECL. *See* Emitter-coupled logic (ECL)
 modified ECL circuit configurations, 1267–1276
- Schottky TTL. *See* Schottky transistor-transistor logic
- TTL. *See* Transistor-transistor logic
- Bipolar fuse-linked user-programmable ROM, 1223
- Bipolar inverter circuit, 326
- Bipolar junction transistor (BJT), 285–367
 amplifier, 327–330. *See* Basic BJT amplifier
 asymmetry, 287
 basic applications, 323–330
 basic transistor principle, 286
 bias. *See* Bipolar transistor biasing
 breakdown voltage, 299–301
 circuit symptoms/conventions, 293–295
 common-emitter circuit, 302–306
 common base bipolar circuits, 313–322
 cross section, 287
 current-voltage characteristics, 295–297
 current-voltage relationship, 294
 dc analysis, 301–322
 design application (diode thermometer with bipolar transistor), 348–350
 digital logic, 325–327
 doped region, 286
 frequency response, 502–514
 leakage currents, 298–299
 load line, 306–307
 modes of operation, 307–310, 350
 multistage circuits, 344–348
 npn transistor. *See* npn bipolar transistor
 pn junction, 286
 pnp transistor, 292–293
 power, 561–565
 simple geometry, 286
 small-signal two-port network, as, 378
 switch, 323–325
 voltage transfer characteristics, 310–313
- Bipolar linear amplifier, 371–396
 ac equivalent circuit, 377–378
 alternative form of equivalent circuit, 380
 base-emitter loop, 377
 bipolar transistor inverter circuit, 371
 collector-emitter loop, 377
 common-emitter current gain, 381
 data sheet, 393–396
 graphical analysis, 372–378
 h -parameters, 390–392
 harmonic distortion, 375–376
 hybrid- π equivalent circuit, 380
 hybrid- π equivalent circuit (Early effect), 385–389
 input base-emitter port, 378–379
 other small-signal parameters/equivalent circuits, 390–396
 small signal, 374–375
 small-signal hybrid- π equivalent circuit, 378–381
 small-signal voltage gain, 381–384
- T-model, 396
- Bipolar NOR logic circuit response, 326
- Bipolar NOR logic gate, 326
- Bipolar operational amplifier circuit, 812, 950–970
 bias circuit and input stage, 954–956
 circuit description, 950–953
 dc analysis, 953–960
 frequency response, 968–969
 gain stage, 952–953, 956–957, 963–966
 input diff-amp, 951–952
 input stage, 961–963
 output resistance, 966–968
 output stage, 953, 958–960
 overall gain, 966
 short-circuit protection circuitry, 960
 small-signal analysis, 961–968
- Bipolar transistor biasing, 330
 bias stability, 333–339
 integrated circuit biasing, 342–343
 positive and negative biasing, 339–341
 single base resistor biasing, 331–333
 voltage divider biasing, 333–339
- Bipolar transistor current sources, 688–707
 basic three-transistor current source, 693–696
 cascade current source, 696–697
 improved current-source circuits, 693–698
 multitransistor current mirror, 704–706
 two-transistor current source, 688–693
 Widlar current source, 698–704
 Wilson current source, 697–698
- Bipolar transistor inverter circuit, 371
- Bistable multivibrator, 1087
- BJT amplifier. *See* Basic BJT amplifier
- BJT current source circuits, 698
- BJT diff-amp with active load, 791–792
- BJT diff-amp with three-transistor active load, 806
- BJT differential amplifier with three-transistor active load, 792
- BJT Hartley oscillator, 1083
- BJT op-amp circuit, 811–815
- BJT small-signal equivalent circuit, 380
- Black, Harold, 852
- Bode, H., 474
- Bode plot, 474–479
- Bode plot of gain magnitude, 492
- Body effect, 143–144, 1160–1162
- Body-effect parameter, 144
- Bottom-up design process, 616
- Brattain, Walter, 1
- Break-point frequency, 475
- Breakdown, 31
- Breakdown effect, 144–145
- Breakdown voltage, 30–32, 52, 299–301
- Bridge circuit, 663, 664
- Bridge power amplifier, 1113

Bridge rectifier, 72
 Built-in potential barrier, 24
 Butterworth filter
 defined, 1065
 four-pole, 1069
 higher-order, 1068–1070
 N-th order, 1069
 three-pole, 1069
 two-pole high-pass, 1068
 two-pole low-pass, 1065–1067
 Butterworth *N*th-order high-pass filter, 1068–1070
 Butterworth *N*th-order low-pass filter, 1068–1070
 Bypass capacitor effects, 496–499

C

CA3140 BiCMOS circuit description, 983–984
 CA3140 BiCMOS data, 984
 CA3140 BiCMOS op-amp equivalent circuit, 983
 CA3140 dc analysis, 984–986
 CA3140 small-signal analysis
 frequency response, 988–989
 gain stage, 987
 input stage, 986–987
 overall gain, 988
 Capacitance
 depletion layer, 26
 diffusion, 46
 equivalent input gate, 519
 equivalent Miller, 816
 junction, 26
 Miller, 510–512, 518–520
 overlap, 515
 parasitic, 515
 Capacitance of parallel plate capacitor, 126n
 Capacitor, 1331
 compensation, 922
 coupling, 331, 399
 emitter bypass, 496
 filter, 76
 input coupling, 485–488
 MOS, 127, 128
 output coupling, 488–490, 490–492
 source bypass, 224–226
 Capacitor charging current, 78
 Capacitor impedance, 217, 399
 Capacitor values. *See* Standard resistor and capacitor values
 Carbon resistors, 1329–1330
 Carrier, 17
 Carrier velocity saturation, 1150
 Cascade, 175, 436
 Cascade circuit (multistage amplifier), 254–256
 Cascade configuration
 multistage BJT analysis, 436–438
 multitransistor circuit, 176–177
 Cascaded clocked CMOS logic circuit, 1193
 Cascaded-stage amplifiers. *See* Differential and multistage amplifiers
 Cascade, 178
 Cascade circuit (multistage amplifier), 256–257
 Cascode configuration
 multistage BJT analysis, 440–442
 multitransistor circuit, 178–179
 Cascode current mirror, 711–713
 Cascode current source, 696–697
 Channel inversion charge, 515
 Channel length modulation, 142, 1150
 Channel-length modulation, 515
 Channel length modulation factor, 1169
 Channel-length modulation parameter, 142
 Channel region, 128
 Chip select signal, 1210
 Circuit

ac equivalent, 377–378
 active load, 719–726
 analog, 206, 370
 BiCMOS, 801–805
 bipolar digital. *See* Bipolar digital circuits
 bipolar transistor inverter, 371
 BJT current-source, 698
 clamper, 94–97
 clipper, 91–94
 CMOS full-adder, 1207–1208
 CMOS op-amp, 970–981
 common-base, 431
 common-drain, 227
 common-emitter, 302–306
 common-emitter amplifier, 399
 common-gate, 234
 common-source, 146–153
 comparator, 1086
 diode. *See* Diode circuit
 diode logic, 104–105
 emitter-follower, 421
 equivalent preamplifier, 398
 expanded hybrid- π equivalent, 502–503
 555, 1100–1107
 high-frequency equivalent, 514–516
 hybrid- π equivalent, 380
 LED, 107–108
 MOSFET current-source, 711
 MOSFET digital. *See* MOSFET digital circuits
 MOSFET high-frequency, 519
 multi-MOSFET current-source, 711–715
 multiple-diode, 97–105
 multistage MOSFET, 175–179
 Norton equivalent, 396
 one-bit full-adder, 1207–1208
 op-amp. *See* Operational amplifier circuits
 op-amp voltage reference, 661
 phase-shift oscillator, 1078
 photodiode, 106
 piezoelectric crystal equivalent, 1083
 precision half-wave rectifier, 654
 rectifier. *See* Rectifier circuit
 reference, 1260–1262
 Schmitt trigger, 1084–1095
 simplified BJT op-amp, 811–815
 small-signal equivalent, 210–214, 378, 380, 399
 SRAM read/write, 1218–1220
 summing op-amp, 636
 switched-capacitor, 1072
 Thevenin equivalent, 333, 396
 two-port equivalent, 403
 voltage doubler, 83–84
 voltage-follower op-amp, 640
 Circuit element matching, 950
 Circuit load line, 37
 Circuits with active loads, 719–726
 dc analysis - BJT active load circuit, 720–722
 dc analysis - MOSFET active load circuit, 724–725
 discussion, 726
 voltage gain - BJT active load circuit, 722–724
 voltage gain - MOSFET active load circuit, 725
 Clammer, 94–97
 Clammer circuit, 94–97
 Clamping, 94
 Class-A operation, 572–576
 Class-A power amplifier, 586–591
 inductively coupled amplifier, 587–588
 transformer-coupled common-emitter amplifier, 588–589
 transformer-coupled emitter-follower amplifier, 589–591
 Class-AB operation, 581–585
 Class-AB output stage, 581
 Class-AB push-pull complementary output stages, 591–601

