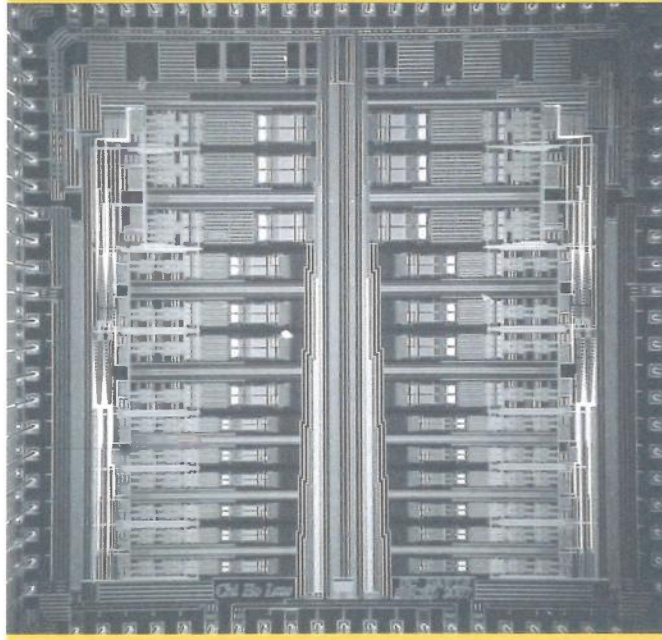


ANALYSIS AND DESIGN OF ANALOG INTEGRATED CIRCUITS

Paul R. Gray ■ Paul J. Hurst ■ Stephen H. Lewis ■ Robert G. Meyer

Fifth Edition



ANALYSIS AND DESIGN OF ANALOG INTEGRATED CIRCUITS

Fifth Edition

PAUL R. GRAY

University of California, Berkeley

PAUL J. HURST

University of California, Davis

STEPHEN H. LEWIS

University of California, Davis

ROBERT G. MEYER

University of California, Berkeley



WILEY

New York / Chichester / Weinheim / Brisbane / Singapore / Toronto

PUBLISHER	Don Fowley
ACQUISITIONS EDITOR	Daniel Sayre
SENIOR PRODUCTION EDITOR	Valerie A. Vargas
EXECUTIVE MARKETING MANAGER	Christopher Ruel
DESIGNER	Arthur Medina
PRODUCTION MANAGEMENT SERVICES	Elm Street Publishing Services
EDITORIAL ASSISTANT	Carolyn Weisman
MEDIA EDITOR	Lauren Sapira

Cover courtesy of Chi Ho Law.

This book was set in 10/12 Times Roman by Thomson Digital and printed and bound by Hamilton Printing Company. The cover was printed by Phoenix Color, Inc.

This book was printed on acid-free paper. ♻

Copyright 2009 © John Wiley & Sons, Inc. All rights reserved.

No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning or otherwise, except as permitted under Sections 107 or 108 of the 1976 United States Copyright Act, without either the prior written permission of the Publisher, or authorization through payment of the appropriate per-copy fee to the Copyright Clearance Center Inc., 222 Rosewood Drive, Danvers, MA 01923, website www.copyright.com. Requests to the Publisher for permission should be addressed to the Permissions Department, John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030-5774, (201) 748-6011, fax (201) 748-6008, website <http://www.wiley.com/go/permissions>. To order books or for customer service please call 1-800-CALL-WILEY (255-5945).

<http://www.wiley.com/college/gray>

Library of Congress Cataloging-in-Publication Data

Analysis and design of analog integrated circuits / Paul R. Gray . . . [et al.]. — 5th ed.
p. cm.

Includes bibliographical references and index.

ISBN 978-0-470-24599-6 (cloth: alk. paper)

1. Linear integrated circuits-Computer-aided design. 2. Metal oxide semiconductors-Computer-aided design. 3. Bipolar transistors-Computer-aided design.
I. Gray, Paul R., 1942-

TK7874.A588 2009

621.3815-dc21

08-043583

Printed in the United States of America

10 9 8 7 6 5 3 2 1

To Liz, Barbara, Robin, and Judy

Preface

Since the publication of the first edition of this book, the field of analog integrated circuits has developed and matured. The initial groundwork was laid in bipolar technology, followed by a rapid evolution of MOS analog integrated circuits. Thirty years ago, CMOS technologies were fast enough to support applications only at audio frequencies. However, the continuing reduction of the minimum feature size in integrated-circuit (IC) technologies has greatly increased the maximum operating frequencies, and CMOS technologies have become fast enough for many new applications as a result. For example, the bandwidth in some video applications is about 4 MHz, requiring bipolar technologies as recently as about twenty-three years ago. Now, however, CMOS easily can accommodate the required bandwidth for video and is being used for radio-frequency applications. Today, bipolar integrated circuits are used in some applications that require very low noise, very wide bandwidth, or driving low-impedance loads.

In this fifth edition, coverage of the bipolar 741 op amp has been replaced with a low-voltage bipolar op amp, the NE5234, with rail-to-rail common-mode input range and almost rail-to-rail output swing. Analysis of a fully differential CMOS folded-cascode operational amplifier (op amp) is now included in Chapter 12. The 560B phase-locked loop, which is no longer commercially available, has been deleted from Chapter 10.

The SPICE computer analysis program is now readily available to virtually all electrical engineering students and professionals, and we have included extensive use of SPICE in this edition, particularly as an integral part of many problems. We have used computer analysis as it is most commonly employed in the engineering design process—both as a more accurate check on hand calculations, and also as a tool to examine complex circuit behavior beyond the scope of hand analysis.

An in-depth look at SPICE as an indispensable tool for IC robust design can be found in *The SPICE Book*, 2nd ed., published by J. Wiley and Sons. This text contains many worked out circuit designs and verification examples linked to the multitude of analyses available in the most popular versions of SPICE. The *SPICE Book* conveys the role of simulation as an integral part of the design process, but not as a replacement for solid circuit-design knowledge.

This book is intended to be useful both as a text for students and as a reference book for practicing engineers. For class use, each chapter includes many worked problems; the problem sets at the end of each chapter illustrate the practical applications of the material in the text. All of the authors have extensive industrial experience in IC design and in the teaching of courses on this subject; this experience is reflected in the choice of text material and in the problem sets.

Although this book is concerned largely with the analysis and design of ICs, a considerable amount of material also is included on applications. In practice, these two subjects are closely linked, and a knowledge of both is essential for designers and users of ICs. The latter compose the larger group by far, and we believe that a working knowledge of IC design is a great advantage to an IC user. This is particularly apparent when the user must choose from among a number of competing designs to satisfy a particular need. An understanding of the IC structure is then useful in evaluating the relative desirability of the different designs under extremes of environment or in the presence of variations in supply voltage. In addition, the IC user is in a

much better position to interpret a manufacturer's data if he or she has a working knowledge of the internal operation of the integrated circuit.

The contents of this book stem largely from courses on analog integrated circuits given at the University of California at the Berkeley and Davis campuses. The courses are senior-level electives and first-year graduate courses. The book is structured so that it can be used as the basic text for a sequence of such courses. The more advanced material is found at the end of each chapter or in an appendix so that a first course in analog integrated circuits can omit this material without loss of continuity. An outline of each chapter is given below with suggestions for material to be covered in such a first course. It is assumed that the course consists of three hours of lecture per week over a fifteen-week semester and that the students have a working knowledge of Laplace transforms and frequency-domain circuit analysis. It is also assumed that the students have had an introductory course in electronics so that they are familiar with the principles of transistor operation and with the functioning of simple analog circuits. Unless otherwise stated, each chapter requires three to four lecture hours to cover.

Chapter 1 contains a summary of bipolar transistor and MOS transistor device physics. We suggest spending one week on selected topics from this chapter, with the choice of topics depending on the background of the students. The material of Chapters 1 and 2 is quite important in IC design because there is significant interaction between circuit and device design, as will be seen in later chapters. A thorough understanding of the influence of device fabrication on device characteristics is essential.

Chapter 2 is concerned with the technology of IC fabrication and is largely descriptive. One lecture on this material should suffice if the students are assigned the chapter to read.

Chapter 3 deals with the characteristics of elementary transistor connections. The material on one-transistor amplifiers should be a review for students at the senior and graduate levels and can be assigned as reading. The section on two-transistor amplifiers can be covered in about three hours, with greatest emphasis on differential pairs. The material on device mismatch effects in differential amplifiers can be covered to the extent that time allows.

In Chapter 4, the important topics of current mirrors and active loads are considered. These configurations are basic building blocks in modern analog IC design, and this material should be covered in full, with the exception of the material on band-gap references and the material in the appendices.

Chapter 5 is concerned with output stages and methods of delivering output power to a load. Integrated-circuit realizations of Class A, Class B, and Class AB output stages are described, as well as methods of output-stage protection. A selection of topics from this chapter should be covered.

Chapter 6 deals with the design of operational amplifiers (op amps). Illustrative examples of dc and ac analysis in both MOS and bipolar op amps are performed in detail, and the limitations of the basic op amps are described. The design of op amps with improved characteristics in both MOS and bipolar technologies are considered. This key chapter on amplifier design requires at least six hours.

In Chapter 7, the frequency response of amplifiers is considered. The zero-value time-constant technique is introduced for the calculations of the -3 -dB frequency of complex circuits. The material of this chapter should be considered in full.

Chapter 8 describes the analysis of feedback circuits. Two different types of analysis are presented: two-port and return-ratio analyses. Either approach should be covered in full with the section on voltage regulators assigned as reading.

Chapter 9 deals with the frequency response and stability of feedback circuits and should be covered up to the section on root locus. Time may not permit a detailed discussion of root locus, but some introduction to this topic can be given.

In a fifteen-week semester, coverage of the above material leaves about two weeks for Chapters 10, 11, and 12. A selection of topics from these chapters can be chosen as follows. Chapter 10 deals with nonlinear analog circuits and portions of this chapter up to Section 10.2 could be covered in a first course. Chapter 11 is a comprehensive treatment of noise in integrated circuits and material up to and including Section 11.4 is suitable. Chapter 12 describes fully differential operational amplifiers and common-mode feedback and may be best suited for a second course.

We are grateful to the following colleagues for their suggestions for and/or evaluation of this book: R. Jacob Baker, Bernhard E. Boser, A. Paul Brokaw, Iwen Chao, John N. Churchill, David W. Cline, Kenneth C. Dyer, Ozan E. Erdoğan, John W. Fattaruso, Weinan Gao, Edwin W. Greeneich, Alex Gros-Balthazard, Tünde Gyurics, Ward J. Helms, Kaveh Hosseini, Timothy H. Hu, Shafiq M. Jamal, John P. Keane, Haideh Khorramabadi, Pak Kim Lau, Thomas W. Matthews, Krishnaswamy Nagaraj, Khalil Najafi, Borivoje Nikolić, Keith O'Donoghue, Robert A. Pease, Lawrence T. Pileggi, Edgar Sánchez-Sinencio, Bang-Sup Song, Richard R. Spencer, Eric J. Swanson, Andrew Y. J. Szeto, Yannis P. Tsividis, Srikanth Vaidianathan, T. R. Viswanathan, Chorng-Kuang Wang, Dong Wang, and Mo Maggie Zhang. We are also grateful to Darrel Akers, Mu Jane Lee, Lakshmi Rao, Nattapol Sitthimahachaikul, Haoyue Wang, and Mo Maggie Zhang for help with proofreading, and to Chi Ho Law for allowing us to use on the cover of this book a die photograph of an integrated circuit he designed. Finally, we would like to thank the staffs at Wiley and Elm Street Publishing Services for their efforts in producing this edition.

The material in this book has been greatly influenced by our association with the late Donald O. Pederson, and we acknowledge his contributions.

Berkeley and Davis, CA, 2008

*Paul R. Gray
Paul J. Hurst
Stephen H. Lewis
Robert G. Meyer*

Contents

CHAPTER 1	
Models for Integrated-Circuit Active Devices 1	
1.1	Introduction 1
1.2	Depletion Region of a pn Junction 1
1.2.1	Depletion-Region Capacitance 5
1.2.2	Junction Breakdown 6
1.3	Large-Signal Behavior of Bipolar Transistors 8
1.3.1	Large-Signal Models in the Forward-Active Region 8
1.3.2	Effects of Collector Voltage on Large-Signal Characteristics in the Forward-Active Region 14
1.3.3	Saturation and Inverse-Active Regions 16
1.3.4	Transistor Breakdown Voltages 20
1.3.5	Dependence of Transistor Current Gain β_F on Operating Conditions 23
1.4	Small-Signal Models of Bipolar Transistors 25
1.4.1	Transconductance 26
1.4.2	Base-Charging Capacitance 27
1.4.3	Input Resistance 28
1.4.4	Output Resistance 29
1.4.5	Basic Small-Signal Model of the Bipolar Transistor 30
1.4.6	Collector-Base Resistance 30
1.4.7	Parasitic Elements in the Small-Signal Model 31
1.4.8	Specification of Transistor Frequency Response 34
1.5	Large-Signal Behavior of Metal-Oxide-Semiconductor Field-Effect Transistors 38
1.5.1	Transfer Characteristics of MOS Devices 38
1.5.2	Comparison of Operating Regions of Bipolar and MOS Transistors 45
1.5.3	Decomposition of Gate-Source Voltage 47
1.5.4	Threshold Temperature Dependence 47
1.5.5	MOS Device Voltage Limitations 48
1.6	Small-Signal Models of MOS Transistors 49
1.6.1	Transconductance 50
1.6.2	Intrinsic Gate-Source and Gate-Drain Capacitance 51
1.6.3	Input Resistance 52
1.6.4	Output Resistance 52
1.6.5	Basic Small-Signal Model of the MOS Transistor 52
1.6.6	Body Transconductance 53
1.6.7	Parasitic Elements in the Small-Signal Model 54
1.6.8	MOS Transistor Frequency Response 55
1.7	Short-Channel Effects in MOS Transistors 59
1.7.1	Velocity Saturation from the Horizontal Field 59
1.7.2	Transconductance and Transition Frequency 63
1.7.3	Mobility Degradation from the Vertical Field 65
1.8	Weak Inversion in MOS Transistors 65
1.8.1	Drain Current in Weak Inversion 66
1.8.2	Transconductance and Transition Frequency in Weak Inversion 69
1.9	Substrate Current Flow in MOS Transistors 71
A.1.1	Summary of Active-Device Parameters 73