biasing (V_{BE} multiplier), 594–596
 Darlington configuration, 599–600
 diode biasing, 592–594
 input buffer transistors, 596–599
 Class-B operation, 576–581
 approximate class-B circuit, 577
 crossover distortion, 577–578
 idealized, 576–577
 idealized power efficiency, 579–581
 Class-C operation, 585–586
 Classical cascode stage, 976
 Clipper, 91–94
 Clipper circuit, 91–94
 Clock, 1105
 Clock frequency, 1071
 Clocked CMOS logic circuits, 1191–1193
 Closed-loop feedback system, 1116
 Closed-loop frequency response, 920–922
 Closed-loop voltage gain, 625, 628, 634, 864
 CMOS, 140–141
 CMOS cascode current-mirror op-amp, 980
 CMOS cascode current-mirror op-amp circuit, 980–981
 CMOS common-gate amplifier, 251–254
 CMOS common-source amplifier, 247–249
 CMOS current-mirror op-amp, 979
 CMOS current-mirror op-amp circuit, 979–980
 CMOS D-type flip-flop, 1206
 CMOS diff-amp, 821–824
 CMOS differential sense amplifier, 1219
 CMOS dynamic shift register, 1204
 CMOS folded cascode amplifier, 977
 CMOS full-adder circuit, 1207–1208
 CMOS inverter, 163–164, 1168–1182
 dc analysis, 1170–1176
 noise margin, 1178–1182
 p-channel MOSFET, 1168
 power dissipation, 1176–1178
 transistor sizing, 1175–1176, 1184–1185
 voltage transfer curve, 1171–1175
 CMOS inverter current, 1175–1176
 CMOS logic circuits, 1183–1191
 basic CMOS NOR and NAND gates, 1183
 clocked, 1191–1193
 complex, 1187–1190
 fanout, 1190
 propagation delay time, 1191
 transistor sizing, 1184–1187
 CMOS master-slave D flip-flop, 1207
 CMOS NAND logic gate, 1184
 CMOS NOR gate, 1183, 1186
 CMOS one-bit full adder, 1208
 CMOS operational amplifier circuits, 970–981
 CMOS cascode current-mirror op-amp circuit, 980–981
 CMOS current-mirror op-amp circuit, 979–980
 folded cascode CMOS op-amp circuit, 976–979
 MC14573 CMOS op-amp circuit, 970–973
 three-stage CMOS op-amp, 973–976
 CMOS pass networks, 1202
 CMOS R-S flip-flop, 1205
 CMOS source-follower amplifier, 249–251
 CMOS SRAM cells, 1214–1217
 CMOS static exclusive-OR logic gate, 1189
 CMOS static RAM cell, 1215
 CMOS transmission gate, 1200–1201
 CMRR, 648, 765–768, 1046–1047
 Collector, 286
 Collector-emitter loop, 377
 Collector-emitter voltage, 1259
 Colpitts oscillator, 1082–1083
 Common-anode display, 107
 Common-base amplifier, 431–435
 Common-base circuit, 431, 525
 Common-base circuit configuration, 295
 Common-base current gain, 289

Common-collector amplifier, 420–431
 input/output impedance, 423–426
 other basic amplifiers, compared, 435–436
 small-signal current gain, 426–431
 small-signal voltage gain, 420–423
 Common-drain amplifier, 227–233, 238
 Common-drain circuit, 227
 Common-emitter amplifier, 398–412, 573
 advanced common-emitter amplifier concepts, 411–412
 basic amplifier circuit, 399–400
 circuit with emitter bypass capacitor, 407–410
 circuit with emitter resistor, 400–407
 other basic amplifiers, compared, 435–436
 Common-emitter bipolar transistor, 391
 Common-emitter circuit, 294, 302–306, 373
 Common-emitter configuration, 290
 Common-emitter current gain, 290–291, 381
 Common-emitter *h*-parameter, 390
 Common-emitter npn bipolar transistor, 502
 Common-emitter npn transistor, 391
 Common-gate amplifier, 234–237, 238
 Common-gate circuit, 234
 Common-mode gain, 648
 Common-mode gain frequency response, 818
 Common-mode input resistance, 775
 Common-mode input signal, 623, 648
 Common-mode input voltage, 648, 754
 Common-mode rejection, 623
 Common-mode rejection ratio (CMRR), 648, 765–768, 1046–1047
 Common-source amplifier, 216–226
 basic configuration, 217–222
 source bypass capacitor, 224–226
 source resistor, 222–224
 summary, 238
 Common-source amplifier with source resistor, 222–224
 Common-source circuit with source bypass capacitor, 224–226
 Common-source NMOS transistor circuit, 220
 Comparator, 1084–1087
 inverting comparator circuit, 1086
 noninverting comparator circuit, 1086
 open-loop comparator circuit, 1087
 threshold, 1101
 trigger, 1103
 Compensation capacitor, 922
 Complementary MOS (CMOS), 140–141
 Complementary MOS (CMOS) inverter, 163–164
 Complementary MOSFET, 140–141
 Complementary push-pull output stage, 577
 Complete amplifier system, 664
 Complex CMOS logic circuit, 1187
 Complex frequency *s*, 472
 Composite transconductance, 802
 Compound semiconductor, 10
 Computer memory, 1208–1225
 address decoder, 1211–1212
 CMOS SRAM cells, 1214–1217
 DRAM cells, 1220–1221
 memory architecture, 1210
 NMOS SRAM cells, 1212–1214
 RAM, 1209, 1212–1221
 ROM, 1209–1210, 1221–1225
 SRAM read/write circuitry, 1218–1220
 Conduction parameter, 133, 1169
 Conductivity, 18
 Constant-current biasing, 170–175
 Constant-current source, 289, 687, 688
 Conversion efficiency, 581
 Conversion factors, 1315
 Corner frequency, 475, 481, 487
 Counting A/D converter, 1230–1231
 Coupling capacitor, 331, 399

Coupling capacitor effects, 485–492
 Covalent bond, 11
 Cross-coupled NMOS sense amplifier, 1219
 Crossover distortion, 577–578, 949
 Crystal matrix, 11
 Crystal oscillator, 1083–1084
 Current amplifier, 397
 Current buffer, 434
 Current divider, 486
 Current mirror, 170, 688
 Current (shunt-series) amplifier, 879–886
 discrete circuit representation, 882–886
 op-amp circuit representation, 880–881
 simple discrete circuit representation, 881–882
 Current-to-voltage converter, 641–642
 Current transfer function, 472
 Cut-in voltage, 38
 Cutoff, 291
 Cutoff frequency, 506–507, 519
 Cyclical frequency, 470n

D

D/A converter, 1226, 1227–1229
 D flip-flop, 1206–1207
 Darlington configuration, 439, 599–600
 Darlington pair configuration, 439–440, 801
 Darlington pair gain stage, 806
 Data converter, 1226–1232
 A/D converter, 1229–1232
 basic A/D and D/A concepts, 1226–1227
 counting A/D converter, 1230–1231
 D/A converter, 1227–1229
 dual-slope A/D converter, 1231–1232
 flash A/D converter, 1229–1230
 R-2R ladder network D/A converter, 1228–1229
 weighted-resistor 4-bit D/A converter, 1227–1228
 Data sheets, 393–396. *See also* Selected manufacturers' data sheets
 dc load line, 413, 573
 dc power supply, 108–110
 dc quantity, 4, 372
 dc transfer characteristics
 BJT differential pair, 758–762
 FET differential pair, 779–784
 Dead band, 577
 Decade, 475
 Decoder, 1211
 Demodulated output signal, 82
 Depletion layer capacitance, 26
 Depletion-load inverter, 1160
 Depletion mode, 138
 Depletion-mode JFET connected as current source, 716
 Depletion-mode MOSFET, 138
 Depletion-mode NMOS device, 161
 Depletion region, 24
 Desensitivity factor, 857
 Design. *See* Applications and design of integrated circuits; Electronic design
 Design applications
 active bandpass filter, 1122–1125
 audio amplifier, 445–449
 CMOS diff-amp, 821–824
 dc power supply, 108–110
 diode thermometer, 54–56
 diode thermometer with bipolar transistor, 348–350
 diode thermometer with MOS transistor, 190–192
 electronic thermometer with instrumentation amplifier, 665–668
 MOSFET feedback circuit, 924–926
 NMOS current source, 734–736
 offset voltage compensation network, 1047–1049
 output stage using MOSFETs, 601–603
 static CMOS logic gate, 1232–1234
 static ECL gate, 1298–1299
 two-stage amplifier, 264–266
 two-stage amplifier with coupling capacitors, 537–539
 two-stage CMOS op-amp, 992–995
 Detector, 82–83
 Detector circuit, 82
 Diff-amp. *See* Differential and multistage amplifiers
 Diff-amp circuit, 1256–1258
 Diff-amp frequency response, 815–821
 active load, 820–821
 common-mode input signal, 817–819
 differential-mode input signal, 816
 emitter-degeneration resistors, 820
 Difference amplifier, 645–649
 Difference amplifier and bridge circuit design, 663–664
 Difference amplifier block diagram, 754
 Differential amplifier circuit, 1256–1258
 Differential amplifier with active load, 790–801
 BJT diff-amp with active load, 791–792
 MOSFET diff-amp with cascode active load, 798–800
 MOSFET differential amplifier with active load, 794–798
 small-signal analysis of BJT active load, 792–794
 Differential and multistage amplifiers, 753–850
 basic BJT differential pair. *See* Basic BJT differential pair
 basic FET differential pair. *See* Basic FET differential pair
 BiCMOS circuits, 801–805
 block diagram, 754
 design application (CMOS diff-amp), 821–824
 diff-amp frequency response, 815–821
 differential amplifier with active load. *See* Differential amplifier with active load
 gain stage/simple output stage, 806–811
 generally, 754
 simplified BJT op-amp circuit, 811–815
 Differential input resistance, 646
 Differential mode gain, 762
 Differential-mode gain, 818
 Differential-mode input resistance, 774–775
 Differential-mode input voltage, 754, 760
 Differential voltage gain, 623
 Differentiator, 653
 Diffusion, 17
 Diffusion capacitance, 46
 Diffusion conductance, 45
 Diffusion current density, 19–21
 Diffusion resistance, 45, 379
 Digital circuits, 3–4
 Digital electronics, 1141–1314
 bipolar digital circuits. *See* Bipolar digital circuits
 logic functions/logic gates, 1141–1142
 logic levels, 1143
 MOSFET digital circuits. *See* MOSFET digital circuits
 noise margin, 1143–1144
 preview, 1141
 propagation delay time, 1144
 switching time, 1144
 Digital logic, 325–327
 Digital logic gate, 167–168
 Digital signals, 3–4
 Digital-to-analog (D/A) converter, 1226, 1227–1229

Diode. *See also* Diode types
 back-to-back Zener, 1094
 ideal, 34, 35
 LED, 49
 piecewise linear, 37–41
 pn junction, 9, 29–34
 Schottky, 50
 speedup, 1295
 varactor, 26

Diode AND logic circuit, 105
 Diode biasing, 592–594
 Diode breakdown voltage, 52
 Diode charging time, 78, 79
 Diode circuit. *See* Diode circuit; Diode circuit - ac equivalent circuit; Diode circuit - dc analysis; Diode types
 clamp circuit, 94–97
 clipper circuit, 91–94
 design application (dc power supply), 108–110
 LED circuit, 107–108
 logic circuit, 104–105
 multiple-diode circuit, 97–105
 photodiode circuit, 106
 rectifier circuit. *See* Rectifier circuit
 Zener diode circuit, 84–90

Diode circuit - ac equivalent circuit
 circuit analysis, 45–46
 current-voltage relationships, 43–45
 frequency response, 46–47
 sinusoidal analysis, 43–47
 small-signal equivalent circuit, 47

Diode circuit - dc analysis, 34–43
 computer simulation/analysis, 41–42
 iteration/graphical analysis techniques, 36–37
 piecewise linear model, 37–41
 summary, 42–43

Diode clamp circuit, 94–97
 Diode clipper circuit, 91–94
 Diode conduction time, 78, 79
 Diode current, 78
 Diode logic circuit, 104–105
 Diode OR logic circuit, 104
 Diode rectifier, 35
 Diode thermometer, 54–56
 Diode thermometer with bipolar transistor, 348–350
 Diode thermometer with MOS transistor, 190–192

Diode types, 48–53. *See also* Diode
 LED, 49
 photodiode, 49
 pn junction diode, 29–34
 Schottky barrier diode, 50–51
 solar cell, 48–49
 Zener diode, 51–53

Direct bandgap material, 49
 Direct bandgap semiconductor, 107

Discrete circuit representation
 current amplifier, 882–886
 transconductance amplifier, 888–892
 transresistance amplifier, 895–901
 voltage amplifier, 873–879

Discrete-circuit series-shunt feedback topology, 874

Discrete electronic circuit, 3

Discrete-transistor shunt-series feedback circuit, 883

Discrete transistor shunt-shunt feedback circuit, 895

DMOS device, 566

DMOS process, 566

Dominant low-frequency pole, 968

Dominant pole, 500, 919

Domino circuit, 1193

Domino CMOS logic circuit, 1193

Donor impurity, 14

Doped semiconductor, 15

Doping, 14

Double-diffused process, 566

Drain current, 142

Drain terminal, 128
 Drain-to-source resistance, 567
 Drain-to-source voltage, 130, 131, 1297
 DRAM, 144, 1209
 DRAM cells, 1220–1221
 Drift, 17
 Drift current density, 17
 Driver transistor, 158
 Dual-slope A/D converter, 1231–1232
 Duty cycle, 1097
 Dynamic RAM (DRAM), 1209
 Dynamic RAM (DRAM) cells, 1220–1221
 Dynamic random access memory (DRAM), 144
 Dynamic shift register, 1203–1204

E

Early, J. M., 296
 Early effect, 296
 Early voltage, 296, 695
 Ebers-Moll model, 310
 ECL. *See* Emitter-coupled logic (ECL)
 ECL OR/NOR logic gate voltage transfer characteristics, 1266
 ECL series gating, 1274, 1275
 EEPROM, 1210, 1225
 Einstein relation, 21
 Electrically erasable ROM (EEPROM), 1210, 1225
 Electron accumulation layer, 128
 Electron diffusion coefficient, 20
 Electron-hole recombination, 22
 Electron inversion layer, 128
 Electron mobility, 17
 Electron-volt, 12n
 Electronic circuit, 2
 Electronic design, 615–618
 bottom-up design process, 616
 circuit performance estimation/simulation, 618
 design approach, 615–616
 flowchart, 617
 generally, 617–618
 prototype circuit, 618
 system design, 616–617
 top-down design process, 615–616
 Electronic noise, 859
 Electronic power supply, 68
 Electronic thermometer. *See* Thermometer
 Electronic thermometer with instrumentation amplifier, 665–668
 Electronics, 1
 Elemental semiconductor, 10
 Emitter, 286
 Emitter bypass capacitor, 407–410, 496
 Emitter-coupled logic (ECL), 753, 1256–1275
 alternative ECL gates, 1270–1273
 basic ECL logic gate, 1258–1259
 dc analysis, 1266
 diff-amp circuit, 1256–1258
 ECL logic gate with emitter followers, 1259–1260
 fanout, 1263–1264
 low-power ECL, 1267–1269
 modified ECL circuit configurations, 1267–1276
 negative supply voltage, 1265–1266
 noise margin, 1266–1267
 power dissipation, 1262–1263
 propagation delay time, 1263, 1275–1276
 reference circuit, 1260–1262
 series gating, 1273–1275
 voltage transfer characteristics, 1266–1267
 Emitter degeneration, 820
 Emitter-degeneration resistors, 820

Emitter-follower amplifier. *See* Common-collector amplifier

Emitter-follower circuit, 421, 533

Emitter resistor, 400

Energy band diagram, 12

Enhancement-load device, 157–165

Enhancement-load inverter, 1160

Enhancement mode, 128

Enhancement-mode MESFET, 183

Enhancement-mode MOSFET, 130

Enhancement-mode NMOS transistor, 1222

Enhancement-mode p-channel MESFET, 183

Enhancement-mode pn JFET, 183

EPROM, 1224

EPROM transistor, 1224–1125

Equilibrium pn junction, 23–25

Equivalent input gate capacitance, 519

Equivalent Miller capacitance, 816

Equivalent preamplifier circuit, 398

Equivalent Thevenin resistance, 333

Equivalent Thevenin voltage, 333

Equivalent two-port networks, 397

Erasable programmable ROM (EPROM), 1224

Error amplifier, 1116, 1120

Euler's identity, 4

Excess carrier lifetime, 22

Excess carriers, 21–22

Excess electron, 22

Excess hole, 22

Exclusive-OR function, 1166, 1167

Expanded hybrid- π equivalent circuit, 502–503

Exponential amplifier, 656

External offset compensation network, 1038–1039

Extrinsic semiconductor, 15

Extrinsic semiconductors, 14–16

F

Fanout

CMOS logic circuits, 1190

ECL, 1263–1264

NMOS logic circuits, 1167

TTL, 1285–1286

Feedback amplifier, 854

Feedback and stability, 851–946

bandwidth extension, 857–859

basic feedback concepts, 853–862

computer simulation, 853

current (shunt-series) amplifier, 879–886

design application (MOSFET feedback circuit), 924–926

frequency compensation, 918–924

gain sensitivity, 856–857

ideal closed-loop signal gain, 854–856

ideal feedback topologies, 863–873

loop gain, 901–907

noise sensitivity, 859–861

nonlinear distortion, 861–862

positive/negative feedback, 852–853

series-series configuration, 870–871

series-shunt configuration, 863–866

shunt-series configuration, 867–870

shunt-shunt configuration, 871–872

stability of feedback circuit. *See* Stability of feedback circuit

transconductance (series-series amplifier), 886–892

transresistance (shunt-shunt amplifier), 833–901

voltage (series-shunt) amplifier, 873–879

Feedback circuit, 1075

Feedback effect, 507

Feedback transfer function, 854

FET. *See* Field-effect transistor
 FET amplifier. *See* Basic FET amplifier
 FET current sources, 707–719
 basic two-transistor MOSFET current source, 707–711
 bias-independent current source, 715–716
 cascode current mirror, 711–713
 JFET current source, 716–718
 multi-MOSFET current-source circuit, 711–715
 wide-swing current mirror, 713–714
 Wilson current mirror, 713
 FET differential pair. *See* Basic FET differential pair
 Field effect, 126
 Field-effect transistor, 125–204
 CMOS, 140–141
 constant-current biasing, 170–175
 current source. *See* FET current sources
 design application (diode thermometer with MOS transistor), 190–192
 differential pair, 779–790
 frequency response, 514–520
 JFET, 180–190
 MESFET, 182–190
 MOSFET. *See* MOSFET
 MOSFET dc circuit analysis, 146–165
 NMOS, 128–134, 141
 PMOS, 134–137, 141
 Field oxide, 129
 Filter, 75–76
 active. *see* Active filter
 band-reject, 1063
 bandpass, 1063
 Butterworth. *See* Butterworth filter
 four-pole, 1069
 general two-pole active, 1064
 high-pass, 1062
 higher-order Butterworth, 1068–1070
 low-pass, 1063
 one-pole, 1063
 switched-capacitor, 1070–1074
 three-pole, 1069
 two-pole high-pass Butterworth, 1068
 two-pole low-pass Butterworth, 1065–1067
 Filter capacitor, 76
 Filter circuit, 75
 Finite differential-mode gain, 633
 Finite open-loop gain, 633, 1013–1023
 Finite output resistance, 142–143
 First-order functions, 473–474
 First-order transfer function, 473–474
 Five-input NOR logic address decoder, 1211
 555 circuit, 1100–1107
 astable multivibrator, 1105–1107
 basic operation, 1101–1103
 monostable multivibrator, 1103–1105
 other applications, 1107
 555 IC timer circuit, 1101
 555 monolithic integrated circuit timer, 1100
 Flash A/D converter, 1229–1230
 Flash memory, 1210
 Flip-flop
 D, 1206–1207
 defined, 1205
 J-K, 1251
 R-S, 1205–1206
 Floating-gate electrically erasable programmable ROM, 1225
 Foldback characteristic, 1121
 Folded cascode CMOS op-amp circuit, 976–979
 Folded-cascode stage, 976
 Forbidden bandgap, 12
 Forced beta, 309
 Forward-active operating mode, 287
 Forward bias, 27
 Forward-biased pn junction, 27, 28
 Forward diode resistance, 38