CHAPTER 2		
Bipolar, MOS, and BiCMOS		
Integrated-Circuit Technology 78		
2.1	Introduction 78	
2.2	Basic Processes in Integrated-Circuit Fabrication 79	
2.2.1	Electrical Resistivity of Silicon 79	
2.2.2	Solid-State Diffusion 80	
2.2.3	Electrical Properties of Diffused Layers 82	
2.2.4	Photolithography 84	
2.2.5	Epitaxial Growth 86	
2.2.6	Ion Implantation 87	
2.2.7	Local Oxidation 87	
2.2.8	Polysilicon Deposition 87	
2.3	High-Voltage Bipolar Integrated-Circuit Fabrication 88	
2.4	Advanced Bipolar Integrated-Circuit Fabrication 92	
2.5	Active Devices in Bipolar Analog Integrated Circuits 95	
2.5.1	Integrated-Circuit <i>npn</i> Transistors 96	
2.5.2	Integrated-Circuit <i>pnp</i> Transistors 107	
2.6	Passive Components in Bipolar Integrated Circuits 115	
2.6.1	Diffused Resistors 115	
2.6.2	Epitaxial and Epitaxial Pinch Resistors 119	
2.6.3	Integrated-Circuit Capacitors 120	
2.6.4	Zener Diodes 121	
2.6.5	Junction Diodes 122	
2.7	Modifications to the Basic Bipolar Process 123	
2.7.1	Dielectric Isolation 123	
2.7.2	Compatible Processing for High-Performance Active Devices 124	
2.7.3	High-Performance Passive Components 127	
2.8	MOS Integrated-Circuit Fabrication 127	
2.9	Active Devices in MOS Integrated Circuits 131	
2.9.1	<i>n</i> -Channel Transistors 131	
2.9.2	<i>p</i> -Channel Transistors 144	
2.9.3	Depletion Devices 144	
2.9.4	Bipolar Transistors 145	
2.10	Passive Components in MOS Technology 146	
2.10.1	Resistors 146	
2.10.2	Capacitors in MOS Technology 148	
2.10.3	Latchup in CMOS Technology 151	
2.11	BiCMOS Technology 152	
2.12	Heterojunction Bipolar Transistors 153	
2.13	Interconnect Delay 156	
2.14	Economics of Integrated-Circuit Fabrication 156	
2.14.1	Yield Considerations in Integrated-Circuit Fabrication 157	
2.14.2	Cost Considerations in Integrated-Circuit Fabrication 159	
A.2.1	SPICE Model-Parameter Files 162	
CHAPTER 3		
Single-Transistor and Multiple-Transistor Amplifiers 169		
3.1	Device Model Selection for Approximate Analysis of Analog Circuits 170	
3.2	Two-Port Modeling of Amplifiers 171	
3.3	Basic Single-Transistor Amplifier Stages 173	
3.3.1	Common-Emitter Configuration 174	
3.3.2	Common-Source Configuration 178	
3.3.3	Common-Base Configuration 182	
3.3.4	Common-Gate Configuration 185	
3.3.5	Common-Base and Common-Gate Configurations with Finite r_o 187	
3.3.5.1	Common-Base and Common-Gate Input Resistance 187	
3.3.5.2	Common-Base and Common-Gate Output Resistance 189	

3.3.6	Common-Collector Configuration (Emitter Follower)	191	3.5.6.8	Offset Voltage Drift in the Source-Coupled Pair	236
3.3.7	Common-Drain Configuration (Source Follower)	194	3.5.6.9	Small-Signal Characteristics of Unbalanced Differential Amplifiers	237
3.3.8	Common-Emitter Amplifier with Emitter Degeneration	196	A.3.1	Elementary Statistics and the Gaussian Distribution	244
3.3.9	Common-Source Amplifier with Source Degeneration	199			
3.4	Multiple-Transistor Amplifier Stages	201	CHAPTER 4	Current Mirrors, Active Loads, and References	251
3.4.1	The CC-CE, CC-CC, and Darlington Configurations	201	4.1	Introduction	251
3.4.2	The Cascode Configuration	205	4.2	Current Mirrors	251
3.4.2.1	The Bipolar Cascode	205	4.2.1	General Properties	251
3.4.2.2	The MOS Cascode	207	4.2.2	Simple Current Mirror	253
3.4.3	The Active Cascode	210	4.2.2.1	Bipolar	253
3.4.4	The Super Source Follower	212	4.2.2.2	MOS	255
3.5	Differential Pairs	214	4.2.3	Simple Current Mirror with Beta Helper	258
3.5.1	The dc Transfer Characteristic of an Emitter-Coupled Pair	214	4.2.3.1	Bipolar	258
3.5.2	The dc Transfer Characteristic with Emitter Degeneration	216	4.2.3.2	MOS	260
3.5.3	The dc Transfer Characteristic of a Source-Coupled Pair	217	4.2.4	Simple Current Mirror with Degeneration	260
3.5.4	Introduction to the Small-Signal Analysis of Differential Amplifiers	220	4.2.4.1	Bipolar	260
3.5.5	Small-Signal Characteristics of Balanced Differential Amplifiers	223	4.2.4.2	MOS	261
3.5.6	Device Mismatch Effects in Differential Amplifiers	229	4.2.5	Cascode Current Mirror	261
3.5.6.1	Input Offset Voltage and Current	230	4.2.5.1	Bipolar	261
3.5.6.2	Input Offset Voltage of the Emitter-Coupled Pair	230	4.2.5.2	MOS	264
3.5.6.3	Offset Voltage of the Emitter-Coupled Pair: Approximate Analysis	231	4.2.6	Wilson Current Mirror	272
3.5.6.4	Offset Voltage Drift in the Emitter-Coupled Pair	233	4.2.6.1	Bipolar	272
3.5.6.5	Input Offset Current of the Emitter-Coupled Pair	233	4.2.6.2	MOS	275
3.5.6.6	Input Offset Voltage of the Source-Coupled Pair	234	4.3	Active Loads	276
3.5.6.7	Offset Voltage of the Source-Coupled Pair: Approximate Analysis	235	4.3.1	Motivation	276
			4.3.2	Common-Emitter–Common-Source Amplifier with Complementary Load	277
			4.3.3	Common-Emitter–Common-Source Amplifier with Depletion Load	280
			4.3.4	Common-Emitter–Common-Source Amplifier with Diode-Connected Load	282
			4.3.5	Differential Pair with Current-Mirror Load	285
			4.3.5.1	Large-Signal Analysis	285
			4.3.5.2	Small-Signal Analysis	286
			4.3.5.3	Common-Mode Rejection Ratio	291

6.2	Deviations from Ideality in Real Operational Amplifiers	415			
6.2.1	Input Bias Current	415			
6.2.2	Input Offset Current	416			
6.2.3	Input Offset Voltage	416			
6.2.4	Common-Mode Input Range	416			
6.2.5	Common-Mode Rejection Ratio (CMRR)	417			
6.2.6	Power-Supply Rejection Ratio (PSRR)	418			
6.2.7	Input Resistance	420			
6.2.8	Output Resistance	420			
6.2.9	Frequency Response	420			
6.2.10	Operational-Amplifier Equivalent Circuit	420			
6.3	Basic Two-Stage MOS Operational Amplifiers	421			
6.3.1	Input Resistance, Output Resistance, and Open-Circuit Voltage Gain	422			
6.3.2	Output Swing	423			
6.3.3	Input Offset Voltage	424			
6.3.4	Common-Mode Rejection Ratio	427			
6.3.5	Common-Mode Input Range	427			
6.3.6	Power-Supply Rejection Ratio (PSRR)	430			
6.3.7	Effect of Overdrive Voltages	434			
6.3.8	Layout Considerations	435			
6.4	Two-Stage MOS Operational Amplifiers with Cascodes	438			
6.5	MOS Telescopic-Cascode Operational Amplifiers	439			
6.6	MOS Folded-Cascode Operational Amplifiers	442			
6.7	MOS Active-Cascode Operational Amplifiers	446			
6.8	Bipolar Operational Amplifiers	448			
6.8.1	The dc Analysis of the NE5234 Operational Amplifier	452			
6.8.2	Transistors that Are Normally Off	467			
6.8.3	Small-Signal Analysis of the NE5234 Operational Amplifier	469			
6.8.4	Calculation of the Input Offset Voltage and Current of the NE5234	477			
			CHAPTER 7		
			Frequency Response of Integrated Circuits	490	
			7.1	Introduction	490
			7.2	Single-Stage Amplifiers	490
			7.2.1	Single-Stage Voltage Amplifiers and the Miller Effect	490
			7.2.1.1	The Bipolar Differential Amplifier: Differential-Mode Gain	495
			7.2.1.2	The MOS Differential Amplifier: Differential-Mode Gain	499
			7.2.2	Frequency Response of the Common-Mode Gain for a Differential Amplifier	501
			7.2.3	Frequency Response of Voltage Buffers	503
			7.2.3.1	Frequency Response of the Emitter Follower	505
			7.2.3.2	Frequency Response of the Source Follower	511
			7.2.4	Frequency Response of Current Buffers	514
			7.2.4.1	Common-Base Amplifier Frequency Response	516
			7.2.4.2	Common-Gate Amplifier Frequency Response	517
			7.3	Multistage Amplifier Frequency Response	518
			7.3.1	Dominant-Pole Approximation	518
			7.3.2	Zero-Value Time Constant Analysis	519
			7.3.3	Cascode Voltage-Amplifier Frequency Response	524
			7.3.4	Cascode Frequency Response	527
			7.3.5	Frequency Response of a Current Mirror Loading a Differential Pair	534
			7.3.6	Short-Circuit Time Constants	536
			7.4	Analysis of the Frequency Response of the NE5234 Op Amp	539
			7.4.1	High-Frequency Equivalent Circuit of the NE5234	539
			7.4.2	Calculation of the -3 -dB Frequency of the NE5234	540
			7.4.3	Nondominant Poles of the NE5234	542

7.5	Relation Between Frequency Response and Time Response	542	9.3	Instability and the Nyquist Criterion	626
CHAPTER 8			9.4	Compensation	633
Feedback			9.4.1	Theory of Compensation	633
8.1	Ideal Feedback Equation	553	9.4.2	Methods of Compensation	637
8.2	Gain Sensitivity	555	9.4.3	Two-Stage MOS Amplifier Compensation	643
8.3	Effect of Negative Feedback on Distortion	555	9.4.4	Compensation of Single-Stage CMOS Op Amps	650
8.4	Feedback Configurations	557	9.4.5	Nested Miller Compensation	654
8.4.1	Series-Shunt Feedback	557	9.5	Root-Locus Techniques	664
8.4.2	Shunt-Shunt Feedback	560	9.5.1	Root Locus for a Three-Pole Transfer Function	665
8.4.3	Shunt-Series Feedback	561	9.5.2	Rules for Root-Locus Construction	667
8.4.4	Series-Series Feedback	562	9.5.3	Root Locus for Dominant-Pole Compensation	676
8.5	Practical Configurations and the Effect of Loading	563	9.5.4	Root Locus for Feedback-Zero Compensation	677
8.5.1	Shunt-Shunt Feedback	563	9.6	Slew Rate	681
8.5.2	Series-Series Feedback	569	9.6.1	Origin of Slew-Rate Limitations	681
8.5.3	Series-Shunt Feedback	579	9.6.2	Methods of Improving Slew-Rate in Two-Stage Op Amps	685
8.5.4	Shunt-Series Feedback	583	9.6.3	Improving Slew-Rate in Bipolar Op Amps	687
8.5.5	Summary	587	9.6.4	Improving Slew-Rate in MOS Op Amps	688
8.6	Single-Stage Feedback	587	9.6.5	Effect of Slew-Rate Limitations on Large-Signal Sinusoidal Performance	692
8.6.1	Local Series-Series Feedback	587	A.9.1	Analysis in Terms of Return-Ratio Parameters	693
8.6.2	Local Series-Shunt Feedback	591	A.9.2	Roots of a Quadratic Equation	694
8.7	The Voltage Regulator as a Feedback Circuit	593	CHAPTER 10		
8.8	Feedback Circuit Analysis Using Return Ratio	599	Nonlinear Analog Circuits		
8.8.1	Closed-Loop Gain Using Return Ratio	601	10.1	Introduction	704
8.8.2	Closed-Loop Impedance Formula Using Return Ratio	607	10.2	Analog Multipliers Employing the Bipolar Transistor	704
8.8.3	Summary—Return-Ratio Analysis	612	10.2.1	The Emitter-Coupled Pair as a Simple Multiplier	704
8.9	Modeling Input and Output Ports in Feedback Circuits	613	10.2.2	The dc Analysis of the Gilbert Multiplier Cell	706

10.2.3	The Gilbert Cell as an Analog Multiplier	708	11.6.2	Effect of Practical Feedback on Noise Performance	765
10.2.4	A Complete Analog Multiplier	711	11.7	Noise Performance of Other Transistor Configurations	771
10.2.5	The Gilbert Multiplier Cell as a Balanced Modulator and Phase Detector	712	11.7.1	Common-Base Stage Noise Performance	771
10.3	Phase-Locked Loops (PLL)	716	11.7.2	Emitter-Follower Noise Performance	773
10.3.1	Phase-Locked Loop Concepts	716	11.7.3	Differential-Pair Noise Performance	773
10.3.2	The Phase-Locked Loop in the Locked Condition	718	11.8	Noise in Operational Amplifiers	776
10.3.3	Integrated-Circuit Phase-Locked Loops	727	11.9	Noise Bandwidth	782
10.4	Nonlinear Function Synthesis	731	11.10	Noise Figure and Noise Temperature	786
CHAPTER 11			11.10.1	Noise Figure	786
Noise in Integrated Circuits			11.10.2	Noise Temperature	790
11.1	Introduction	736	CHAPTER 12		
11.2	Sources of Noise	736	Fully Differential Operational Amplifiers		
11.2.1	Shot Noise	736	12.1	Introduction	796
11.2.2	Thermal Noise	740	12.2	Properties of Fully Differential Amplifiers	796
11.2.3	Flicker Noise ($1/f$ Noise)	741	12.3	Small-Signal Models for Balanced Differential Amplifiers	799
11.2.4	Burst Noise (<i>Popcorn Noise</i>)	742	12.4	Common-Mode Feedback	804
11.2.5	Avalanche Noise	743	12.4.1	Common-Mode Feedback at Low Frequencies	805
11.3	Noise Models of Integrated-Circuit Components	744	12.4.2	Stability and Compensation Considerations in a CMFB Loop	810
11.3.1	Junction Diode	744	12.5	CMFB Circuits	811
11.3.2	Bipolar Transistor	745	12.5.1	CMFB Using Resistive Divider and Amplifier	812
11.3.3	MOS Transistor	746	12.5.2	CMFB Using Two Differential Pairs	816
11.3.4	Resistors	747	12.5.3	CMFB Using Transistors in the Triode Region	819
11.3.5	Capacitors and Inductors	747	12.5.4	Switched-Capacitor CMFB	821
11.4	Circuit Noise Calculations	748	12.6	Fully Differential Op Amps	823
11.4.1	Bipolar Transistor Noise Performance	750	12.6.1	A Fully Differential Two-Stage Op Amp	823
11.4.2	Equivalent Input Noise and the Minimum Detectable Signal	754	12.6.2	Fully Differential Telescopic Cascode Op Amp	833
11.5	Equivalent Input Noise Generators	756			
11.5.1	Bipolar Transistor Noise Generators	757			
11.5.2	MOS Transistor Noise Generators	762			
11.6	Effect of Feedback on Noise Performance	764			
11.6.1	Effect of Ideal Feedback on Noise Performance	764			

xiv **Symbol Convention**

12.6.3 Fully Differential Folded-Cascode Op Amp	834	12.9 Analysis of a CMOS Fully Differential Folded-Cascode Op Amp	845
12.6.4 A Differential Op Amp with Two Differential Input Stages	835	12.9.1 DC Biasing	848
12.6.5 Neutralization	835	12.9.2 Low-Frequency Analysis	850
12.7 Unbalanced Fully Differential Circuits	838	12.9.3 Frequency and Time Responses in a Feedback Application	856
12.8 Bandwidth of the CMFB Loop	844	Index	871

Symbol Convention

Unless otherwise stated, the following symbol convention is used in this book. *Bias* or *dc* quantities, such as transistor collector current I_C and collector-emitter voltage V_{CE} , are represented by uppercase symbols with uppercase subscripts. Small-signal quantities, such as the incremental change in transistor collector current i_c , are represented by lowercase symbols with lowercase subscripts. Elements such as transconductance g_m in small-signal equivalent circuits are represented in the same way. Finally, quantities such as *total* collector current I_c , which represent the sum of the bias quantity *and* the signal quantity, are represented by an uppercase symbol with a lowercase subscript.