Forward transconductance, 760
 4-bit weighted-resistor D/A converter, 1227
 Four-pole high-pass Butterworth filter, 1070
 Four-pole low-pass Butterworth filter, 1070
 Four-terminal NMOS device, 214
 FPBW, 1029
 Frequency compensation, 918–924
 basic theory, 918–920
 closed-loop frequency response, 920–922
 Miller compensation, 922–924
 problem-solving technique, 920
 Frequency response, 469–558
 amplifier, 470–472
 bipolar transistor, 502–514
 Bode plot, 474–479
 bypass capacitor effects, 496–499
 cascode circuit, 528–532
 common-base circuit, 524–528
 common-emitter circuit, 521–524
 common-gate circuit, 524
 common-source circuit, 524
 coupling and bypass capacitors, 499–501
 coupling and load capacitors, 493–496
 coupling capacitor effects, 485–492
 cutoff frequency, 506–507
 emitter-follower circuit, 532–536
 expanded hybrid- π equivalent circuit, 502–503
 FET, 514–520
 first-order functions, 473–474
 high-frequency amplifier design, 536
 high-frequency equivalent circuit, 514–516
 input coupling capacitor, 485–488
 load capacitor effects, 492–493
 Miller capacitance, 510–512, 518–520
 Miller effect, 507–509, 512–513, 518–520
 multistage circuits, 485
 open-circuit time constant, 480–482
 output coupling capacitor (common-source circuit), 488–490
 output coupling capacitor (emitter-follower circuit), 490–492
 s-domain analysis, 472–473
 short-circuit current gain, 503–505
 short-circuit time constant, 480–482
 source-follower circuit, 532
 system transfer functions, 472–485
 time constant technique, 488
 time response, 483–485
 transistor amplifiers with circuit capacitors, 485–501
 transistor circuits, 520–537
 unity-gain bandwidth, 516–518
 Frequency-selective network, 1075
 Full-power-bandwidth (FPBW), 1029
 Full-wave bridge rectifier, 73
 Full-wave rectification, 71–74
 Full-wave rectifier, 72

G

g_m mismatch
 BJT differential pair, 770
 FET differential pair, 788
 GaAs MESFET, 182
 Gain-bandwidth product, 495, 1025–1026
 Gain margin, 917
 Gain sensitivity, 856–857
 Gallium arsenide, 10, 13
 Gallium phosphide, 10
 Gate, 126
 Gate potential, 1169
 Gate-to-channel space charge region, 181
 Gate-to-source voltage, 130, 225
 General CMOS design, 1188
 General constants and conversion factors, 1315

General op-amp circuit design, 948–950
 General two-pole active filter, 1064
 General two-pole low-pass filter, 1065
 Generalized CMOS clocked logic circuit, 1193
 Generalized current mirror, 705, 706
 Generalized inverting amplifier, 652
 Generalized op-amp summing amplifier, 659
 Generalized three-stage amplifier, 436
 Generalized two-stage amplifier, 175
 Germanium, 10, 13
 Graphical analysis technique, 36

H

h-parameters, 390–392
 Half-wave rectification, 68–71
 Half-wave rectifier, 69
 Half-wave rectifier circuit, 70
 Harmonic distortion, 375–376. *See also* Total harmonic distortion (THD)
 Hartley oscillator, 1083
 Heat sink, 567–571
 Hester, Richard, 384
 HEXFET structure, 566
 High-frequency amplifier design, 536
 High-frequency equivalent circuit, 514–516
 High-frequency radio-frequency (RF) circuit design, 572
 High-frequency range, 470, 471
 High-frequency response of transistor circuits, 520–537
 cascode circuit, 528–532
 common-base circuit, 524–528
 common-emitter circuit, 521–524
 common-gate circuit, 524
 common-source circuit, 524
 emitter-follower circuit, 532–536
 high-frequency amplifier design, 536
 source-follower circuit, 532
 High-pass filter, 1062
 High-pass network, 476
 Higher-order Butterworth filter, 1068–1070
 Historical overview, 1–2
 Hole, 13
 Hole diffusion coefficient, 20
 Hole inversion layer, 128
 Hole mobility, 18
 Hybrid FET op-amp
 LF155 series, 990–991
 LH002/42/52 series, 989–990
 Hybrid- π , 378
 Hybrid- π equivalent circuit, 303, 380
 Hybrid- π equivalent circuit (Early effect), 385–389
 Hysteresis effect, 1089

I

IC biasing. *See* Integrated circuit biasing and active loads
 IC npn bipolar transistor, 287
 IC op-amp, 622. *See also* Ideal op-amp
 IC power amplifiers. *See* Integrated circuit power amplifiers
 IC resistor, 693
 Ideal closed-loop signal gain, 854–856
 Ideal closed-loop voltage gain, 634
 Ideal current-voltage relationships, 28–29
 Ideal difference amplifier, 645
 Ideal diode, 34, 35
 Ideal feedback topologies, 863–873
 Ideal op-amp, 621–686
 antilog amplifier, 656

applications, 641–657
 circuit design, 658–665
 circuits. *See* Operational amplifier circuits
 comparator, contrasted, 1085
 current-to-voltage converter, 641–642
 defined, 621
 derivation of name, 621
 design application (electronic thermometer with instrumentation amplifier), 665–668
 difference amplifier, 645–649
 difference amplifier and bridge circuit design, 663–664
 differentiator, 653
 evolution, 622
 exponential amplifier, 656
 generally, 622
 ideal parameters, 622–625
 instrumentation amplifier, 650–651
 integrator, 652–653
 inverting amplifier, 627–635
 log amplifier, 655
 negative feedback analysis, 626
 non-linear circuit applications, 654–657
 noninverting amplifier, 638–639
 OTA, 657–658
 output currents, 627
 output voltage swing, 627
 practical specifications, 626–627
 precision half-wave rectifier, 654–655
 PSpice modeling, 627
 reference voltage source design, 661–663
 summing amplifier, 636–638
 summing op-amp circuit design, 659–661
 T-network, 631–633
 voltage follower, 639–640
 voltage-to-current converter, 642–645
 Ideal series-series feedback topology, 870
 Ideal series-shunt feedback topology, 864
 Ideal short-circuit output current, 519
 Ideal shunt-series feedback topology, 867
 Ideal shunt-shunt feedback topology, 871
 Ideal voltage reference circuit, 85–88
 Idealized class-B operation, 576–577
 Idealized power efficiency, 579–581
 Identical transistor characteristics, 687
 Impedance, 217
 capacitor, 217, 399
 input. *See* Input impedance
 output. *See* Output impedance
 Impedance buffer, 640
 Impedance transformer, 425, 640
 Incremental current gain, 380
 Indirect bandgap material, 49
 Indium phosphide, 10
 Induced positive space-charge region, 128
 Inductively coupled amplifier, 587–588
 Input base-emitter port, 378–379
 Input bias current, 1042–1044
 Input buffer-inverter pair, 1211
 Input buffer transistors, 596–599
 Input coupling capacitor, 485–488
 Input dc offset voltage, 1030
 Input impedance
 common-base amplifier, 433–434
 common-drain amplifier, 231–233
 common-gate amplifier, 236
 emitter-follower amplifier, 423–424
 Input load line, 306–307
 Input offset current, 1042
 Input offset current temperature coefficient, 1046
 Input offset voltage, 1031
 Input offset voltage drift, 1045
 Input resistance, 629, 774, 808
 Input signal-to-noise ratio, 859
 Input stage offset voltage effects, 1030–1037
 Input transistor, 1280–1282

Instantaneous gate-to-source voltage, 258
 Instantaneous power dissipation, 562, 579
 Instantaneous value, 4
 Instrumentation amplifier, 650–651, 665–668
 Insulator, 12
 Integrated circuit biasing, 342–343
 Integrated circuit biasing and active loads, 687–751
 bipolar transistor current sources. *See* Bipolar transistor current sources
 circuits with active loads. *See* Circuits with active loads
 design application (NMOS current source), 734–736
 FET current sources. *See* FET current sources
 small-signal analysis - active load circuits, 726–734
 Integrated circuit npn bipolar transistor, 287
 Integrated circuit operational amplifier, 622. *See also* Ideal op-amp
 Integrated circuit power amplifiers, 1107–1113
 bridge power amplifier, 1113
 LM380 power amplifier, 1108–1111
 PA12 power amplifier, 1111–1113
 Integrated circuit resistor, 693
 Integrator, 652–653
 Interdigitated bipolar transistor, 565
 Interdigitated structure, 565
 Intrinsic carrier concentration, 13
 Intrinsic semiconductors, 10–14
 Inverse-active mode, 310
 Inverse Laplace transform, 484
 Inverse resistance reflection rule, 424
 Inverter
 BiCMOS, 1296–1297
 bipolar circuit, 323
 CMOS, 1168–1182
 NMOS, 1165–1166, 1148–1163
 Inverter gain, 1155
 Inverting amplifier, 627–635
 Inverting amplifier closed-loop gain, 1014–1016
 Inverting amplifier closed-loop input resistance, 1019–1021
 Inverting comparator circuit, 1086
 Inverting input terminal, 623
 Inverting Schmitt trigger, 1087–1090
 Inverting summing amplifier, 637
 Iteration, 36

J

J-K flip-flop, 1251
 JFET, 180–190
 JFET amplifier, 258–263
 JFET cascode current source, 717
 JFET current source, 716–718
 JFET differential amplifier, 788–789
 JFET logic circuit, 1147
 JFET operational amplifier circuit, 989–992
 JFET source-follower circuit, 262
 Junction area, 106
 Junction capacitance, 26
 Junction field-effect transistor (JFET), 180–190

K

Kilby, Jack, 1
 Kirchhoff's current law
 expanded hybrid- π equivalent circuit, 391
 Miller effect/capacitance, 518
 npn transistor, 291
 system transfer functions, 473
 Kirchhoff's voltage law
 clamper, 94
 common-emitter circuit, 305

expanded hybrid- π equivalent circuit, 391
 iteration/graphical analysis technique, 36
 load line/modes of operation, 307
 piecewise linear model, 41
 voltage doubler circuit, 83