Models for Integrated-Circuit Active Devices

1.1 Introduction

The analysis and design of integrated circuits depend heavily on the utilization of suitable models for integrated-circuit components. This is true in hand analysis, where fairly simple models are generally used, and in computer analysis, where more complex models are encountered. Since any analysis is only as accurate as the model used, it is essential that the circuit designer have a thorough understanding of the origin of the models commonly utilized and the degree of approximation involved in each.

This chapter deals with the derivation of large-signal and small-signal models for integrated-circuit devices. The treatment begins with a consideration of the properties of *pn* junctions, which are basic parts of most integrated-circuit elements. Since this book is primarily concerned with circuit analysis and design, no attempt has been made to produce a comprehensive treatment of semiconductor physics. The emphasis is on summarizing the basic aspects of semiconductor-device behavior and indicating how these can be modeled by equivalent circuits.

1.2 Depletion Region of a *pn* Junction

The properties of reverse-biased *pn* junctions have an important influence on the characteristics of many integrated-circuit components. For example, reverse-biased *pn* junctions exist between many integrated-circuit elements and the underlying substrate, and these junctions all contribute voltage-dependent parasitic capacitances. In addition, a number of important characteristics of active devices, such as breakdown voltage and output resistance, depend directly on the properties of the depletion region of a reverse-biased *pn* junction. Finally, the basic operation of the junction field-effect transistor is controlled by the width of the depletion region of a *pn* junction. Because of its importance and application to many different problems, an analysis of the depletion region of a reverse-biased *pn* junction is considered below. The properties of forward-biased *pn* junctions are treated in Section 1.3 when bipolar-transistor operation is described.

Consider a *pn* junction under reverse bias as shown in Fig. 1.1. Assume *constant doping densities* of N_D atoms/cm³ in the *n*-type material and N_A atoms/cm³ in the *p*-type material. (The characteristics of junctions with nonconstant doping densities will be described later.) Due to the difference in carrier concentrations in the *p*-type and *n*-type regions, there exists a region at the junction where the mobile holes and electrons have been removed, leaving the fixed acceptor and donor ions. Each acceptor atom carries a negative charge and each donor atom carries a positive charge, so that the region near the junction is one of significant space charge and resulting high electric field. This is called the *depletion region* or *space-charge*

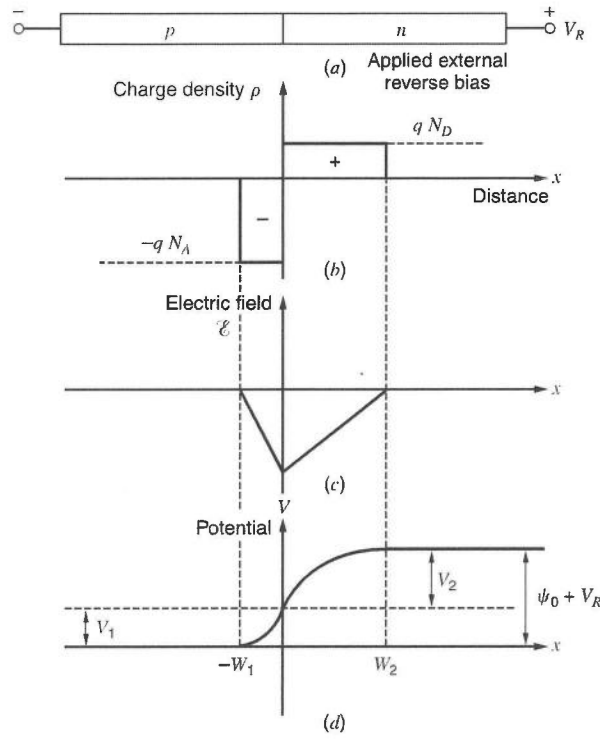


Figure 1.1 The abrupt junction under reverse bias V_R . (a) Schematic. (b) Charge density. (c) Electric field. (d) Electrostatic potential.

region. It is assumed that the edges of the depletion region are sharply defined as shown in Fig. 1.1, and this is a good approximation in most cases.

For zero applied bias, there exists a voltage ψ_0 across the junction called the *built-in potential*. This potential opposes the diffusion of mobile holes and electrons across the junction in equilibrium and has a value¹

$$\psi_0 = V_T \ln \frac{N_A N_D}{n_i^2} \quad (1.1)$$

where

$$V_T = \frac{kT}{q} \simeq 26 \text{ mV} \quad \text{at } 300^\circ\text{K}$$

the quantity n_i is the intrinsic carrier concentration in a pure sample of the semiconductor and $n_i \simeq 1.5 \times 10^{10} \text{ cm}^{-3}$ at 300°K for silicon.

In Fig. 1.1 the built-in potential is augmented by the applied reverse bias, V_R , and the total voltage across the junction is $(\psi_0 + V_R)$. If the depletion region penetrates a distance W_1 into the p -type region and W_2 into the n -type region, then we require

$$W_1 N_A = W_2 N_D \quad (1.2)$$

because the total charge per unit area on either side of the junction must be equal in magnitude but opposite in sign.

Poisson's equation in one dimension requires that

$$\frac{d^2 V}{dx^2} = -\frac{\rho}{\epsilon} = \frac{qN_A}{\epsilon} \quad \text{for} \quad -W_1 < x < 0 \quad (1.3)$$

where ρ is the charge density, q is the electron charge (1.6×10^{-19} coulomb), and ϵ is the permittivity of the silicon (1.04×10^{-12} farad/cm). The permittivity is often expressed as

$$\epsilon = K_S \epsilon_0 \quad (1.4)$$

where K_S is the dielectric constant of silicon and ϵ_0 is the permittivity of free space (8.86×10^{-14} F/cm). Integration of (1.3) gives

$$\frac{dV}{dx} = \frac{qN_A}{\epsilon} x + C_1 \quad (1.5)$$

where C_1 is a constant. However, the electric field $\mathcal{E}$ is given by

$$\mathcal{E} = -\frac{dV}{dx} = -\left(\frac{qN_A}{\epsilon} x + C_1\right) \quad (1.6)$$

Since there is zero electric field outside the depletion region, a boundary condition is

$$\mathcal{E} = 0 \quad \text{for} \quad x = -W_1$$

and use of this condition in (1.6) gives

$$\mathcal{E} = -\frac{qN_A}{\epsilon}(x + W_1) = -\frac{dV}{dx} \quad \text{for} \quad -W_1 < x < 0 \quad (1.7)$$

Thus the dipole of charge existing at the junction gives rise to an electric field that varies linearly with distance.

Integration of (1.7) gives

$$V = \frac{qN_A}{\epsilon} \left(\frac{x^2}{2} + W_1 x \right) + C_2 \quad (1.8)$$

If the zero for potential is arbitrarily taken to be the potential of the neutral p -type region, then a second boundary condition is

$$V = 0 \quad \text{for} \quad x = -W_1$$

and use of this in (1.8) gives

$$V = \frac{qN_A}{\epsilon} \left(\frac{x^2}{2} + W_1 x + \frac{W_1^2}{2} \right) \quad \text{for} \quad -W_1 < x < 0 \quad (1.9)$$

At $x = 0$, we define $V = V_1$, and then (1.9) gives

$$V_1 = \frac{qN_A}{\epsilon} \frac{W_1^2}{2} \quad (1.10)$$

If the potential difference from $x = 0$ to $x = W_2$ is V_2 , then it follows that

$$V_2 = \frac{qN_D}{\epsilon} \frac{W_2^2}{2} \quad (1.11)$$

and thus the total voltage across the junction is

$$\psi_0 + V_R = V_1 + V_2 = \frac{q}{2\epsilon} (N_A W_1^2 + N_D W_2^2) \quad (1.12)$$

4 Chapter 1 ■ Models for Integrated-Circuit Active Devices

Substitution of (1.2) in (1.12) gives

$$\psi_0 + V_R = \frac{qW_1^2 N_A}{2\epsilon} \left(1 + \frac{N_A}{N_D} \right) \quad (1.13)$$

From (1.13), the penetration of the depletion layer into the p -type region is

$$W_1 = \left[\frac{2\epsilon(\psi_0 + V_R)}{qN_A \left(1 + \frac{N_A}{N_D} \right)} \right]^{1/2} \quad (1.14)$$

Similarly,

$$W_2 = \left[\frac{2\epsilon(\psi_0 + V_R)}{qN_D \left(1 + \frac{N_D}{N_A} \right)} \right]^{1/2} \quad (1.15)$$

Equations 1.14 and 1.15 show that the depletion regions extend into the p -type and n -type regions in *inverse* relation to the impurity concentrations and in proportion to $\sqrt{\psi_0 + V_R}$. If either N_D or N_A is much larger than the other, the depletion region exists almost entirely in the *lightly doped* region.

■ EXAMPLE

An abrupt pn junction in silicon has doping densities $N_A = 10^{15}$ atoms/cm³ and $N_D = 10^{16}$ atoms/cm³. Calculate the junction built-in potential, the depletion-layer depths, and the maximum field with 10 V reverse bias.

From (1.1)

$$\psi_0 = 26 \ln \frac{10^{15} \times 10^{16}}{2.25 \times 10^{20}} \text{ mV} = 638 \text{ mV} \quad \text{at } 300^\circ\text{K}$$

From (1.14) the depletion-layer depth in the p -type region is

$$W_1 = \left(\frac{2 \times 1.04 \times 10^{-12} \times 10.64}{1.6 \times 10^{-19} \times 10^{15} \times 1.1} \right)^{1/2} = 3.5 \times 10^{-4} \text{ cm} \\ = 3.5 \text{ } \mu\text{m} \quad (\text{where } 1 \text{ } \mu\text{m} = 1 \text{ micrometer} = 10^{-6} \text{ m})$$

The depletion-layer depth in the more heavily doped n -type region is

$$W_2 = \left(\frac{2 \times 1.04 \times 10^{-12} \times 10.64}{1.6 \times 10^{-19} \times 10^{16} \times 11} \right)^{1/2} = 0.35 \times 10^{-4} \text{ cm} = 0.35 \text{ } \mu\text{m}$$

Finally, from (1.7) the maximum field that occurs for $x = 0$ is

$$\mathcal{E}_{\max} = -\frac{qN_A}{\epsilon} W_1 = -1.6 \times 10^{-19} \times \frac{10^{15} \times 3.5 \times 10^{-4}}{1.04 \times 10^{-12}} \\ = -5.4 \times 10^4 \text{ V/cm}$$

■ Note the large magnitude of this electric field.

1.2.1 Depletion-Region Capacitance

Since there is a *voltage-dependent charge* Q associated with the depletion region, we can calculate a small-signal capacitance C_j as follows:

$$C_j = \frac{dQ}{dV_R} = \frac{dQ}{dW_1} \frac{dW_1}{dV_R} \quad (1.16)$$

Now

$$dQ = AqN_A dW_1 \quad (1.17)$$

where A is the cross-sectional area of the junction. Differentiation of (1.14) gives

$$\frac{dW_1}{dV_R} = \left[\frac{\epsilon}{2qN_A \left(1 + \frac{N_A}{N_D}\right) (\psi_0 + V_R)} \right]^{1/2} \quad (1.18)$$

Use of (1.17) and (1.18) in (1.16) gives

$$C_j = A \left[\frac{q\epsilon N_A N_D}{2(N_A + N_D)} \right]^{1/2} \frac{1}{\sqrt{\psi_0 + V_R}} \quad (1.19)$$

The above equation was derived for the case of reverse bias V_R applied to the diode. However, it is valid for positive bias voltages as long as the forward current flow is small. Thus, if V_D represents the bias on the junction (positive for forward bias, negative for reverse bias), then (1.19) can be written as

$$C_j = A \left[\frac{q\epsilon N_A N_D}{2(N_A + N_D)} \right]^{1/2} \frac{1}{\sqrt{\psi_0 - V_D}} \quad (1.20)$$

$$= \frac{C_{j0}}{\sqrt{1 - \frac{V_D}{\psi_0}}} \quad (1.21)$$

where C_{j0} is the value of C_j for $V_D = 0$.

Equations 1.20 and 1.21 were derived using the assumption of constant doping in the p -type and n -type regions. However, many practical diffused junctions more closely approach a *graded* doping profile as shown in Fig. 1.2. In this case, a similar calculation yields

$$C_j = \frac{C_{j0}}{\sqrt[3]{1 - \frac{V_D}{\psi_0}}} \quad (1.22)$$

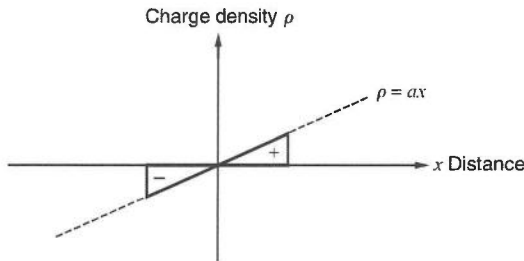


Figure 1.2 Charge density versus distance in a graded junction.

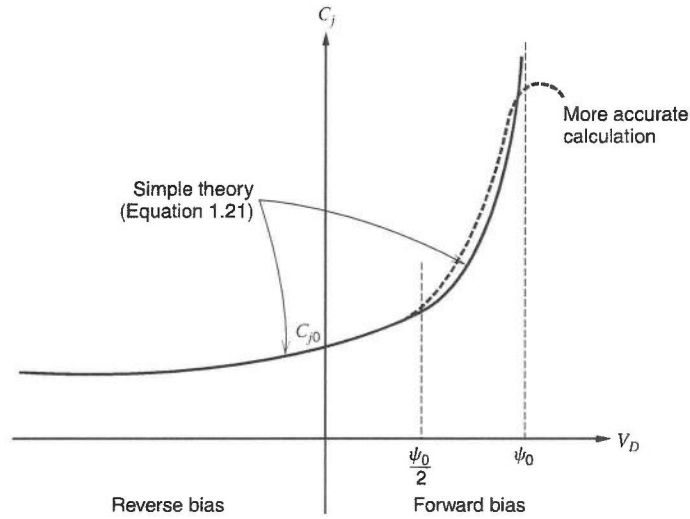


Figure 1.3 Behavior of pn junction depletion-layer capacitance C_j as a function of bias voltage V_D .

Note that both (1.21) and (1.22) predict values of C_j approaching infinity as V_D approaches ψ_0 . However, the current flow in the diode is then appreciable and the equations are no longer valid. A more exact analysis^{2,3} of the behavior of C_j as a function of V_D gives the result shown in Fig. 1.3. For forward bias voltages up to about $\psi_0/2$, the values of C_j predicted by (1.21) are very close to the more accurate value. As an approximation, some computer programs approximate C_j for $V_D > \psi_0/2$ by a linear extrapolation of (1.21) or (1.22).

■ EXAMPLE

If the zero-bias capacitance of a diffused junction is 3 pF and $\psi_0 = 0.5$ V, calculate the capacitance with 10 V reverse bias. Assume the doping profile can be approximated by an abrupt junction.

From (1.21)

$$C_j = \frac{3}{\sqrt{1 + \frac{10}{0.5}}} \text{ pF} = 0.65 \text{ pF}$$

■

1.2.2 Junction Breakdown

From Fig. 1.1c it can be seen that the maximum electric field in the depletion region occurs at the junction, and for an abrupt junction (1.7) yields a value

$$\mathcal{E}_{\max} = -\frac{qN_A}{\epsilon} W_1 \quad (1.23)$$

Substitution of (1.14) in (1.23) gives

$$|\mathcal{E}_{\max}| = \left[\frac{2qN_A N_D V_R}{\epsilon (N_A + N_D)} \right]^{1/2} \quad (1.24)$$

where ψ_0 has been neglected. Equation 1.24 shows that the maximum field increases as the doping density increases and the reverse bias increases. Although useful for indicating the

functional dependence of $\mathcal{E}_{\max}$ on other variables, this equation is strictly valid for an ideal plane junction only. Practical junctions tend to have edge effects that cause somewhat higher values of $\mathcal{E}_{\max}$ due to a concentration of the field at the curved edges of the junction.

Any reverse-biased *pn* junction has a small reverse current flow due to the presence of minority-carrier holes and electrons in the vicinity of the depletion region. These are swept across the depletion region by the field and contribute to the leakage current of the junction. As the reverse bias on the junction is increased, the maximum field increases and the carriers acquire increasing amounts of energy between lattice collisions in the depletion region. At a critical field $\mathcal{E}_{\text{crit}}$ the carriers traversing the depletion region acquire sufficient energy to create new hole-electron pairs in collisions with silicon atoms. This is called the *avalanche process* and leads to a sudden increase in the reverse-bias leakage current since the newly created carriers are also capable of producing avalanche. The value of $\mathcal{E}_{\text{crit}}$ is about 3×10^5 V/cm for junction doping densities in the range of 10^{15} to 10^{16} atoms/cm³, but it increases slowly as the doping density increases and reaches about 10^6 V/cm for doping densities of 10^{18} atoms/cm³.

A typical *I-V* characteristic for a junction diode is shown in Fig. 1.4, and the effect of avalanche breakdown is seen by the large increase in reverse current, which occurs as the reverse bias approaches the breakdown voltage *BV*. This corresponds to the maximum field $\mathcal{E}_{\max}$ approaching $\mathcal{E}_{\text{crit}}$. It has been found empirically⁴ that if the normal reverse bias current of the diode is I_R with no avalanche effect, then the actual reverse current near the breakdown voltage is

$$I_{RA} = MI_R \quad (1.25)$$

where *M* is the *multiplication factor* defined by

$$M = \frac{1}{1 - \left(\frac{V_R}{BV}\right)^n} \quad (1.26)$$

In this equation, V_R is the reverse bias on the diode and *n* has a value between 3 and 6.

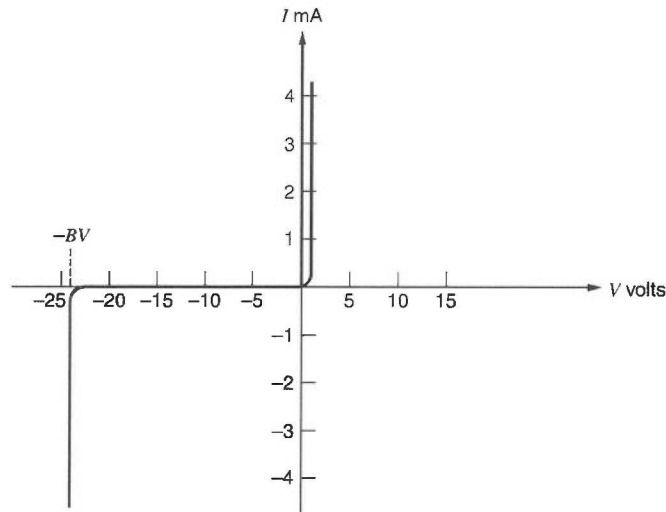


Figure 1.4 Typical *I-V* characteristic of a junction diode showing avalanche breakdown.

The operation of a *pn* junction in the breakdown region is not inherently destructive. However, the avalanche current flow must be limited by external resistors in order to prevent excessive power dissipation from occurring at the junction and causing damage to the device. Diodes operated in the avalanche region are widely used as voltage references and are called *Zener diodes*. There is another, related process called *Zener breakdown*,⁵ which is different from the avalanche breakdown described above. Zener breakdown occurs only in very heavily doped junctions where the electric field becomes large enough (even with small reverse-bias voltages) to strip electrons away from the valence bonds. This process is called *tunneling*, and there is no multiplication effect as in avalanche breakdown. Although the Zener breakdown mechanism is important only for breakdown voltages below about 6 V, all breakdown diodes are commonly referred to as Zener diodes.

The calculations so far have been concerned with the breakdown characteristic of plane abrupt junctions. Practical diffused junctions differ in some respects from these results and the characteristics of these junctions have been calculated and tabulated for use by designers.⁵ In particular, edge effects in practical diffused junctions can result in breakdown voltages as much as 50 percent below the value calculated for a plane junction.

■ EXAMPLE

An abrupt plane *pn* junction has doping densities $N_A = 5 \times 10^{15}$ atoms/cm³ and $N_D = 10^{16}$ atoms/cm³. Calculate the breakdown voltage if $\mathcal{E}_{\text{crit}} = 3 \times 10^5$ V/cm.

The breakdown voltage is calculated using $\mathcal{E}_{\text{max}} = \mathcal{E}_{\text{crit}}$ in (1.24) to give

$$\begin{aligned} BV &= \frac{\epsilon (N_A + N_D)}{2qN_A N_D} \mathcal{E}_{\text{crit}}^2 \\ &= \frac{1.04 \times 10^{-12} \times 15 \times 10^{15}}{2 \times 1.6 \times 10^{-19} \times 5 \times 10^{15} \times 10^{16}} \times 9 \times 10^{10} \text{ V} \\ &= 88 \text{ V} \end{aligned}$$

■

1.3 Large-Signal Behavior of Bipolar Transistors

In this section, the large-signal or dc behavior of bipolar transistors is considered. Large-signal models are developed for the calculation of total currents and voltages in transistor circuits, and such effects as breakdown voltage limitations, which are usually not included in models, are also considered. Second-order effects, such as current-gain variation with collector current and Early voltage, can be important in many circuits and are treated in detail.

The sign conventions used for bipolar transistor currents and voltages are shown in Fig. 1.5. All bias currents for both *npn* and *pnp* transistors are assumed positive going into the device.

1.3.1 Large-Signal Models in the Forward-Active Region

A typical *npn* planar bipolar transistor structure is shown in Fig. 1.6a, where collector, base, and emitter are labeled C, B, and E, respectively. The method of fabricating such transistor structures is described in Chapter 2. It is shown there that the impurity doping density in the base and the emitter of such a transistor is not constant but varies with distance from the top surface. However, many of the characteristics of such a device can be predicted by analyzing the idealized transistor structure shown in Fig. 1.6b. In this structure, the base and emitter

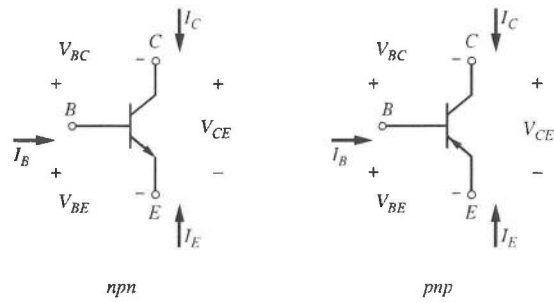


Figure 1.5 Bipolar transistor sign convention.

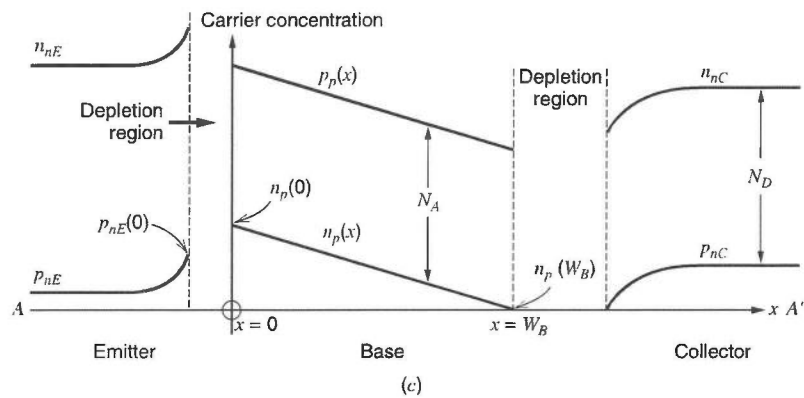
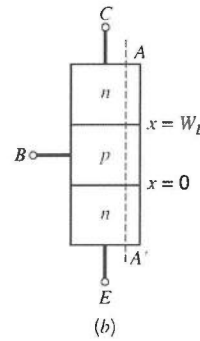
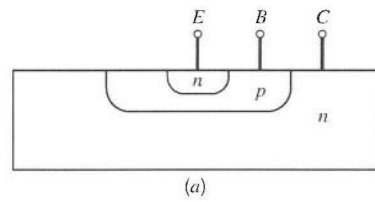


Figure 1.6 (a) Cross section of a typical *npn* planar bipolar transistor structure. (b) Idealized transistor structure. (c) Carrier concentrations along the cross section AA' of the transistor in (b). Uniform doping densities are assumed. (Not to scale.)

doping densities are assumed constant, and this is sometimes called a *uniform-base* transistor. Where possible in the following analyses, the equations for the uniform-base analysis are expressed in a form that applies also to nonuniform-base transistors.

A cross section AA' is taken through the device of Fig. 1.6b and carrier concentrations along this section are plotted in Fig. 1.6c. Hole concentrations are denoted by p and electron concentrations by n with subscripts p or n representing p -type or n -type regions. The n -type emitter and collector regions are distinguished by subscripts E and C , respectively. The carrier concentrations shown in Fig. 1.6c apply to a device in the *forward-active region*. That is, the base-emitter junction is forward biased and the base-collector junction is reverse biased. The minority-carrier concentrations in the base at the edges of the depletion regions can be calculated from a Boltzmann approximation to the Fermi-Dirac distribution function to give⁶

$$n_p(0) = n_{po} \exp \frac{V_{BE}}{V_T} \quad (1.27)$$

$$n_p(W_B) = n_{po} \exp \frac{V_{BC}}{V_T} \simeq 0 \quad (1.28)$$

where W_B is the width of the base from the base-emitter depletion layer edge to the base-collector depletion layer edge and n_{po} is the equilibrium concentration of electrons in the base. Note that V_{BC} is negative for an nnp transistor in the forward-active region and thus $n_p(W_B)$ is very small. Low-level injection conditions are assumed in the derivation of (1.27) and (1.28). This means that the minority-carrier concentrations are always assumed much smaller than the majority-carrier concentration.

If *recombination* of holes and electrons in the base is small, it can be shown that⁷ the minority-carrier concentration $n_p(x)$ in the base varies *linearly* with distance. Thus a straight line can be drawn joining the concentrations at $x = 0$ and $x = W_B$ in Fig. 1.6c.

For charge neutrality in the base, it is necessary that

$$N_A + n_p(x) = p_p(x) \quad (1.29)$$

and thus

$$p_p(x) - n_p(x) = N_A \quad (1.30)$$

where $p_p(x)$ is the hole concentration in the base and N_A is the base doping density that is assumed constant. Equation 1.30 indicates that the hole and electron concentrations are separated by a constant amount and thus $p_p(x)$ also varies linearly with distance.

Collector current is produced by minority-carrier electrons in the base diffusing in the direction of the concentration gradient and being swept across the collector-base depletion region by the field existing there. The diffusion current density due to electrons in the base is

$$J_n = qD_n \frac{dn_p(x)}{dx} \quad (1.31)$$

where D_n is the diffusion constant for electrons. From Fig. 1.6c

$$J_n = -qD_n \frac{n_p(0)}{W_B} \quad (1.32)$$

If I_C is the collector current and is taken as positive flowing *into* the collector, it follows from (1.32) that

$$I_C = qAD_n \frac{n_p(0)}{W_B} \quad (1.33)$$

where A is the cross-sectional area of the emitter. Substitution of (1.27) into (1.33) gives

$$I_C = \frac{qAD_n n_{po}}{W_B} \exp \frac{V_{BE}}{V_T} \quad (1.34)$$

$$= I_S \exp \frac{V_{BE}}{V_T} \quad (1.35)$$

where

$$I_S = \frac{qAD_n n_{po}}{W_B} \quad (1.36)$$

and I_S is a constant used to describe the transfer characteristic of the transistor in the forward-active region. Equation 1.36 can be expressed in terms of the base doping density by noting that⁸ (see Chapter 2)

$$n_{po} = \frac{n_i^2}{N_A} \quad (1.37)$$

and substitution of (1.37) in (1.36) gives

$$I_S = \frac{qAD_n n_i^2}{W_B N_A} = \frac{qA \bar{D}_n n_i^2}{Q_B} \quad (1.38)$$

where $Q_B = W_B N_A$ is the number of doping atoms in the base per unit area of the emitter and n_i is the intrinsic carrier concentration in silicon. In this form (1.38) applies to both uniform- and nonuniform-base transistors and D_n has been replaced by $\bar{D}_n$, which is an average effective value for the electron diffusion constant in the base. This is necessary for nonuniform-base devices because the diffusion constant is a function of impurity concentration. Typical values of I_S as given by (1.38) are from 10^{-14} to 10^{-16} A.

Equation 1.35 gives the collector current as a function of base-emitter voltage. The base current I_B is also an important parameter and, at moderate current levels, consists of two major components. One of these (I_{B1}) represents recombination of holes and electrons in the base and is proportional to the minority-carrier charge Q_e in the base. From Fig. 1.6c, the minority-carrier charge in the base is

$$Q_e = \frac{1}{2} n_p(0) W_B q A \quad (1.39)$$

and we have

$$I_{B1} = \frac{Q_e}{\tau_b} = \frac{1}{2} \frac{n_p(0) W_B q A}{\tau_b} \quad (1.40)$$

where τ_b is the minority-carrier lifetime in the base. I_{B1} represents a flow of majority holes from the base lead into the base region. Substitution of (1.27) in (1.40) gives

$$I_{B1} = \frac{1}{2} \frac{n_{po} W_B q A}{\tau_b} \exp \frac{V_{BE}}{V_T} \quad (1.41)$$

The second major component of base current (usually the dominant one in integrated-circuit npn devices) is due to injection of holes from the base into the emitter. This current component depends on the gradient of minority-carrier holes in the emitter and is⁹

$$I_{B2} = \frac{qAD_p}{L_p} p_{nE}(0) \quad (1.42)$$

where D_p is the diffusion constant for holes and L_p is the diffusion length (assumed small) for holes in the emitter. $p_{nE}(0)$ is the concentration of holes in the emitter at the edge of the

depletion region and is

$$p_{nE}(0) = p_{nEo} \exp \frac{V_{BE}}{V_T} \quad (1.43)$$

If N_D is the donor atom concentration in the emitter (assumed constant), then

$$p_{nEo} \simeq \frac{n_i^2}{N_D} \quad (1.44)$$

The emitter is deliberately doped much more heavily than the base, making N_D large and p_{nEo} small, so that the base-current component, I_{B2} , is minimized.