L

Large signal analysis, 371
 Large-value resistors, 1070
 Leakage currents, 298–299
 Least significant bit (LSB), 1226
 LED, 49
 LED circuit, 107–108
 LF155 BiFET op-amp, 990–991
 LF155 BiFET op-amp input stages, 991
 LH002/42/52 series hybrid JFET op-amp, 990
 LH002/42/52 series op-amp, 989–990
 LH0042C data, 990
 Light-emitting diode (LED), 49
 Light-emitting diode (LED) circuit, 107
 Limiter circuit, 91
 Line regulation, 1114
 Linear amplifier, 3, 205, 206. *See also* Basic FET amplifier; Bipolar linear amplifier, 370
 Linear ramp generator, 1107
 Linear term, 374
 LM741 op-amp, 995. *See also* Bipolar operational amplifier circuit, 1325–1327
 LM380 power amplifier, 1108–1111
 LM380 power amplifier characteristics, 1110
 LM555 timer circuit, 1101
 LM78LXX series three-terminal positive voltage regulator, 1119
 Load capacitor effects, 492–493
 Load curve, 240
 Load line, 37, 153–155, 306–310
 Load regulation, 89, 1115
 Loading effect, 219, 403
 Log amplifier, 655
 Logic 1 level, 1259
 Logic circuits, 104–105
 BiCMOS digital circuits, 1296–1298
 bipolar. *See* Bipolar digital circuits
 CMOS. *See* CMOS logic circuits
 emitter-coupled logic (ECL), 1256–1267
 JFET, 1147
 NMOS, 1163–1168
 NMOS transistor digital, 167–168
 sequential, 1202–1208
 Logic function, 104–105, 1208
 Logic functions/logic gates, 1141–1142
 Logic levels, 1143
 Long-channel device, 141
 Loop gain, 855, 901–907
 Low-frequency dominant pole, 988
 Low-frequency range, 470, 471
 Low-pass filter, 1063
 Low-pass network, 478
 Low-power ECL, 1267–1269
 Low-power Schottky TTL circuits, 1293–1295
 Low-power Schottky TTL NAND logic circuit, 1293
 Lower corner, 481
 Lower corner frequency, 501
 Lower cutoff frequency, 501
 Lowercase letter/lowercase subscript, 4
 Lowercase letter/uppercase subscript, 4
 LSB, 1226

M

Majority carrier, 16
 Majority-carrier device, 180

- Manufacturers' data sheets. *See* Selected manufacturers' data sheets
- Mask-programmable ROM, 1209, 1222
- Matched or identical transistor characteristics, 687
- Matched transistor, 950
- Maximally flat magnitude filter, 1065
- Maximum attainable conversion efficiency, 574
- Maximum average output signal power, 590
- Maximum gain asymptote, 493
- Maximum output symmetrical swing, 415
- Maximum possible average power, 581
- Maximum power conversion efficiency, 590
- Maximum power transfer theorem, 574
- Maximum rated collector current, 561
- Maximum rated current/maximum rated voltage, 560n
- Maximum rated power, 563
- Maximum safe power dissipation, 569, 570
- Maximum symmetrical swing, 415–420
- MC14573 CMOS op-amp circuit
- circuit description, 970
 - dc analysis, 970–972
 - small-signal analysis, 972–973
- MC14573 CMOS op-amp equivalent circuit, 971
- Memory. *See* Computer memory
- Memory architecture, 1210
- Memory cell, 1209
- MESFET, 182–190
- Metal-oxide-semiconductor field-effect transistor. *See* MOSFET
- Metal-semiconductor field-effect transistor (MESFET), 182–190
- Metallurgical junction, 23
- Microelectronics, 1
- Microphone, 663
- Microphone system, 645
- Midband range, 470, 471
- Midband small-signal voltage gain, 255
- Midband voltage gain, 527
- Miller capacitance
- cascode circuit, 529
 - common-emitter/common-source circuits, 521
 - expanded hybrid-p equivalent circuit, 510–512
 - FET, 518–520
- Miller compensation, 922–924, 968
- Miller effect
- expanded hybrid-p equivalent circuit, 507–509
 - FET, 518–520
 - importance, 523
 - physical effect, 512–513
- Miller multiplication factor, 821, 922
- Minimum β , 1279
- Minority carrier, 16
- Minority carrier electron concentration, 288
- 3 dB frequency, 475
- Model, 34
- Modified Darlington pair configuration, 801
- Modified ECL circuit configurations, 1267–1276
- Modified ECL logic gate, 1268
- Modified MOSFET Wilson current source, 713
- Modified totem-pole output stage, 1287–1288
- Modified Wilson current source, 713
- Monostable multivibrator, 1098–1100, 1103–1105
- MOS capacitor (n-type substrate), 128
- MOS capacitor (p-type substrate), 127
- MOS capacitor structure, 127
- MOS field-effect transistor. *See* MOSFET
- MOSFET, 126–146
- body effect, 143–144
 - breakdown effect, 144–145
 - circuit symbols/conventions, 137–138
 - CMOS, 140–141
 - common-source circuit, 146–153
 - dc circuit analysis, 146–165
 - digital circuits. *See* MOSFET digital circuits
 - digital logic gate, 167–168
 - finite output resistance, 142–143
 - ideal current-voltage characteristics - PMOS, 135–137
 - load line/modes of operation, 153–155
 - multistage circuit, 175–179
 - n-channel depletion-load device, 161–163
 - n-channel enhancement-mode, 138–140
 - n-channel enhancement-load device, 157–161
 - NMOS, 128–134, 141
 - NMOS inverter, 165–166
 - nonideal current-voltage characteristics, 142–145
 - p-channel depletion-mode, 140
 - p-channel enhancement-load device, 163–164
 - PMOS, 134–137, 141
 - power, 565–567
 - short-channel effects, 141–142
 - small-signal amplifier, 168–169
 - subthreshold conduction, 144
 - summary of transistor operation, 141
 - switching, 165–166
 - temperature effect, 145
 - two-terminal MOS structure, 126–128
- MOSFET amplifier
- graphical analysis, 206–209
 - load line, 207
 - modeling the body effect, 214–215
 - small-signal equivalent circuit, 210–214
 - small-signal parameters, 206–210
 - transistor parameters, 207–209
- MOSFET cascode current mirror, 711
- MOSFET cascode current source, 786
- MOSFET cascode diff-amp with cascode active load, 800
- MOSFET class-AB output stage, 584–585
- MOSFET Colpitts oscillator, 1082
- MOSFET common-source circuit, 492
- MOSFET current source, 709
- MOSFET current-source circuit, 711
- MOSFET dc analysis, 146–165
- MOSFET diff-amp with cascode active load, 798–800
- MOSFET differential amplifier with active load, 794–798
- MOSFET digital circuits, 1147–1254
- clocked CMOS logic circuits, 1191–1193
 - CMOS inverter. *See* CMOS inverter
 - CMOS logic circuits. *See* CMOS logic circuits
 - data converter. *See* Data converter; Sequential logic circuits
 - design application (static CMOS logic gate), 1232–1234
 - memory. *See* Computer memory
 - NMOS inverter. *See* NMOS inverter
 - NMOS logic circuits, 1163–1168
 - sequential logic circuits. *See* Sequential logic circuits
 - transmission gates. *See* Transmission gates
- MOSFET feedback circuit, 924–926
- MOSFET high-frequency circuit, 519
- MOSFET integrated circuit amplifier, 155
- MOSFET small-signal amplifier, 168–169
- MOSFET Wilson current source, 713
- Most significant bit (MSB), 1226
- MSB, 1226
- Multi-MOSFET current-source circuit, 711–715
- Multiooutput transistor current source, 704
- Multiple-diode circuit, 97–105
- Multistage amplifiers, 254–258. *See also*
- Differential and multistage amplifiers
 - cascade circuit, 254–256
 - cascode circuit, 256–257
- Multistage BJT amplifier
- cascade configuration, 436–438
- cascode configuration, 440–442
- Darlington pair configuration, 439–440
- Multistage BJT analysis
- cascade configuration, 436–438
 - cascode configuration, 440–442
- Multistage MOSFET circuit, 175–179
- Multistage transistor circuit, 344–348
- Multitransistor circuit
- cascade configuration, 176–177
 - cascode configuration, 178–179
- Multitransistor current mirror, 704–706
- Multivibrator
- astable, 1096, 1105–1107
 - bistable, 1087
 - defined/types, 1084
 - monostable, 1098–1100, 1103–1105
 - one-shot, 1098, 1103
 - Schmitt trigger monostable, 1098
- μ A-709, 622
- μ A-741, 622
- Music, 3

N

- N-bit D/A converter, 1229
- n-channel depletion-load device, 161–163
- n-channel depletion-mode MOSFET, 138–140
- n-channel enhancement-load device, 157–161
- n-channel enhancement-mode MESFET, 183
- n-channel enhancement-mode MESFET circuit, 189
- n-channel enhancement-mode MOSFET, 128–134, 624, 1153
- n-channel JFET, 181, 184
- n-channel JFET circuit, 186
- n-channel MESFET, 182
- n-channel MOSFET, 128–134, 141, 1148–1150
- n-channel MOSFET detailed cross section, 1148
- n-channel MOSFET simplified view, 1148
- n-type semiconductor, 14, 17
- NDS9410 n-channel enhancement-mode MOSFET, 1322–1324
- Near-avalanche breakdown, 145
- Negative and positive biasing, 339–341
- Negative feedback, 626, 852–853
- Negative feedback analysis, 626
- NMOS, 128–134, 141
- NMOS amplifier with active loads, 247–254
- NMOS cascade circuit, 255
- NMOS cascode circuit, 256
- NMOS common-source circuit, 225
- NMOS current source, 734–736
- NMOS depletion-load amplifier, 244–247
- NMOS enhancement-load amplifier, 239–244
- NMOS exclusive-OR logic gate, 1167
- NMOS input buffer-inverter, 1211
- NMOS inverter, 165–166, 1148–1163
- body effect, 1160–1162
 - depletion load, 1157–1160
 - n-channel MOSFET, 1148–1150
 - resistor load, 1150–1157
- NMOS inverter circuit, 165
- NMOS inverter with depletion load, 1157–1160
- NMOS inverter with resistor load, 1150–1157
- NMOS logic circuits, 1163–1168
- NMOS NAND logic gate, 1165
- NMOS NOR logic gate, 167, 1163–1165
- NMOS pass networks, 1198–1199
- NMOS R-S flip-flop, 1205
- NMOS shift register voltage, 1204
- NMOS 16 x 1 mask-programmable ROM, 1222
- NMOS source follower, 227
- NMOS source-follower circuit, 250
- NMOS SRAM cells, 1212–1214
- NMOS transistor digital logic circuit, 167–168