Substitution of (1.43) and (1.44) in (1.42) gives

$$I_{B2} = \frac{qAD_p}{L_p} \frac{n_i^2}{N_D} \exp \frac{V_{BE}}{V_T} \quad (1.45)$$

The total base current, I_B , is the sum of I_{B1} and I_{B2} :

$$I_B = I_{B1} + I_{B2} = \left(\frac{1}{2} \frac{n_{po} W_B q A}{\tau_b} + \frac{qAD_p}{L_p} \frac{n_i^2}{N_D} \right) \exp \frac{V_{BE}}{V_T} \quad (1.46)$$

Although this equation was derived assuming uniform base and emitter doping, it gives the correct functional dependence of I_B on device parameters for practical double-diffused nonuniform-base devices. Second-order components of I_B , which are important at low current levels, are considered later.

Since I_C in (1.35) and I_B in (1.46) are both proportional to $\exp(V_{BE}/V_T)$ in this analysis, the base current can be expressed in terms of collector current as

$$I_B = \frac{I_C}{\beta_F} \quad (1.47)$$

where β_F is the forward current gain. An expression for β_F can be calculated by substituting (1.34) and (1.46) in (1.47) to give

$$\beta_F = \frac{\frac{qAD_n n_{po}}{W_B}}{\frac{1}{2} \frac{n_{po} W_B q A}{\tau_b} + \frac{qAD_p n_i^2}{L_p N_D}} = \frac{1}{\frac{W_B^2}{2\tau_b D_n} + \frac{D_p}{D_n} \frac{W_B}{L_p} \frac{N_A}{N_D}} \quad (1.48)$$

where (1.37) has been substituted for n_{po} . Equation 1.48 shows that β_F is maximized by minimizing the base width W_B and maximizing the ratio of emitter to base doping densities N_D/N_A . Typical values of β_F for npn transistors in integrated circuits are 50 to 500, whereas lateral pnp transistors (to be described in Chapter 2) have values 10 to 100. Finally, the emitter current is

$$I_E = -(I_C + I_B) = -\left(I_C + \frac{I_C}{\beta_F}\right) = -\frac{I_C}{\alpha_F} \quad (1.49)$$

where

$$\alpha_F = \frac{\beta_F}{1 + \beta_F} \quad (1.50)$$

The value of α_F can be expressed in terms of device parameters by substituting (1.48) in (1.50) to obtain

$$\alpha_F = \frac{1}{1 + \frac{1}{\beta_F}} = \frac{1}{1 + \frac{W_B^2}{2\tau_b D_n} + \frac{D_p W_B N_A}{D_n L_p N_D}} \simeq \alpha_T \gamma \quad (1.51)$$

where

$$\alpha_T = \frac{1}{1 + \frac{W_B^2}{2\tau_b D_n}} \quad (1.51a)$$

$$\gamma = \frac{1}{1 + \frac{D_p W_B N_A}{D_n L_p N_D}} \quad (1.51b)$$

The validity of (1.51) depends on $W_B^2/2\tau_b D_n \ll 1$ and $(D_p/D_n)(W_B/L_p)(N_A/N_D) \ll 1$, and this is always true if β_F is large [see (1.48)]. The term γ in (1.51) is called the *emitter injection efficiency* and is equal to the ratio of the electron current (*npn* transistor) injected into the base from the emitter to the total hole and electron current crossing the base-emitter junction. Ideally $\gamma \rightarrow 1$, and this is achieved by making N_D/N_A large and W_B small. In that case very little reverse injection occurs from base to emitter.

The term α_T in (1.51) is called the *base transport factor* and represents the fraction of carriers injected into the base (from the emitter) that reach the collector. Ideally $\alpha_T \rightarrow 1$ and this is achieved by making W_B small. It is evident from the above development that fabrication changes that cause α_T and γ to approach unity also maximize the value of β_F of the transistor.

The results derived above allow formulation of a large-signal model of the transistor suitable for bias-circuit calculations with devices in the forward-active region. One such circuit is shown in Fig. 1.7 and consists of a base-emitter diode to model (1.46) and a controlled collector-current generator to model (1.47). Note that the collector voltage ideally has no influence on the collector current and the collector node acts as a high-impedance current source. A simpler version of this equivalent circuit, which is often useful, is shown in Fig. 1.7b, where the input diode has been replaced by a battery with a value $V_{BE(on)}$, which is usually 0.6 to 0.7 V. This represents the fact that in the forward-active region the base-emitter voltage varies very little because of the steep slope of the exponential characteristic. In some circuits the temperature coefficient of $V_{BE(on)}$ is important, and a typical value for this is $-2 \text{ mV}/^\circ\text{C}$. The equivalent circuits of Fig. 1.7 apply for *npn* transistors. For *pnp* devices the corresponding equivalent circuits are shown in Fig. 1.8.

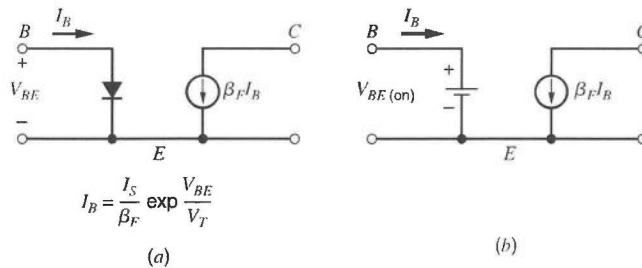


Figure 1.7 Large-signal models of *npn* transistors for use in bias calculations. (a) Circuit incorporating an input diode. (b) Simplified circuit with an input voltage source.

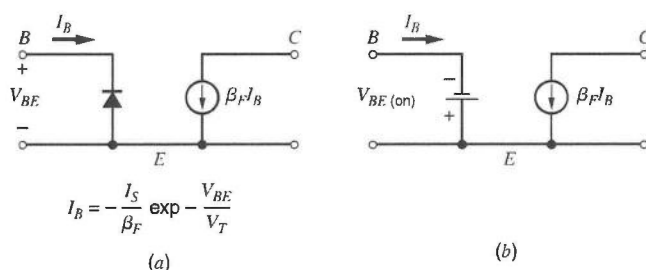


Figure 1.8 Large-signal models of *pnp* transistors corresponding to the circuits of Fig. 1.7.

1.3.2 Effects of Collector Voltage on Large-Signal Characteristics in the Forward-Active Region

In the analysis of the previous section, the collector-base junction was assumed reverse biased and ideally had no effect on the collector currents. This is a useful approximation for first-order calculations, but is not strictly true in practice. There are occasions when the influence of collector voltage on collector current is important, and this will now be investigated.

The collector voltage has a dramatic effect on the collector current in two regions of device operation. These are the saturation (V_{CE} approaches zero) and breakdown (V_{CE} very large) regions that will be considered later. For values of collector-emitter voltage V_{CE} between these extremes, the collector current increases slowly as V_{CE} increases. The reason for this can be seen from Fig. 1.9, which is a sketch of the minority-carrier concentration in the base of the transistor. Consider the effect of changes in V_{CE} on the carrier concentration for constant V_{BE} . Since V_{BE} is constant, the change in V_{CB} equals the change in V_{CE} and this causes an increase in the collector-base depletion-layer width as shown. The change in the base width of the transistor, ΔW_B , equals the change in the depletion-layer width and causes an increase ΔI_C in the collector current.

From (1.35) and (1.38) we have

$$I_C = \frac{qA\bar{D}_n n_i^2}{Q_B} \exp \frac{V_{BE}}{V_T} \quad (1.52)$$

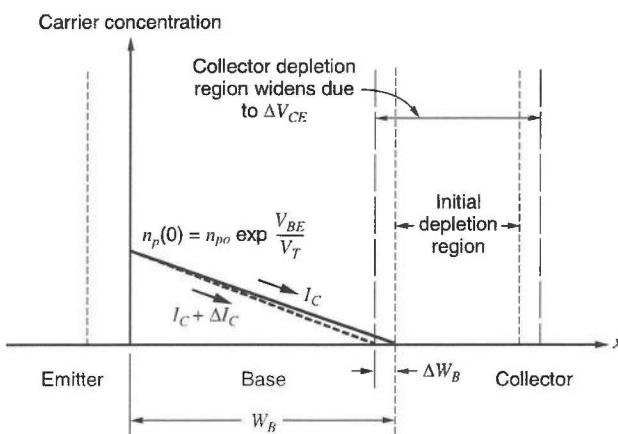


Figure 1.9 Effect of increases in V_{CE} on the collector depletion region and base width of a bipolar transistor.

Differentiation of (1.52) yields

$$\frac{\partial I_C}{\partial V_{CE}} = -\frac{qA\bar{D}_n n_i^2}{Q_B^2} \left(\exp \frac{V_{BE}}{V_T} \right) \frac{dQ_B}{dV_{CE}} \quad (1.53)$$

and substitution of (1.52) in (1.53) gives

$$\frac{\partial I_C}{\partial V_{CE}} = -\frac{I_C}{Q_B} \frac{dQ_B}{dV_{CE}} \quad (1.54)$$

For a uniform-base transistor $Q_B = W_B N_A$, and (1.54) becomes

$$\frac{\partial I_C}{\partial V_{CE}} = -\frac{I_C}{W_B} \frac{dW_B}{dV_{CE}} \quad (1.55)$$

Note that since the base width *decreases* as V_{CE} increases, dW_B/dV_{CE} in (1.55) is negative and thus $\partial I_C/\partial V_{CE}$ is positive. The magnitude of dW_B/dV_{CE} can be calculated from (1.18) for a uniform-base transistor. This equation predicts that dW_B/dV_{CE} is a function of the bias value of V_{CE} , but the variation is typically small for a reverse-biased junction and dW_B/dV_{CE} is often assumed constant. The resulting predictions agree adequately with experimental results.

Equation 1.55 shows that $\partial I_C/\partial V_{CE}$ is proportional to the collector-bias current and inversely proportional to the transistor base width. Thus narrow-base transistors show a greater dependence of I_C on V_{CE} in the forward-active region. The dependence of $\partial I_C/\partial V_{CE}$ on I_C results in typical transistor output characteristics as shown in Fig. 1.10. In accordance with the assumptions made in the foregoing analysis, these characteristics are shown for constant values of V_{BE} . However, in most integrated-circuit transistors the base current is dependent only on V_{BE} and not on V_{CE} , and thus constant-base-current characteristics can often be used in the following calculation. The reason for this is that the base current is usually dominated by the I_{B2} component of (1.45), which has no dependence on V_{CE} . Extrapolation of the characteristics of Fig. 1.10 back to the V_{CE} axis gives an intercept V_A called the Early voltage, where

$$V_A = \frac{I_C}{\frac{\partial I_C}{\partial V_{CE}}} \quad (1.56)$$

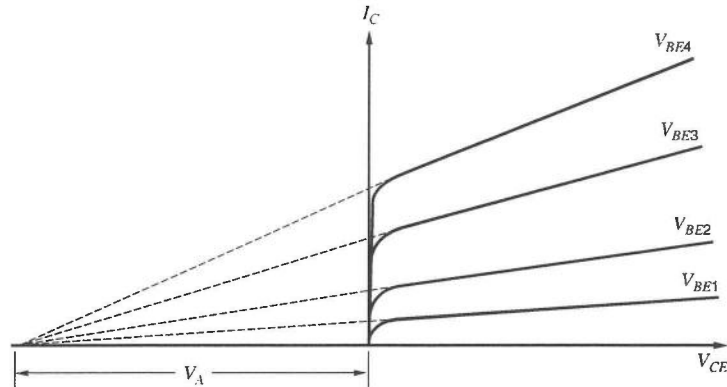


Figure 1.10 Bipolar transistor output characteristics showing the Early voltage, V_A .

Substitution of (1.55) in (1.56) gives

$$V_A = -W_B \frac{dV_{CE}}{dW_B} \quad (1.57)$$

which is a constant, independent of I_C . Thus all the characteristics extrapolate to the same point on the V_{CE} axis. The variation of I_C with V_{CE} is called the Early effect, and V_A is a common model parameter for circuit-analysis computer programs. Typical values of V_A for integrated-circuit transistors are 15 to 100 V. The inclusion of Early effect in dc bias calculations is usually limited to computer analysis because of the complexity introduced into the calculation. However, the influence of the Early effect is often dominant in small-signal calculations for high-gain circuits and this point will be considered later.

Finally, the influence of Early effect on the transistor large-signal characteristics in the forward-active region can be represented approximately by modifying (1.35) to

$$I_C = I_S \left(1 + \frac{V_{CE}}{V_A} \right) \exp \frac{V_{BE}}{V_T} \quad (1.58)$$

This is a common means of representing the device output characteristics for computer simulation.

1.3.3 Saturation and Inverse-Active Regions

Saturation is a region of device operation that is usually avoided in analog circuits because the transistor gain is very low in this region. Saturation is much more commonly encountered in digital circuits, where it provides a well-specified output voltage that represents a logic state.

In saturation, both emitter-base and collector-base junctions are forward biased. Consequently, the collector-emitter voltage V_{CE} is quite small and is usually in the range 0.05 to 0.3 V. The carrier concentrations in a saturated npn transistor with uniform base doping are shown in Fig. 1.11. The minority-carrier concentration in the base at the edge of the depletion region is again given by (1.28) as

$$n_p(W_B) = n_{p0} \exp \frac{V_{BC}}{V_T} \quad (1.59)$$

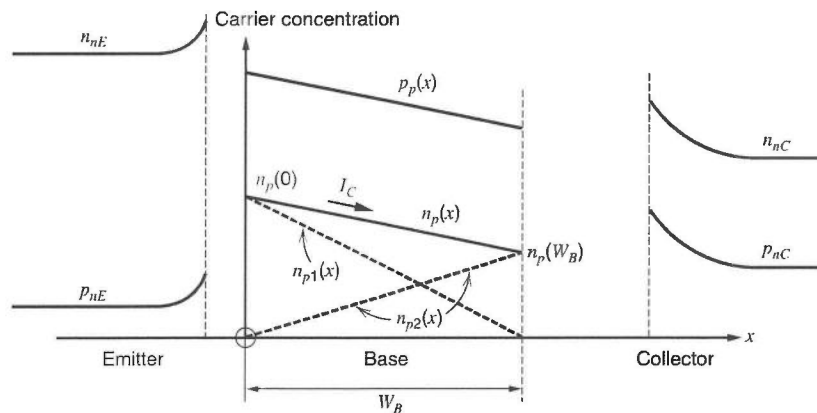


Figure 1.11 Carrier concentrations in a saturated npn transistor. (Not to scale.)

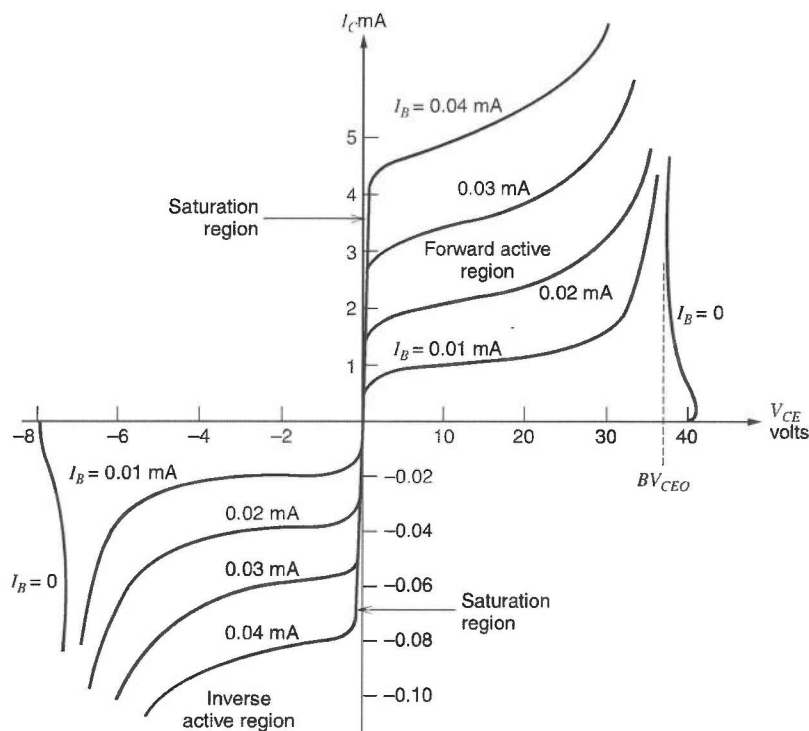


Figure 1.12 Typical I_C - V_{CE} characteristics for an *npn* bipolar transistor. Note the different scales for positive and negative currents and voltages.

but since V_{BC} is now positive, the value of $n_p(W_B)$ is no longer negligible. Consequently, changes in V_{CE} with V_{BE} held constant (which cause equal changes in V_{BC}) directly affect $n_p(W_B)$. Since the collector current is proportional to the slope of the minority-carrier concentration in the base [see (1.31)], it is also proportional to $[n_p(0) - n_p(W_B)]$ from Fig. 1.11. Thus changes in $n_p(W_B)$ directly affect the collector current, and the collector node of the transistor appears to have a *low impedance*. As V_{CE} is decreased in saturation with V_{BE} held constant, V_{BC} increases, as does $n_p(W_B)$ from (1.59). Thus, from Fig. 1.11, the collector current decreases because the slope of the carrier concentration decreases. This gives rise to the saturation region of the $I_C - V_{CE}$ characteristic shown in Fig. 1.12. The slope of the $I_C - V_{CE}$ characteristic in this region is largely determined by the resistance in series with the collector lead due to the finite resistivity of the *n*-type collector material. A useful model for the transistor in this region is shown in Fig. 1.13 and consists of a fixed voltage source to represent $V_{BE(on)}$, and a fixed voltage source to represent the collector-emitter voltage $V_{CE(sat)}$. A more accurate but more complex model includes a resistor in series with the collector. This resistor can have a value ranging from 20 to 500 Ω , depending on the device structure.

An additional aspect of transistor behavior in the saturation region is apparent from Fig. 1.11. For a given collector current, there is now a much larger amount of stored charge in the base than there is in the forward-active region. Thus the base-current contribution represented by (1.41) will be larger in saturation. In addition, since the collector-base junction is now forward biased, there is a new base-current component due to injection of carriers from the

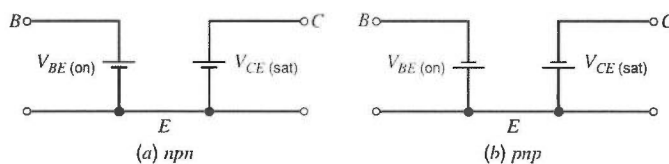


Figure 1.13 Large-signal models for bipolar transistors in the saturation region.

base to the collector. These two effects result in a base current I_B in saturation, which is larger than in the forward-active region for a given collector current I_C . Ratio I_C/I_B in saturation is often referred to as the *forced* β and is always less than β_F . As the forced β is made lower with respect to β_F , the device is said to be more *heavily saturated*.

The minority-carrier concentration in saturation shown in Fig. 1.11 is a straight line joining the two end points, assuming that recombination is small. This can be represented as a linear superposition of the two dotted distributions as shown. The justification for this is that the terminal currents depend *linearly* on the concentrations $n_p(0)$ and $n_p(W_B)$. This picture of device carrier concentrations can be used to derive some general equations describing transistor behavior. Each of the distributions in Fig. 1.11 is considered separately and the two contributions are combined. The *emitter* current that would result from $n_{p1}(x)$ above is given by the classical diode equation

$$I_{EF} = -I_{ES} \left(\exp \frac{V_{BE}}{V_T} - 1 \right) \quad (1.60)$$

where I_{ES} is a constant that is often referred to as the *saturation current* of the junction (no connection with the transistor saturation previously described). Equation 1.60 predicts that the junction current is given by $I_{EF} \simeq I_{ES}$ with a reverse-bias voltage applied. However, in practice (1.60) is applicable only in the forward-bias region, since second-order effects dominate under reverse-bias conditions and typically result in a junction current several orders of magnitude larger than I_{ES} . The junction current that flows under reverse-bias conditions is often called the *leakage current* of the junction.

Returning to Fig. 1.11, we can describe the *collector* current resulting from $n_{p2}(x)$ alone as

$$I_{CR} = -I_{CS} \left(\exp \frac{V_{BC}}{V_T} - 1 \right) \quad (1.61)$$

where I_{CS} is a constant. The total collector current I_C is given by I_{CR} plus the fraction of I_{EF} that reaches the collector (allowing for recombination and reverse emitter injection). Thus

$$I_C = \alpha_F I_{ES} \left(\exp \frac{V_{BE}}{V_T} - 1 \right) - I_{CS} \left(\exp \frac{V_{BC}}{V_T} - 1 \right) \quad (1.62)$$

where α_F has been defined previously by (1.51). Similarly, the total emitter current is composed of I_{EF} plus the fraction of I_{CR} that reaches the emitter with the transistor acting in an inverted mode. Thus

$$I_E = -I_{ES} \left(\exp \frac{V_{BE}}{V_T} - 1 \right) + \alpha_R I_{CS} \left(\exp \frac{V_{BC}}{V_T} - 1 \right) \quad (1.63)$$

where α_R is the ratio of emitter to collector current with the transistor operating *inverted* (i.e., with the collector-base junction forward biased and emitting carriers into the base and the emitter-base junction reverse biased and collecting carriers). Typical values of α_R are 0.5 to 0.8.

An inverse current gain β_R is also defined

$$\beta_R = \frac{\alpha_R}{1 - \alpha_R} \quad (1.64)$$

and has typical values 1 to 5. This is the current gain of the transistor when operated inverted and is much lower than β_F because the device geometry and doping densities are designed to maximize β_F . The inverse-active region of device operation occurs for V_{CE} negative in an npn transistor and is shown in Fig. 1.12. In order to display these characteristics adequately in the same figure as the forward-active region, the negative voltage and current scales have been expanded. The inverse-active mode of operation is rarely encountered in analog circuits.

Equations 1.62 and 1.63 describe npn transistor operation in the saturation region when V_{BE} and V_{BC} are both positive, and also in the forward-active and inverse-active regions. These equations are the *Ebers-Moll* equations. In the forward-active region, they degenerate into a form similar to that of (1.35), (1.47), and (1.49) derived earlier. This can be shown by putting V_{BE} positive and V_{BC} negative in (1.62) and (1.63) to obtain

$$I_C = \alpha_F I_{ES} \left(\exp \frac{V_{BE}}{V_T} - 1 \right) + I_{CS} \quad (1.65)$$

$$I_E = -I_{ES} \left(\exp \frac{V_{BE}}{V_T} - 1 \right) - \alpha_R I_{CS} \quad (1.66)$$

Equation 1.65 is similar in form to (1.35) except that leakage currents that were previously neglected have now been included. This minor difference is significant only at high temperatures or very low operating currents. Comparison of (1.65) with (1.35) allows us to identify $I_S = \alpha_F I_{ES}$, and it can be shown¹⁰ in general that

$$\alpha_F I_{ES} = \alpha_R I_{CS} = I_S \quad (1.67)$$

where this expression represents a reciprocity condition. Use of (1.67) in (1.62) and (1.63) allows the Ebers-Moll equations to be expressed in the general form

$$I_C = I_S \left(\exp \frac{V_{BE}}{V_T} - 1 \right) - \frac{I_S}{\alpha_R} \left(\exp \frac{V_{BC}}{V_T} - 1 \right) \quad (1.62a)$$

$$I_E = -\frac{I_S}{\alpha_F} \left(\exp \frac{V_{BE}}{V_T} - 1 \right) + I_S \left(\exp \frac{V_{BC}}{V_T} - 1 \right) \quad (1.63a)$$

This form is often used for computer representation of transistor large-signal behavior.

The effect of leakage currents mentioned above can be further illustrated as follows. In the forward-active region, from (1.66)

$$I_{ES} \left(\exp \frac{V_{BE}}{V_T} - 1 \right) = -I_E - \alpha_R I_{CS} \quad (1.68)$$

Substitution of (1.68) in (1.65) gives

$$I_C = -\alpha_F I_E + I_{CO} \quad (1.69)$$

where

$$I_{CO} = I_{CS}(1 - \alpha_R \alpha_F) \quad (1.69a)$$

and I_{CO} is the collector-base leakage current with the emitter open. Although I_{CO} is given theoretically by (1.69a), in practice, surface leakage effects dominate when the collector-base junction is reverse biased and I_{CO} is typically several orders of magnitude larger than the value

given by (1.69a). However, (1.69) is still valid if the appropriate measured value for I_{CO} is used. Typical values of I_{CO} are from 10^{-10} to 10^{-12} A at 25°C , and the magnitude doubles about every 8°C . As a consequence, these leakage terms can become very significant at high temperatures. For example, consider the base current I_B . From Fig. 1.5 this is

$$I_B = -(I_C + I_E) \quad (1.70)$$

If I_E is calculated from (1.69) and substituted in (1.70), the result is

$$I_B = \frac{1 - \alpha_F}{\alpha_F} I_C - \frac{I_{CO}}{\alpha_F} \quad (1.71)$$

But from (1.50)

$$\beta_F = \frac{\alpha_F}{1 - \alpha_F} \quad (1.72)$$

and use of (1.72) in (1.71) gives

$$I_B = \frac{I_C}{\beta_F} - \frac{I_{CO}}{\alpha_F} \quad (1.73)$$

Since the two terms in (1.73) have opposite signs, the effect of I_{CO} is to *decrease* the magnitude of the external base current at a given value of collector current.

■ EXAMPLE

If I_{CO} is 10^{-10} A at 24°C , estimate its value at 120°C .

Assuming that I_{CO} doubles every 8°C , we have

$$\begin{aligned} I_{CO}(120^\circ\text{C}) &= 10^{-10} \times 2^{12} \\ &= 0.4 \mu\text{A} \end{aligned}$$

■

1.3.4 Transistor Breakdown Voltages

In Section 1.2.2 the mechanism of avalanche breakdown in a *pn* junction was described. Similar effects occur at the base-emitter and base-collector junctions of a transistor and these effects limit the maximum voltages that can be applied to the device.

First consider a transistor in the common-base configuration shown in Fig. 1.14a and supplied with a constant emitter current. Typical $I_C - V_{CB}$ characteristics for an *npn* transistor in such a connection are shown in Fig. 1.14b. For $I_E = 0$ the collector-base junction breaks down at a voltage BV_{CBO} , which represents collector-base breakdown with the emitter open. For finite values of I_E , the effects of avalanche multiplication are apparent for values of V_{CB} below BV_{CBO} . In the example shown, the effective common-base current gain $\alpha_F = I_C/I_E$ becomes larger than unity for values of V_{CB} above about 60 V. Operation in this region (but below BV_{CBO}) can, however, be safely undertaken if the device power dissipation is not excessive. The considerations of Section 1.2.2 apply to this situation, and neglecting leakage currents, we can calculate the collector current in Fig. 1.14a as

$$I_C = -\alpha_F I_E M \quad (1.74)$$

where M is defined by (1.26) and thus

$$I_C = -\alpha_F I_E \frac{1}{1 - \left(\frac{V_{CB}}{BV_{CBO}} \right)^n} \quad (1.75)$$

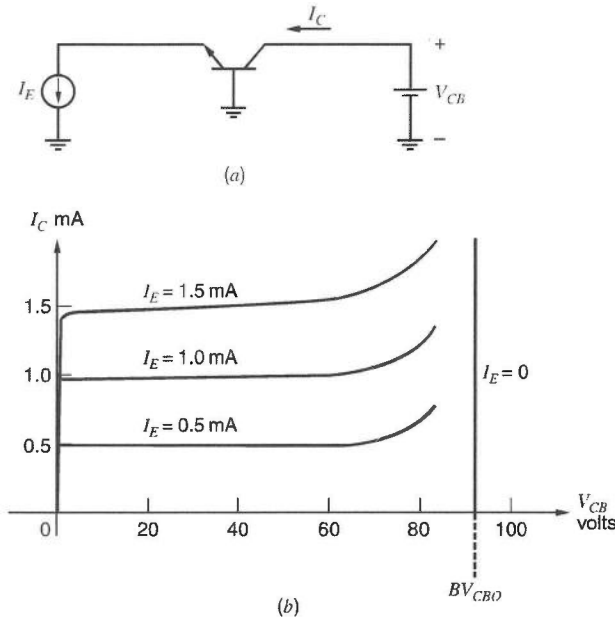


Figure 1.14 Common-base transistor connection. (a) Test circuit. (b) $I_C - V_{CB}$ characteristics.

One further point to note about the common-base characteristics of Fig. 1.14b is that for low values of V_{CB} where avalanche effects are negligible, the curves show very little of the Early effect seen in the common-emitter characteristics. Base widening still occurs in this configuration as V_{CB} is increased, but unlike the common-emitter connection, it produces little change in I_C . This is because I_E is now fixed instead of V_{BE} or I_B , and in Fig. 1.9, this means the slope of the minority-carrier concentration at the emitter edge of the base is fixed. Thus the collector current remains almost unchanged.

Now consider the effect of avalanche breakdown on the common-emitter characteristics of the device. Typical characteristics are shown in Fig. 1.12, and breakdown occurs at a value BV_{CEO} , which is sometimes called the sustaining voltage LV_{CEO} . As in previous cases, operation near the breakdown voltage is destructive to the device only if the current (and thus the power dissipation) becomes excessive.

The effects of avalanche breakdown on the common-emitter characteristics are more complex than in the common-base configuration. This is because hole-electron *pairs* are produced by the avalanche process and the holes are swept into the base, where they effectively contribute to the base current. In a sense, the avalanche current is then *amplified* by the transistor. The base current is still given by

$$I_B = -(I_C + I_E) \quad (1.76)$$

Equation 1.74 still holds, and substitution of this in (1.76) gives

$$I_C = \frac{M\alpha_F}{1 - M\alpha_F} I_B \quad (1.77)$$

where

$$M = \frac{1}{1 - \left(\frac{V_{CB}}{BV_{CBO}} \right)^n} \quad (1.78)$$

Equation 1.77 shows that I_C approaches infinity as $M\alpha_F$ approaches unity. That is, the effective β approaches infinity because of the additional base-current contribution from the avalanche process itself. The value of BV_{CEO} can be determined by solving

$$M\alpha_F = 1 \quad (1.79)$$

If we assume that $V_{CB} \simeq V_{CE}$, this gives

$$\frac{\alpha_F}{1 - \left(\frac{BV_{CEO}}{BV_{CBO}} \right)^n} = 1 \quad (1.80)$$

and this results in

$$\frac{BV_{CEO}}{BV_{CBO}} = \sqrt[n]{1 - \alpha_F}$$

and thus

$$BV_{CEO} \simeq \frac{BV_{CBO}}{\sqrt[n]{\beta_F}} \quad (1.81)$$

Equation 1.81 shows that BV_{CEO} is less than BV_{CBO} by a substantial factor. However, the value of BV_{CBO} , which must be used in (1.81), is the *plane* junction breakdown of the collector-base junction, neglecting any edge effects. This is because it is only collector-base avalanche current actually under the emitter that is amplified as described in the previous calculation. However, as explained in Section 1.2.2, the measured value of BV_{CBO} is usually determined by avalanche in the curved region of the collector, which is remote from the active base. Consequently, for typical values of $\beta_F = 100$ and $n = 4$, the value of BV_{CEO} is about one-half of the measured BV_{CBO} and not 30 percent as (1.81) would indicate.