- NMOS transmission gate, 1194–1198
 Noise, 859
 Noise margin, 1143–1144, 1178–1182, 1266–1267
 Noise sensitivity, 859–861
 Nondominant-pole frequency, 982
 Nonideal current-voltage characteristics, 142–145
 Nonideal effects in op-amp circuits, 1009–1060
 basic bipolar diff-amp stage, 1030–1032
 bias current compensation, 1043–1044
 bias current effects, 1042–1043
 bipolar active load diff-amp stage, 1032–1035
 CMRR, 1046–1047
 design application (offset voltage compensation network), 1047–1049
 external offset compensation network, 1038–1039
 finite open-loop gain, 1013–1023
 frequency response, 1023–1030
 gain-bandwidth product, 1025–1026
 input bias current, 1042–1044
 input offset current temperature coefficient, 1046
 input stage offset voltage effects, 1030–1037
 input voltage limitations, 1011–1013
 inverting amplifier closed-loop gain, 1014–1016
 inverting amplifier closed-loop input resistance, 1019–1021
 MOSFET diff-amp stage, 1035–1037
 noninverting amplifier closed-loop gain, 1016–1017
 nonzero output resistance, 1021–1023
 offset-null terminals, 1039–1041
 offset voltage, 1030–1041
 offset voltage compensation, 1037–1041
 offset voltage temperature coefficient, 1045–1046
 open-loop/closed-loop frequency response, 1023–1025
 output voltage limitations, 1013
 practical op-amp parameters, 1010–1011
 slew rate, 1026–1029
 temperature effects, 1045–1046
 Noninverting amplifier, 638–639
 Noninverting amplifier closed-loop gain, 1016–1017
 Noninverting comparator circuit, 1086
 Noninverting input terminal, 623
 Noninverting Schmitt trigger circuit, 1090–1091
 Nonlinear distortion, 861–862
 Nonsaturation region, 132, 1149, 1169
 Nonsinusoidal oscillators/timing circuits, 1096–1107
 555 circuit. *See* 555 circuit
 monostable multivibrator, 1098–1100
 Schmitt trigger oscillator, 1096–1098
 Nonzero output resistance, 1021–1023
 NOR decoder, 1211
 NOR logic function, 326
 Norton equivalent circuit, 396, 509
 Norton equivalent signal source, 235
 Notation, 4–5
 Noyce, Robert, 1
 npn bipolar transistor, 502
 base current, 289–290
 circuit symbol, 294
 collector current, 289
 common-emitter configuration, 291
 common-emitter current gain, 290–291
 cross section, 287
 current relationships, 291
 current-voltage relationships, 294
 emitter current, 288–289
 generally, 286
 minority carrier concentration, 288
 notation, 290n
 other current components, 289n
 simple block diagram, 294
 simple geometry, 286
 summary of transistor operation, 291–292
 transistor currents, 287–288
 npn common-emitter circuit, 400
 npn general purpose amplifier, 1318–1319
 Null technique, 952
 Nyquist diagram, 913
 Nyquist plot, 913, 914
 Nyquist stability criterion, 913–916
- O**
- Octave, 475
 Offset compensation network, 1038
 Offset-null terminal, 1038
 Offset-null terminals, 1039–1041
 Offset voltage, 1030–1041
 Offset voltage compensation, 1037–1041
 Offset voltage compensation network, 1047–1049
 Offset voltage temperature coefficient, 1045–1046
 Ohmic contact, 51
 Ohm's law, 34
 On resistance, 566
 One-bit full-adder circuit, 1207, 1208
 One-pole filter, 1063
 One-pole low-pass filter, 1072
 One-shot multivibrator, 1098, 1103
 One-sided output, 762, 763–765
 One-transistor dynamic RAM cell, 1220
 Op-amp. *See* Ideal op-amp
 Op-amp applications, 641–657
 Op-amp circuit design, 658–665
 Op-amp circuits. *See* Operational amplifier circuits
 Op-amp difference amplifier, 645
 Op-amp differentiation, 653
 Op-amp integrator, 652
 Op-amp series-shunt feedback circuit, 873
 Op-amp shunt-series feedback configuration, 881
 Op-amp shunt-shunt feedback configuration, 884
 Op-amp voltage reference circuit, 661
 Op-amp with input bias currents, 1042
 Open-circuit time constant, 480
 Open-loop comparator circuit, 1087
 Open-loop gain, 625
 Open-loop transresistance gain factor, 895
 Operational amplifier. *See* Ideal op-amp
 Operational amplifier circuits, 947–1008
 BiCMOS operational amplifier circuit, 981–989
 CA3140 BiCMOS circuit description, 983–984
 circuit element matching, 950
 CMOS op-amp circuits. *See* CMOS operational amplifier circuits
 folded cascode CMOS op-amp circuit, 976–979
 general design philosophy, 948–949
 general op-amp circuit design, 948–950
 hybrid FET op-amp, LF155 series, 990–991
 hybrid FET op amp, LH002/42/52 series, 989–990
 JFET operational amplifier circuit, 989–992
 MC14573 CMOS circuit, 970–973
 problem-solving technique, 969
 741 op-amp. *See* Bipolar operational amplifier circuit
 three-stage CMOS op-amp, 973–976
 Operational transconductance amplifier (OTA), 657–658
 Optical transmission system, 49
 Optoisolator, 108
 OR logic function, 104
 Oscillator, 1074–1084
 basic principles, 1075–1076
 Colpitts, 1082–1083
 crystal, 1083–1084
 defined, 1075
 Hartley, 1083
 nonsinusoidal, 1096–1107
 phase-shift, 1076–1078
 Pierce, 1084
 Schmitt trigger, 1096–1098
 Wien-bridge, 1079–1082
 OTA, 657–658
 Output coupling capacitor
 common-source circuit, 488–490
 emitter-follower circuit, 490–492
 Output dc offset voltage, 1030
 Output impedance
 common-base amplifier, 433–434
 common-drain amplifier, 231–233
 common-gate amplifier, 236
 emitter-follower amplifier, 423–424
 Output resistance, 185, 297
 Output signal-to-noise ratio, 859
 Output stage linearity characteristic, 559
 Output stage using MOSFETs, 601–603
 Output stages and power amplifiers, 559–614
 class-A operation, 572–576
 class-A power amplifier, 586–591
 class-AB operation, 581–585
 class-AB push-pull complementary output stages, 591–601
 class-B operation, 576–581
 class-C operation, 585–586
 design application (output stage using MOSFETs), 601–603
 power transistors. *See* Power transistors
 scope of text, 572
 THD, 560
 Output voltage swing, 627
 Overlap capacitance, 515
- P**
- p-channel depletion-mode MOSFET, 140
 p-channel enhancement-load device, 163–164
 p-channel enhancement-mode MOSFET (PMOS), 134–137, 141
 p-channel JFET, 180, 184
 p-channel JFET circuit, 187
 p-type semiconductor, 15, 17
 PA12 power amplifier, 1111–1113
 Parallel A/D converter, 1229–1230
 Parallel-based clipper, 92
 Parallel-based diode clipper circuit, 92
 Parallel-plate capacitor, 127
 Parasitic capacitance, 515
 Pass transistor logic, 1198
 Passive device, 2
 Passive limiter, 91
 Peak inverse voltage (PIV), 32
 Phase margin, 916
 Phase quantity, 4
 Phase shift, 383
 Phase-shift oscillator, 1076
 Phase-shift oscillator circuit, 1078
 Phasor quantity, 208, 372
 Photocurrent, 48, 106
 Photodetector, 49
 Photodiode, 49
 Photodiode circuit, 106
 Photon flux density, 106
 Physical constants and conversion factors, 1315
 Piecewise linear diode, 37–41
 Pierce oscillator, 1084
 Piezoelectric crystal, 1083
 Piezoelectric crystal circuit symbol, 1083
 Piezoelectric crystal equivalent circuit, 1083
 Pinchoff, 181
 Pinchoff voltage, 183
 PIV, 32

- PMOS, 134–137, 141
 PMOS device, 1171–1174, 1177, 1188
 PMOS transistor, 1176
 pn JFET, 180–182
 pn junction, 23–34
 - diode, 29–34
 - equilibrium, 23–25
 - forward-biased, 27, 28
 - ideal current-voltage relationships, 28–29
 - reverse-biased, 25–27
 pn junction diode, 9, 29–34
 - breakdown voltage, 30–32
 - switching transient, 32–33
 - temperature effects, 30
 pn junction FET, 180
 pn junction solar cell, 48
 pnp bipolar transistor
 - base current, 293
 - circuit symbol, 294
 - collector current, 293
 - electron and hole currents, 293
 - emitter current, 292
 - forward active mode operation, 292–293
 - simple block diagram, 294
 - simple geometry, 286
 pnp general purpose amplifier, 1320–1321
 Pole-splitting, 924
 Polycrystalline silicon layer, 126
 Polysilicon fuse, 1223
 Positive and negative biasing, 339–341
 Positive feedback, 852, 1087
 Positive logic system, 326
 Positive voltage regulator
 - basic circuit description, 1119–1121
 - protection devices, 1121–1122
 - temperature compensation, 1121
 - three-terminal regulator, 1122
 Power BJT, 561–565, 567
 Power conversion efficiency, 573
 Power derating curve, 569, 570
 Power dissipation
 - BJT, 562
 - class-B operation, 579, 580
 - CMOS inverter, 1176–1178
 - ECL, 1262–1263
 - emitter-coupled logic (ECL), 1262–1263
 - maximum safe, 569, 570
 - RAM cells, 1213
 - transistors, 559
 Power MOSFET, 565–567
 Power transistors, 560–571
 - heat sink, 567–571
 - power BJT, 561–565, 567
 - power MOSFET, 565–567
 Precharge cycle, 1193
 Precision half-wave rectifier, 654–655
 Precision half-wave rectifier circuit, 654
 Precision resistors (1% tolerance), 1329–1330
 Pressure transducer, 663
 Principle of superposition, 371, 636, 1085
 Problem-solving technique
 - bipolar ac analysis, 384
 - bipolar dc analysis, 309
 - BJT current source circuits, 698
 - Bode plot of gain magnitude, 492
 - diff-amps with resistive loads, 774
 - frequency compensation, 920
 - maximum symmetrical swing, 416–417
 - MOSFET ac analysis, 212–214
 - MOSFET current-source circuit, 711
 - multistage circuits, 815
 - operational amplifier circuits, 969
 Problems, answers to, 1337–1358
 Process conduction parameter, 134
 Programmable ROM (PROM), 1209, 1222–1223
 PROM, 1209, 1222–1223
- Propagation delay time, 1144
 - CMOS logic circuit, 1190, 1191
 - CMOS logic circuits, 1191
 - ECL, 1263
 - emitter-coupled logic (ECL), 1263, 1275–1276
 - modified ECL circuit configurations, 1275–1276
 Prototype circuit, 618
 PSpice, 41
 - Pspice circuit schematic, 100
 - Pull-down resistor, 1279
 - Pulse position modulator, 1107
 - Pulse width modulator (PWM), 1107
 - Punch-through, 145
 - PWM, 1107
- Q**
- Q-point, 37, 307, 1158
 Q-point stability, 224
 Quasistatic quantities, 5
 Quiescent collector current, 954
 Quiescent drain current, 261
 Quiescent drain-to-source voltage, 225
 Quiescent gate-to-source voltage, 236
 Quiescent output voltage, 1109
 Quiescent point, 37, 307
 Quiescent source-to-gate voltage, 230
- R**
- R_c mismatch, 769–770
 R_D mismatch, 787–788
R-2R ladder network D/A converter, 1228–1229
 R-S flip-flop, 1205–1206
 Radian frequency, 470n
 RAM, 1209, 1212–1221
 RAM architecture, 1210
 RAM memory cells, 1212–1221
 Random access memory (RAM), 1209, 1212–1221
 Rated power, 570
 RC filter, 75
 Read-only memory (ROM), 1209–1210, 1221–1225
 Reading list, 1333–1336
 - analog integrated circuits, 1334
 - digital circuits and devices, 1335
 - general electronics texts, 1333
 - linear circuit theory, 1333–1334
 - op-amp circuits, 1334–1335
 - semiconductor devices, 1334
 - SPICE and PSpice references, 1336
 Recombination current, 289
 Recovery time, 1099
 Rectification, 68
 Rectifier, 34, 35
 Rectifier circuit
 - bridge rectifier, 72
 - detector, 82–83
 - diode current, 78
 - filter, 75–76
 - full-wave rectification, 71–74
 - half-wave rectification, 68–71
 - ripple voltage, 76, 77
 - voltage doubler circuit, 83–84
 Rectify, 35
 Reference books. *See* Reading list
 Reference circuit, 1260–1262
 Reference current
 - basic three-transistor current source, 694
 - bipolar current-source circuits, 709–710
 - current mirror circuits, 715
 - two-transistor current source, 688, 691
 Reference voltage source design, 661–663
 Region
 - active, 287
 channel, 128
 depletion, 24
 gate-to-channel space charge, 181
 induced positive space-charge, 128
 nonsaturation, 132, 1149, 1169
 saturation, 131, 1149, 1169, 1172
 space-charge, 24
 triode, 132
- Regulation
 - line, 1114
 - load, 89, 1115
 - source, 88, 89
 - voltage regulator, 84
 Regulator
 - basic series-pass voltage, 1116
 - voltage, 84
 - Zener diode voltage regulator circuit, 85, 88