Equation 1.81 explains the shape of the breakdown characteristics of Fig. 1.12 if the dependence of β_F on collector current is included. As V_{CE} is increased from zero with $I_B = 0$, the initial collector current is approximately $\beta_F I_{CO}$ from (1.73); since I_{CO} is typically several picoamperes, the collector current is very small. As explained in the next section, β_F is small at low currents, and thus from (1.81) the breakdown voltage is high. However, as avalanche breakdown begins in the device, the value of I_C increases and thus β_F increases. From (1.81) this causes a *decrease* in the breakdown voltage and the characteristic bends back as shown in Fig. 1.12 and exhibits a negative slope. At higher collector currents, β_F approaches a constant value and the breakdown curve with $I_B = 0$ becomes perpendicular to the V_{CE} axis. The value of V_{CE} in this region of the curve is usually defined to be BV_{CEO} , since this is the maximum voltage the device can sustain. The value of β_F to be used to calculate BV_{CEO} in (1.81) is thus the *peak* value of β_F . Note from (1.81) that high- β transistors will thus have low values of BV_{CEO} .

The base-emitter junction of a transistor is also subject to avalanche breakdown. However, the doping density in the emitter is made very large to ensure a high value of β_F [N_D is made large in (1.45) to reduce I_{B2}]. Thus the base is the more lightly doped side of the junction and determines the breakdown characteristic. This can be contrasted with the collector-base junction, where the collector is the more lightly doped side and results in typical values of BV_{CBO} of 20 to 80 V or more. The base is typically an order of magnitude more heavily doped than the collector, and thus the base-emitter breakdown voltage is much less than BV_{CBO} and is typically about 6 to 8 V. This is designated BV_{EBO} . The breakdown voltage for inverse-active operation shown in Fig. 1.12 is approximately equal to this value because the base-emitter junction is reverse biased in this mode of operation.

The base-emitter breakdown voltage of 6 to 8 V provides a convenient reference voltage in integrated-circuit design, and this is often utilized in the form of a *Zener* diode. However,

care must be taken to ensure that all other transistors in a circuit are protected against reverse base-emitter voltages sufficient to cause breakdown. This is because, unlike collector-base breakdown, base-emitter breakdown is damaging to the device. It can cause a large degradation in β_F , depending on the duration of the breakdown-current flow and its magnitude.¹¹ If the device is used purely as a Zener diode, this is of no consequence, but if the device is an amplifying transistor, the β_F degradation may be serious.

■ EXAMPLE

If the collector doping density in a transistor is 2×10^{15} atoms/cm³ and is much less than the base doping, calculate BV_{CEO} for $\beta = 100$ and $n = 4$. Assume $\mathcal{E}_{\text{crit}} = 3 \times 10^5$ V/cm.

The plane breakdown voltage in the collector can be calculated from (1.24) using $\mathcal{E}_{\text{max}} = \mathcal{E}_{\text{crit}}$:

$$BV_{CBO} = \frac{\epsilon (N_A + N_D)}{2qN_A N_D} \mathcal{E}_{\text{crit}}^2$$

Since $N_D \ll N_A$, we have

$$BV_{CBO}|_{\text{plane}} = \frac{\epsilon}{2qN_D} \mathcal{E}_{\text{crit}}^2 = \frac{1.04 \times 10^{-12}}{2 \times 1.6 \times 10^{-19} \times 2 \times 10^{15}} \times 9 \times 10^{10} \text{ V} = 146 \text{ V}$$

From (1.81)

$$BV_{CEO} = \frac{146}{\sqrt[3]{100}} \text{ V} = 46 \text{ V}$$

■

1.3.5 Dependence of Transistor Current Gain β_F on Operating Conditions

Although most first-order analyses of integrated circuits make the assumption that β_F is constant, this parameter does in fact depend on the operating conditions of the transistor. It was shown in Section 1.3.2, for example, that increasing the value of V_{CE} increases I_C while producing little change in I_B , and thus the effective β_F of the transistor increases. In Section 1.3.4 it was shown that as V_{CE} approaches the breakdown voltage, BV_{CEO} , the collector current increases due to avalanche multiplication in the collector. Equation 1.77 shows that the effective current gain approaches infinity as V_{CE} approaches BV_{CEO} .

In addition to the effects just described, β_F also varies with both temperature and transistor collector current. This is illustrated in Fig. 1.15, which shows typical curves of β_F versus I_C at three different temperatures for an *npn* integrated circuit transistor. It is evident that β_F increases as temperature increases, and a typical temperature coefficient for β_F is $+7000$ ppm/°C (where ppm signifies *parts per million*). This temperature dependence of β_F is due to the effect of the extremely high doping density in the emitter,¹² which causes the emitter injection efficiency γ to increase with temperature.

The variation of β_F with collector current, which is apparent in Fig. 1.15, can be divided into three regions. Region I is the low-current region, where β_F decreases as I_C decreases. Region II is the midcurrent region, where β_F is approximately constant. Region III is the high-current region, where β_F decreases as I_C increases. The reasons for this behavior of β_F with I_C can be better appreciated by plotting base current I_B and collector current I_C on a log scale

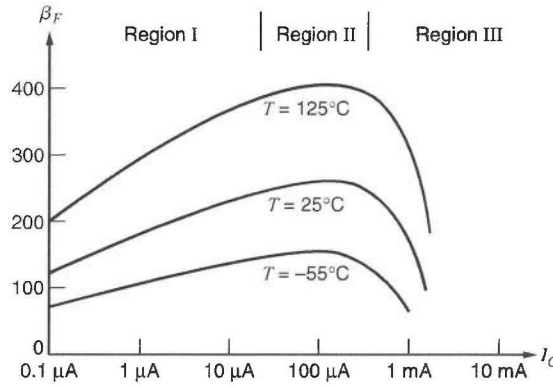


Figure 1.15 Typical curves of β_F versus I_C for an *npn* integrated-circuit transistor with $6 \mu\text{m}^2$ emitter area.

as a function of V_{BE} . This is shown in Fig. 1.16, and because of the log scale on the vertical axis, the value of $\ln \beta_F$ can be obtained directly as the distance between the two curves.

At moderate current levels represented by region II in Figs. 1.15 and 1.16, both I_C and I_B follow the ideal behavior, and

$$I_C = I_S \exp \frac{V_{BE}}{V_T} \quad (1.82)$$

$$I_B \simeq \frac{I_S}{\beta_{FM}} \exp \frac{V_{BE}}{V_T} \quad (1.83)$$

where β_{FM} is the maximum value of β_F and is given by (1.48).

At low current levels, I_C still follows the ideal relationship of (1.82), and the decrease in β_F is due to an additional component in I_B , which is mainly due to recombination of carriers in the base-emitter depletion region and is present at any current level. However, at higher current levels the base current given by (1.83) dominates, and this additional component has

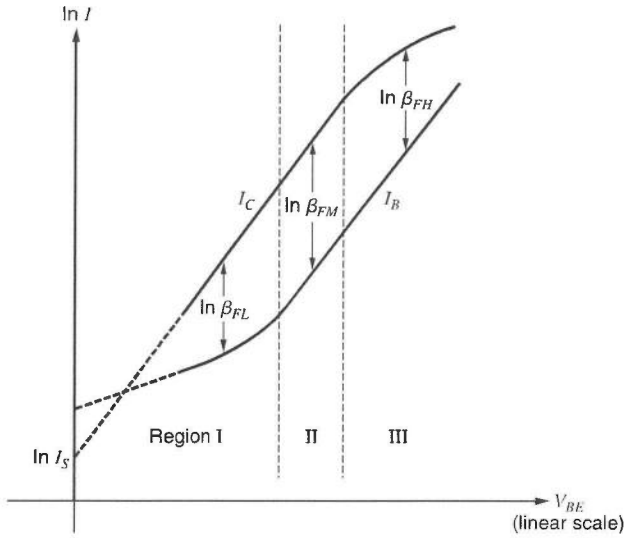


Figure 1.16 Base and collector currents of a bipolar transistor plotted on a log scale versus V_{BE} on a linear scale. The distance between the curves is a direct measure of $\ln \beta_F$.

little effect. The base current resulting from recombination in the depletion region is⁵

$$I_{BX} \simeq I_{SX} \exp \frac{V_{BE}}{m V_T} \quad (1.84)$$

where

$$m \simeq 2$$

At very low collector currents, where (1.84) dominates the base current, the current gain can be calculated from (1.82) and (1.84) as

$$\beta_{FL} \simeq \frac{I_C}{I_{BX}} = \frac{I_S}{I_{SX}} \exp \frac{V_{BE}}{V_T} \left(1 - \frac{1}{m} \right) \quad (1.85)$$

Substitution of (1.82) in (1.85) gives

$$\beta_{FL} \simeq \frac{I_S}{I_{SX}} \left(\frac{I_C}{I_S} \right)^{[1-(1/m)]} \quad (1.86)$$

If $m \simeq 2$, then (1.86) indicates that β_F is proportional to $\sqrt{I_C}$ at very low collector currents.

At high current levels, the base current I_B tends to follow the relationship of (1.83), and the decrease in β_F in region III is due mainly to a decrease in I_C below the value given by (1.82). (In practice the measured curve of I_B versus V_{BE} in Fig. 1.16 may also deviate from a straight line at high currents due to the influence of voltage drop across the base resistance.) The decrease in I_C is due partly to the effect of high-level injection, and at high current levels the collector current approaches⁷

$$I_C \simeq I_{SH} \exp \frac{V_{BE}}{2V_T} \quad (1.87)$$

The current gain in this region can be calculated from (1.87) and (1.83) as

$$\beta_{FH} \simeq \frac{I_{SH}}{I_S} \beta_{FM} \exp \left(-\frac{V_{BE}}{2V_T} \right) \quad (1.88)$$

Substitution of (1.87) in (1.88) gives

$$\beta_{FH} \simeq \frac{I_{SH}^2}{I_S} \beta_{FM} \frac{1}{I_C}$$

Thus β_F decreases rapidly at high collector currents.

In addition to the effect of high-level injection, the value of β_F at high currents is also decreased by the onset of the Kirk effect,¹³ which occurs when the minority-carrier concentration in the collector becomes comparable to the donor-atom doping density. The base region of the transistor then stretches out into the collector and becomes greatly enlarged.

1.4 Small-Signal Models of Bipolar Transistors

Analog circuits often operate with signal levels that are small compared to the bias currents and voltages in the circuit. In these circumstances, *incremental* or *small-signal* models can be derived that allow calculation of circuit gain and terminal impedances without the necessity of including the bias quantities. A hierarchy of models with increasing complexity can be derived, and the more complex ones are generally reserved for computer analysis. Part of the designer's skill is knowing which elements of the model can be omitted when performing hand calculations on a particular circuit, and this point is taken up again later.

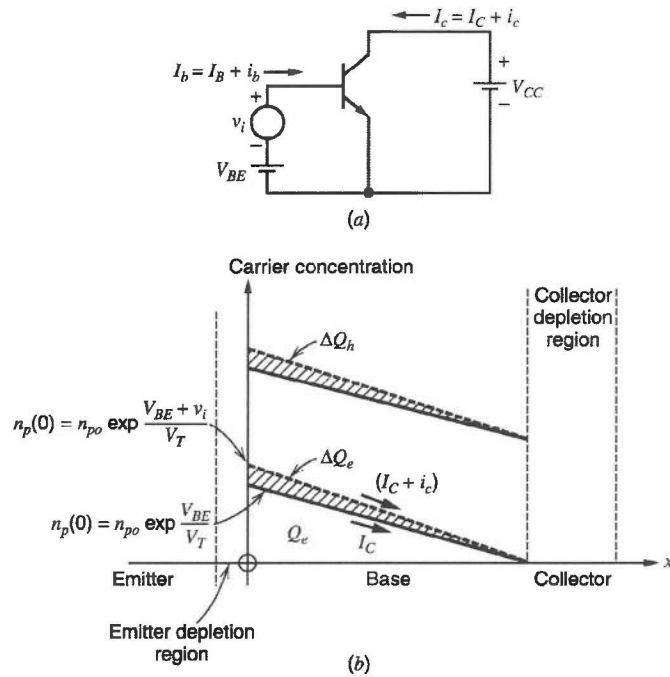


Figure 1.17 Effect of a small-signal input voltage applied to a bipolar transistor. (a) Circuit schematic. (b) Corresponding changes in carrier concentrations in the base when the device is in the forward-active region.

Consider the bipolar transistor in Fig. 1.17a with bias voltages V_{BE} and V_{CC} applied as shown. These produce a quiescent collector current, I_C , and a quiescent base current, I_B , and the device is in the *forward-active region*. A *small-signal* input voltage v_i is applied in series with V_{BE} and produces a small variation in base current i_b and a small variation in collector current i_c . Total values of base and collector currents are I_b and I_c , respectively, and thus $I_b = (I_B + i_b)$ and $I_c = (I_C + i_c)$. The carrier concentrations in the base of the transistor corresponding to the situation in Fig. 1.17a are shown in Fig. 1.17b. With only bias voltages applied, the carrier concentrations are given by the solid lines. Application of the small-signal voltage v_i causes $n_p(0)$ at the emitter edge of the base to increase, and produces the concentrations shown by the dotted lines. These pictures can now be used to derive the various elements in the small-signal equivalent circuit of the bipolar transistor.

1.4.1 Transconductance

The transconductance is defined as

$$g_m = \frac{dI_C}{dV_{BE}} \quad (1.89)$$

Since

$$\Delta I_C = \frac{dI_C}{dV_{BE}} \Delta V_{BE}$$

we can write

$$\Delta I_C = g_m \Delta V_{BE}$$

and thus

$$i_c = g_m v_i \quad (1.90)$$

The value of g_m can be found by substituting (1.35) in (1.89) to give

$$g_m = \frac{d}{dV_{BE}} I_S \exp \frac{V_{BE}}{V_T} = \frac{I_S}{V_T} \exp \frac{V_{BE}}{V_T} = \frac{I_C}{V_T} = \frac{qI_C}{kT} \quad (1.91)$$

The transconductance thus depends linearly on the bias current I_C and is 38 mA/V for $I_C = 1$ mA at 25°C for any bipolar transistor of either polarity (*nnp* or *npn*), of any size, and made of any material (Si, Ge, GaAs).