Resistance

 - on, 566
 - common-mode input, 775
 - differential input, 646
 - differential-mode input, 774–775
 - diffusion, 379
 - finite output, 142–143
 - input, 629, 774
 - nonzero output, 1021–1023
 - output, 185, 297
 - reverse-biased diffusion, 390
 - series, 390
 - small-signal input, 391–392
 - small-signal transistor output, 385
 - thermal, 567
 - Thevenin, 399
 - Zener, 84
 Resistance reflection rule, 401, 423
 Resistive voltage divider, 1116
 Resistivity, 18
 Resistor, 687
 Resistor biasing technique, 687
 Resistors. *See* Standard resistor and capacitor values
 Restoring logic family, 1155
 Reverse bias, 25
 Reverse-bias saturation current, 28, 30
 Reverse-biased diffusion resistance, 390
 Reverse-biased drain-to-substrate pn junction, 130
 Reverse-biased pn junction, 25–27
 Reverse-saturation current, 666
 Ripple voltage, 76–77, 77
 Roll-off, 1063
 ROM, 1209–1210, 1221–1225
 Row and column decoder, 1211
 RS flip-flop, 1101

S

s-domain analysis, 472–473
 Safe operating area (SOA), 563
 Safe power dissipation, 569, 570
 Sampled-data system, 1074
 Saturation mode, 307
 Saturation region, 131, 1149, 1169, 1172
 Schmitt trigger, 1084
 Schmitt trigger application, 1092–1094
 Schmitt trigger circuit, 1084–1095

 - applied reference voltages, 1091–1092
 - basic inverting Schmitt trigger, 1087–1090
 - comparator, 1084–1087
 - limiters, 1094–1095
 - noninverting Schmitt trigger circuit, 1090–1091
 - Schmitt trigger application, 1092–1094
 - Schmitt trigger monostable multivibrator, 1098
 - Schmitt trigger oscillator, 1096–1098
 - Schottky barrier diode, 50–51
 - Schottky clamped transistor, 1289–1291
 - Schottky clamped transistor equivalent circuit, 1290

- Schottky clamped transistor symbol, 1290
 Schottky diode, 50
 Schottky transistor-transistor logic, 1289–1296
 advanced Schottky TTL circuits, 1295–1296
 low-power Schottky TTL circuits, 1293–1295
 Schottky clamped transistor, 1289–1291
 Schottky TTL NAND circuit, 1291–1293
 Schottky TTL NAND circuit, 1291–1293
 Second breakdown, 562
 Selected manufacturers' data sheets, 1317–1327
 LM741 op-amp, 1325–1327
 NDS9410 n-channel enhancement-mode MOSFET, 1322–1324
 npn general purpose amplifier, 1318–1319
 pnp general purpose amplifier, 1320–1321
 single n-channel enhancement-mode FET, 1322–1324
 2N2222 npn bipolar transistor, 1318–1319
 2N2907 pnp bipolar transistor, 1320–1321
 Semiconductor constants, 1315
 Semiconductor device, 5
 Semiconductor materials and diodes, 9–66
 design application (diode thermometer), 54–56
 diode circuits. *See Diode circuits*; Diode circuits - ac equivalent circuit; Diode circuits - dc analysis; Diode types
 drift and diffusion currents, 17–21
 excess carriers, 21–22
 extrinsic semiconductors, 14–16
 intrinsic semiconductors, 10–14
 pn junction. *See pn junction*
 semiconductor materials and properties, 10–22
 Semiconductor materials and properties, 10–22
 Semiconductor memory. *See Computer memory*
 Sequential logic circuits, 1202–1208
 CMOS full-adder circuit, 1207–1208
 D flip-flop, 1206–1207
 dynamic shift register, 1203–1204
 R-S flip-flop, 1205–1206
 Series gating, 1273–1275
 Series-pass regulator, 1116–1118
 Series-pass transistor, 1116
 Series resistance, 390
 Series-series amplifier, 886–892
 Series-series configuration, 870–871
 Series-shunt amplifier, 873–879
 Series-shunt configuration, 863–866
 741 op-amp. *See Bipolar operational amplifier circuit*
 Seven-segment LED display, 107
 Shift register, 1203
 Shockley, William, 1
 Short-channel device, 141
 Short-channel effects, 141–142
 Short-circuit current gain, 503–505
 Short-circuit protection circuitry, 960
 Short-circuit protection device, 960
 Short-circuit time constant, 480
 Shunt-series amplifier, 879–886
 Shunt-series configuration, 867–870
 Shunt-shunt amplifier, 833–901
 Shunt-shunt configuration, 871–872
 Signal, 206, 370
 Signal frequency. *See Frequency response*
 Signal processing, 67
 Signal-to-noise ratio, 859
 Silicon, 10, 13
 Silicon dioxide, 1224
 Simple all-bipolar op-amp circuit, 1012
 Simple antilog (exponential) amplifier, 656
 Simple bipolar operational amplifier, 812, 949
 Simple BJT amplifier with active load, 720
 Simple emitter-follower output stage, 806
 Simple filter circuit, 75
 Simple high-pass filter, 1062
 Simple log amplifier, 655
 Simple NMOS transistor digital logic circuit, 167–168
 Simple op-amp voltage reference circuit, 661
 Simple series-pass regulator, 1116–1118
 Simple voltage-to-current converter, 642
 Simple Zener diode voltage reference circuit, 84–90
 Simplified BJT op-amp circuit, 811–815
 Simulation models, 618
 Simulation program with integrated circuit emphasis (SPICE), 41
 Single base resistor biasing, 331–333
 Single-diode clipper circuit, 91
 Single n-channel enhancement-mode FET, 1322–1324
 Single-stage common-emitter amplifier, 909
 Single-stage integrated circuit MOSFET amplifier, 238–254
 CMOS common-gate amplifier, 251–254
 CMOS common-source amplifier, 247–249
 CMOS source-follower amplifier, 249–251
 load line, 239
 NMOS amplifier with active loads, 247–254
 NMOS depletion-load amplifier, 244–247
 NMOS enhancement-load amplifier, 239–244
 Sinusoidal base current, 375, 382
 Sinusoidal base-emitter voltage, 375
 Sinusoidal collector current, 382
 Sinusoidal input signal, 582
 Sinusoidal signal source, 373
 Sinusoidal source, 206
 Slew rate, 1026–1029
 Small geometry effects
 n-channel MOSFET, 1150
 p-channel MOSFET, 1170
 Small reverse-biased gate voltage, 181
 Small signal, 206, 374, 375
 Small-signal, 370
 Small-signal amplifier, 168–169
 Small-signal analysis, 371
 Small-signal analysis - active load circuits, 726–734
 advanced MOSFET active load, 732–733
 BJT active load circuit, 727–730
 MOSFET active load circuit, 730–732
 Small-signal BJT, 561
 Small-signal collector current, 379
 Small-signal current gain, 391, 426–431, 431–433, 504
 Small-signal differential-mode voltage gain, 793
 Small-signal equivalent circuit, 47, 258–260, 378, 399
 Small-signal equivalent circuit bipolar differential amplifier, 763
 Small-signal hybrid- π equivalent circuit, 378–381
 Small-signal hybrid- π model of npn bipolar transistor, 381
 Small-signal incremental conductance, 45
 Small-signal incremental resistance, 45
 Small-signal input resistance, 391
 Small-signal output admittance, 391
 Small-signal output resistance, 211
 Small-signal power gain, 442
 Small-signal transconductance, 236
 Small-signal transistor output resistance, 385
 Small-signal voltage, 431–433
 Small-signal voltage gain, 381–384, 420–423
 Snapback breakdown, 145
 SOA, 563
 Solar cell, 48–49
 Sound wave, 206, 370
 Source follower, 227
 Source-follower amplifier, 227–233, 238
 Source regulation, 88, 89
 Source-substrate-drain structure, 145
 Source terminal, 128
 Space-charge region, 24
 Speedup diode, 1295
 SPICE, 41
 Squaring network, 1292
 SRAM, 1209, 1212–1220
 SRAM read/write circuitry, 1218–1220
 Stability of feedback circuit, 908–918
 Bode plot, 909–912
 gain margin, 917
 Nyquist stability criterion, 913–916
 phase margin, 916
 stability problem, 908–909
 Stability problem, 908–909
 Stabilize, 335
 Standard class-A amplifier, 572
 Standard resistor and capacitor values, 1329–1331
 capacitors, 1331
 carbon resistors, 1329–1330
 precision resistors (1% tolerance), 1329–1330
 Static CMOS logic circuit, 1183
 Static CMOS logic gate, 1232–1234
 Static ECL gate, 1298–1299
 Static NMOS RAM cells, 1212
 Static RAM (SRAM), 1209, 1212–1220
 Storage time, 33
 Subthreshold conduction, 144
 Subthreshold current, 144
 Summing amplifier, 636–638
 Summing op-amp amplifier circuit, 636
 Summing op-amp circuit design, 659–661
 Sunrayce car, 49
 Superposition principle, 371, 636, 1085
 Switch/switching, 165–166, 323–325
 Switched-capacitor circuit, 1072
 Switched-capacitor filter, 1070–1074
 Switching time, 1144
 Switching transient, 32–33
 Symmetrical sinusoidal signals, 415
 Symmetrical swing, 415–420
 System design, 616–617
 System transfer functions, 472–485
 Bode plots, 474–479
 first-order functions, 473–474
 s-domain analysis, 472–473
 short-circuit and open-circuit time constants, 480–483
 time response, 483–485

T

- T-model, 396
 T-network, 631–633
 Taylor series, 374, 375
 Temperature coefficient of offset voltage, 1045
 Temperature compensation, 1121
 Temperature effect, 145
 Temperature effects, 1045–1046
 Temperature warning light circuit, 617
 THD, 560, 583
 Thermal resistance, 567
 Thermal voltage, 24
 Thermometer
 bipolar transistor, 348–350
 diode, 54–56
 instrumentation amplifier, 665–668
 MOS transistor, 190–192
 Thevenin equivalent circuit, 217, 321, 333, 340, 396, 509
 Thevenin equivalent resistance, 231, 321, 340
 Thevenin equivalent voltage, 321, 340
 Thevenin resistance, 399
 3-bit flash A/D converter, 1229, 1230
 3 db frequency, 475
 Three-input CMOS NAND logic gate, 1187
 Three-stage CMOS op-amp, 973–976
 Three-terminal voltage regulator, 1122

Three-transistor current mirror, 695
 Three-transistor current source, 693–696
 Threshold comparator, 1101
 Threshold voltage, 130, 1150
 Time constant, 480, 1096
 Time constant technique, 488
 Time response, 483–485
 Time-varying base current, 373
 Time-varying signal source, 207, 373
 Timer circuit, 1105
 Top-down design process, 615–616
 Total average power, 580
 Total base current, 443
 Total collector current, 443
 Total gate-to-source voltage, 206
 Total harmonic distortion (THD), 560, 583. *See also* Harmonic distortion
 Total instantaneous collector-emitter voltage, 443
 Total instantaneous power dissipation, 579
 Total instantaneous value, 4, 208, 372
 Totem-pole output stage, 1284–1285
 Transconductance
 back-gate, 215
 basic JFET amplifiers, 259, 262
 calculation of, 208–209
 composite, 802
 defined, 208
 forward, 760
 OTA, 657–658
 proportional to conduction parameter, 209
 proportional to saturation current, 259
 series-series amplifier, 886–892
 small-signal, 236
 small-signal analysis, 262
 small-signal hybrid- π equivalent circuit, 379
 unloaded, 657
 Transconductance amplifier, 397, 870
 Transconductance function, 472
 Transducer, 663
 Transfer functions. *See* System transfer functions
 Transformer-coupled common-emitter amplifier, 588–589
 Transformer-coupled emitter-follower amplifier, 589–591
 Transformer turns ratio, 69
 Transistor
 BJT. *See* Bipolar junction transistor (BJT)
 common-emitter bipolar, 391
 common-emitter npn, 391
 enhancement-mode NMOS, 1222
 EPROM, 1224–1225
 FET. *See* FET
 generally, 687
 input buffer, 596–599
 JFET, 180–190
 matched, 950
 MESFET, 182–190
 MOSFET. *See* MOSFET
 power, 560–571
 Schottky, 1289–1291
 Transistor circuit characteristics, 851
 Transistor equivalent circuit, 304
 Transistor sizing
 CMOS inverter, 1175–1176, 1184–1185
 CMOS logic gates, 1185–1187
 Transistor-transistor logic, 1277–1289
 basic diode-transistor logic gate, 1277–1279
 basic DTL NAND circuit operation, 1277–1279
 basic TTL NAND circuit, 1282–1284
 dc current-voltage analysis, 1282–1284
 fanout, 1285–1286
 input transistor, 1280–1282
 minimum β , 1279
 modified totem-pole output stage, 1287–1288
 pull-down resistor, 1279

totem-pole output stage, 1284–1285
 tristate output, 1288
 Transition point, 154, 1149, 1169, 1172
 Transmission gates, 1194–1202
 CMOS, 1200–1201
 CMOS pass networks, 1202
 NMOS, 1194–1198
 NMOS pass networks, 1198–1199
 Transresistance (shunt-shunt amplifier), 833–901
 Transresistance amplifier, 397, 872
 Transresistance function, 472
 Transresistance transfer function, 895
 Trigger comparator, 1103
 Triode region, 132
 Tristate output, 1288
 TTL. *See* Transistor-transistor logic
 Turn-off time, 33
 Turn-on time, 33
 Turn-on voltage, 38
 Turns ratio, 71, 589
 Two-diode circuit, 99
 Two-input BiCMOS NOR circuit, 1298
 Two-input CMOS NAND logic circuit, 1184
 Two-input CMOS NOR logic circuit, 1183
 Two-input diode AND logic circuit, 105
 Two-input diode OR logic circuit, 104
 Two-input NMOS NOR logic gate, 167, 1163
 Two-pole high-pass Butterworth filter, 1068
 Two-pole low-pass Butterworth filter, 1065–1067
 Two-port equivalent circuit, 403
 Two-sided output, 761
 BJT differential pair, 761, 769–770
 FET differential pair, 787–788
 Two-stage amplifier, 264–266
 Two-stage amplifier with coupling capacitors, 537–539
 Two-stage CMOS op-amp, 992–995
 Two-transistor current mirror, 695
 Two-transistor current source, 343, 688–693
 current relationships, 689
 IC fabrication, 692
 mismatched transistors, 692–693
 output resistance, 690–691
 Two-transistor MOSFET current source, 707–711
 Two-transistor NMOS current source, 707
 2N2222 bipolar transistor, 393–396
 2N2222 npn bipolar transistor, 1318–1319
 2N2907 pnp bipolar transistor, 1320–1321

U

Ultraviolet (UV) light, 1224
 Undefined range, 1179
 Unity-gain bandwidth, 506, 516–518, 1024
 Unity-gain frequency, 517
 Unloaded transconductance, 657
 Upper corner, 481
 Uppercase letter/lowercase subscript, 4
 Uppercase letter/uppercase subscript, 4
 User-programmed ROM, 1222–1223
 UV radiation, 1224

V

V_{BE} , 288n
 V_{BE} multiplier, 590
 Valence electron, 10
 Varactor diode, 26
 Velocity saturation, 142
 Vertical process, 566
 Virtual ground, 625, 628
 Virtual short, 638
 VMOS device, 566

Voice communication, 3
 Volatile memory, 1209
 Voltage
 ac collector-emitter, 373
 ac ripple, 68n
 base-collector, 1259
 breakdown, 52
 collector-emitter, 1259
 common-mode input, 648, 754
 cut-in, 38
 differential-mode input, 754, 760
 diode breakdown, 52
 drain-to-source, 130, 131
 Early, 695
 early, 296
 gate-to-source, 130, 225
 input offset, 1031
 instantaneous gate-to-source, 258
 NMOS shift register, 1204
 pinchoff, 183
 PIV, 32
 quiescent drain-to-source, 225
 quiescent gate-to-source, 236
 ripple, 76–77
 sinusoidal base-emitter, 375
 small reverse-biased gate, 181
 small-signal, 431–433
 temperature coefficient of offset, 1045
 thermal, 24
 Thevenin equivalent, 321, 333, 340, 590
 threshold, 130
 total gate-to-source, 206
 total instantaneous collector-emitter, 443
 turn-on, 38
 zero gate, 181
 Voltage amplifier, 397, 873–879
 Voltage divider biasing, 333–339
 Voltage doubler circuit, 83–84
 Voltage feedback ratio, 391
 Voltage follower, 639–640
 Voltage-follower op-amp circuit, 640
 Voltage notation, 130n, 288n
 Voltage regulator, 84, 1114–1122
 basic regulator description, 1114
 defined, 1114
 LM78LXX series three-terminal positive, 1119
 output resistance/load regulation, 1115–1116
 positive, 1119–1122
 simple series-pass regulator, 1116–1118
 three-terminal, 1122
 Voltage transfer curve, 311, 582, 1171–1175
 Voltage transfer function, 472

W

Weighted-resistor 4-bit D/A converter, 1227–1228
 Wide-swing current mirror, 713–714
 Wide-swing MOSFET cascode current mirror, 714
 Widlar current source, 698–704
 Wien-bridge circuit, 1079–1082
 Wilson current mirror, 713
 Wilson current source, 697–698, 713

Z

z -transform technique, 1074
 Zener breakdown, 32
 Zener diode, 51–53
 Zener diode circuit, 84–90
 ideal voltage reference circuit, 85–88
 Zener resistance/percent regulation, 88–90
 Zener diode voltage regulator circuit, 85, 88
 Zener resistance, 84
 Zero gate voltage, 181