To illustrate the limitations on the use of small-signal analysis, the foregoing relation will be derived in an alternative way. The total collector current in Fig. 1.17a can be calculated using (1.35) as

$$I_c = I_S \exp \frac{V_{BE} + v_i}{V_T} = I_S \exp \frac{V_{BE}}{V_T} \exp \frac{v_i}{V_T} \quad (1.92)$$

But the collector bias current is

$$I_C = I_S \exp \frac{V_{BE}}{V_T} \quad (1.93)$$

and use of (1.93) in (1.92) gives

$$I_c = I_C \exp \frac{v_i}{V_T} \quad (1.94)$$

If $v_i < V_T$, the exponential in (1.94) can be expanded in a power series,

$$I_c = I_C \left[1 + \frac{v_i}{V_T} + \frac{1}{2} \left(\frac{v_i}{V_T} \right)^2 + \frac{1}{6} \left(\frac{v_i}{V_T} \right)^3 + \dots \right] \quad (1.95)$$

Now the incremental collector current is

$$i_c = I_c - I_C \quad (1.96)$$

and substitution of (1.96) in (1.95) gives

$$i_c = \frac{I_C}{V_T} v_i + \frac{1}{2} \frac{I_C}{V_T^2} v_i^2 + \frac{1}{6} \frac{I_C}{V_T^3} v_i^3 + \dots \quad (1.97)$$

If $v_i \ll V_T$, (1.97) reduces to (1.90), and the small-signal analysis is valid. The criterion for use of small-signal analysis is thus $v_i = \Delta V_{BE} \ll 26$ mV at 25°C. In practice, if ΔV_{BE} is less than 10 mV, the small-signal analysis is accurate within about 10 percent.

1.4.2 Base-Charging Capacitance

Figure 1.17b shows that the change in base-emitter voltage $\Delta V_{BE} = v_i$ has caused a change $\Delta Q_e = q_e$ in the minority-carrier charge in the base. By charge-neutrality requirements, there is an equal change $\Delta Q_h = q_h$ in the majority-carrier charge in the base. Since majority carriers

are supplied by the base lead, the application of voltage v_i requires the supply of charge q_h to the base, and the device has an apparent input capacitance

$$C_b = \frac{q_h}{v_i} \quad (1.98)$$

The value of C_b can be related to fundamental device parameters as follows. If (1.39) is divided by (1.33), we obtain

$$\frac{Q_e}{I_C} = \frac{W_B^2}{2D_n} = \tau_F \quad (1.99)$$

The quantity τ_F has the dimension of time and is called the base transit time in the forward direction. Since it is the ratio of the charge in transit (Q_e) to the current flow (I_C), it can be identified as the average time per carrier spent in crossing the base. To a first order it is independent of operating conditions and has typical values 10 to 500 ps for integrated *npn* transistors and 1 to 40 ns for lateral *pnp* transistors. Practical values of τ_F tend to be somewhat lower than predicted by (1.99) for diffused transistors that have nonuniform base doping.¹⁴ However, the functional dependence on base width W_B and diffusion constant D_n is as predicted by (1.99).

From (1.99)

$$\Delta Q_e = \tau_F \Delta I_C \quad (1.100)$$

But since $\Delta Q_e = \Delta Q_h$, we have

$$\Delta Q_h = \tau_F \Delta I_C \quad (1.101)$$

and this can be written

$$q_h = \tau_F i_c \quad (1.102)$$

Use of (1.102) in (1.98) gives

$$C_b = \tau_F \frac{i_c}{v_i} \quad (1.103)$$

and substitution of (1.90) in (1.103) gives

$$C_b = \tau_F g_m \quad (1.104)$$

$$= \tau_F \frac{q I_C}{kT} \quad (1.105)$$

Thus the small-signal, base-charging capacitance is proportional to the collector bias current.

In the inverse-active mode of operation, an equation similar to (1.99) relates stored charge and current via a time constant τ_R . This is typically orders of magnitude larger than τ_F because the device structure and doping are optimized for operation in the forward-active region. Since the saturation region is a combination of forward-active and inverse-active operation, inclusion of the parameter τ_R in a SPICE listing will model the large charge storage that occurs in saturation.

1.4.3 Input Resistance

In the forward-active region, the base current is related to the collector current by

$$I_B = \frac{I_C}{\beta_F} \quad (1.47)$$

Small changes in I_B and I_C can be related using (1.47):

$$\Delta I_B = \frac{d}{dI_C} \left(\frac{I_C}{\beta_F} \right) \Delta I_C \quad (1.106)$$

and thus

$$\beta_0 = \frac{\Delta I_C}{\Delta I_B} = \frac{i_c}{i_b} = \left[\frac{d}{dI_C} \left(\frac{I_C}{\beta_F} \right) \right]^{-1} \quad (1.107)$$

where β_0 is the *small-signal* current gain of the transistor. Note that if β_F is constant, then $\beta_F = \beta_0$. Typical values of β_0 are close to those of β_F , and in subsequent chapters little differentiation is made between these quantities. A single value of β is often assumed for a transistor and then used for both ac and dc calculations.

Equation 1.107 relates the change in base current i_b to the corresponding change in collector current i_c , and the device has a small-signal input resistance given by

$$r_\pi = \frac{v_i}{i_b} \quad (1.108)$$

Substitution of (1.107) in (1.108) gives

$$r_\pi = \frac{v_i}{i_c} \beta_0 \quad (1.109)$$

and use of (1.90) in (1.109) gives

$$r_\pi = \frac{\beta_0}{g_m} \quad (1.110)$$

Thus the small-signal input shunt resistance of a bipolar transistor depends on the current gain and is inversely proportional to I_C .

1.4.4 Output Resistance

In Section 1.3.2 the effect of changes in collector-emitter voltage V_{CE} on the large-signal characteristics of the transistor was described. It follows from that treatment that small changes ΔV_{CE} in V_{CE} produce corresponding changes ΔI_C in I_C , where

$$\Delta I_C = \frac{\partial I_C}{\partial V_{CE}} \Delta V_{CE} \quad (1.111)$$

Substitution of (1.55) and (1.57) in (1.111) gives

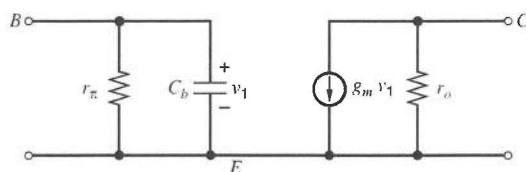
$$\frac{\Delta V_{CE}}{\Delta I_C} = \frac{V_A}{I_C} = r_o \quad (1.112)$$

where V_A is the Early voltage and r_o is the small-signal output resistance of the transistor. Since typical values of V_A are 50 to 100 V, corresponding values of r_o are 50 to 100 k Ω for $I_C = 1$ mA. Note that r_o is inversely proportional to I_C , and thus r_o can be related to g_m , as are many of the other small-signal parameters.

$$r_o = \frac{1}{\eta g_m} \quad (1.113)$$

where

$$\eta = \frac{kT}{qV_A} \quad (1.114)$$



$$r_{\pi} = \frac{\beta}{g_m}, r_o = \frac{1}{\eta g_m}, g_m = \frac{qI_C}{kT}, C_b = \tau_F g_m$$

Figure 1.18 Basic bipolar transistor small-signal equivalent circuit.

If $V_A = 100$ V, then $\eta = 2.6 \times 10^{-4}$ at 25°C . Note that $1/r_o$ is the slope of the output characteristics of Fig. 1.10.

1.4.5 Basic Small-Signal Model of the Bipolar Transistor

Combination of the above small-signal circuit elements yields the small-signal model of the bipolar transistor shown in Fig. 1.18. This is valid for both *npn* and *pnp* devices in the forward-active region and is called the *hybrid- π* model. Collector, base, and emitter nodes are labeled *C*, *B* and *E*, respectively. The elements in this circuit are present in the equivalent circuit of *any* bipolar transistor and are specified by relatively few parameters (β , τ_F , η , I_C). Note that in the evaluation of the small-signal parameters for *pnp* transistors, the *magnitude only* of I_C is used. In the following sections, further elements are added to this model to account for parasitics and second-order effects.

1.4.6 Collector-Base Resistance

Consider the effect of variations in V_{CE} on the minority charge in the base as illustrated in Fig. 1.9. An increase in V_{CE} causes an increase in the collector depletion-layer width and consequent reduction of base width. This causes a reduction in the total minority-carrier charge stored in the base and thus a reduction in base current I_B due to a reduction in I_{B1} given by (1.40). Since an increase ΔV_{CE} in V_{CE} causes a *decrease* ΔI_B in I_B , this effect can be modeled by inclusion of a resistor r_{μ} from collector to base of the model of Fig. 1.18. If V_{BE} is assumed held constant, the value of this resistor can be determined as follows.

$$r_{\mu} = \frac{\Delta V_{CE}}{\Delta I_{B1}} = \frac{\Delta V_{CE}}{\Delta I_C} \frac{\Delta I_C}{\Delta I_{B1}} \quad (1.115)$$

Substitution of (1.112) in (1.115) gives

$$r_{\mu} = r_o \frac{\Delta I_C}{\Delta I_{B1}} \quad (1.116)$$

If the base current I_B is composed entirely of component I_{B1} , then (1.107) can be used in (1.116) to give

$$r_{\mu} = \beta_0 r_o \quad (1.117)$$

This is a lower limit for r_{μ} . In practice, I_{B1} is typically less than 10 percent of I_B [component I_{B2} from (1.42) dominates] in integrated *npn* transistors, and since I_{B1} is very small, the change ΔI_{B1} in I_{B1} for a given ΔV_{CE} and ΔI_C is also very small. Thus a typical value for r_{μ} is greater than $10\beta_0 r_o$. For lateral *pnp* transistors, recombination in the base is more significant, and r_{μ} is in the range $2\beta_0 r_o$ to $5\beta_0 r_o$.

1.4.7 Parasitic Elements in the Small-Signal Model

The elements of the bipolar transistor small-signal equivalent circuit considered so far may be considered basic in the sense that they arise directly from essential processes in the device. However, technological limitations in the fabrication of transistors give rise to a number of parasitic elements that must be added to the equivalent circuit for most integrated-circuit transistors. A cross section of a typical *nnp* transistor in a junction-isolated process is shown in Fig. 1.19. The means of fabricating such devices is described in Chapter 2.

As described in Section 1.2, all *pn* junctions have a voltage-dependent capacitance associated with the depletion region. In the cross section of Fig. 1.19, three depletion-region capacitances can be identified. The base-emitter junction has a depletion-region capacitance C_{je} and the base-collector and collector-substrate junctions have capacitances C_{μ} and C_{cs} , respectively. The base-emitter junction closely approximates an abrupt junction due to the steep rise of the doping density caused by the heavy doping in the emitter. Thus the variation of C_{je} with bias voltage is well approximated by (1.21). The collector-base junction behaves like a graded junction for small bias voltages since the doping density is a function of distance near the junction. However, for larger reverse-bias values (more than about a volt), the junction depletion region spreads into the collector, which is uniformly doped, and thus for devices with thick collectors the junction tends to behave like an abrupt junction with uniform doping. Many modern high-speed processes, however, have very thin collector regions (of the order of one micron), and the collector depletion region can extend all the way to the buried layer for quite small reverse-bias voltages. When this occurs, both the depletion region and the associated capacitance vary quite slowly with bias voltage. The collector-base capacitance C_{μ} thus tends to follow (1.22) for very small bias voltages and (1.21) for large bias voltages in thick-collector devices. In practice, measurements show that the variation of C_{μ} with bias voltage for most devices can be approximated by

$$C_{\mu} = \frac{C_{\mu 0}}{\left(1 - \frac{V}{\psi_0}\right)^n} \quad (1.117a)$$

where V is the forward bias on the junction and n is an exponent between about 0.2 and 0.5. The third parasitic capacitance in a monolithic *nnp* transistor is the collector-substrate capacitance

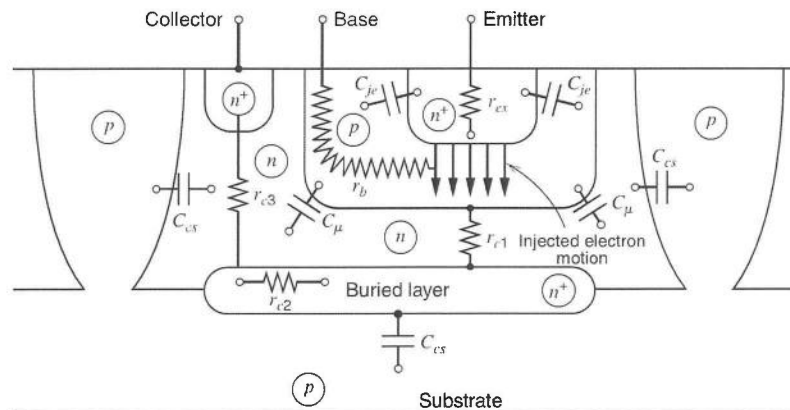


Figure 1.19 Integrated-circuit *nnp* bipolar transistor structure showing parasitic elements. (Not to scale.)

C_{cs} , and for large reverse bias voltages this varies according to the abrupt junction equation (1.21) for junction-isolated devices. In the case of oxide-isolated devices, however, the deep p diffusions used to isolate the devices are replaced by oxide. The sidewall component of C_{cs} then consists of a fixed oxide capacitance. Equation 1.117a may then be used to model C_{cs} , but a value of n less than 0.5 gives the best approximation. In general, (1.117a) will be used to model all three parasitic capacitances with subscripts e , c , and s on n and ψ_0 used to differentiate emitter-base, collector-base, and collector-substrate capacitances, respectively. Typical zero-bias values of these parasitic capacitances for a minimum-size nnp transistor in a modern oxide-isolated process are $C_{je0} \simeq 10$ fF, $C_{\mu 0} \simeq 10$ fF, and $C_{cs0} \simeq 20$ fF. Values for other devices are summarized in Chapter 2.

As described in Chapter 2, lateral pnp transistors have a parasitic capacitance C_{bs} from base to substrate in place of C_{cs} . Note that the substrate is always connected to the most negative voltage supply in the circuit in order to ensure that all isolation regions are separated by reverse-biased junctions. Thus the substrate is an ac ground, and all parasitic capacitance to the substrate is connected to ground in an equivalent circuit.

The final elements to be added to the small-signal model of the transistor are resistive parasitics. These are produced by the finite resistance of the silicon between the top contacts on the transistor and the active base region beneath the emitter. As shown in Fig. 1.19, there are significant resistances r_b and r_c in series with the base and collector contacts, respectively. There is also a resistance r_{ex} of several ohms in series with the emitter lead that can become important at high bias currents. (Note that the collector resistance r_c is actually composed of three parts labeled r_{c1} , r_{c2} , and r_{c3} .) Typical values of these parameters are $r_b = 50$ to 500Ω , $r_{ex} = 1$ to 3Ω , and $r_c = 20$ to 500Ω . The value of r_b varies significantly with collector current because of *current crowding*.¹⁵ This occurs at high collector currents where the dc base current produces a lateral voltage drop in the base that tends to forward bias the base-emitter junction preferentially around the edges of the emitter. Thus the transistor action tends to occur along the emitter periphery rather than under the emitter itself, and the distance from the base contact to the active base region is reduced. Consequently, the value of r_b is reduced, and in a typical nnp transistor, r_b may decrease 50 percent as I_C increases from 0.1 mA to 10 mA.

The value of these parasitic resistances can be reduced by changes in the device structure. For example, a large-area transistor with multiple base and emitter stripes will have a smaller value of r_b . The value of r_c is reduced by inclusion of the low-resistance buried n^+ layer beneath the collector.

The addition of the resistive and capacitive parasitics to the basic small-signal circuit of Fig. 1.18 gives the complete small-signal equivalent circuit of Fig. 1.20. The internal base node

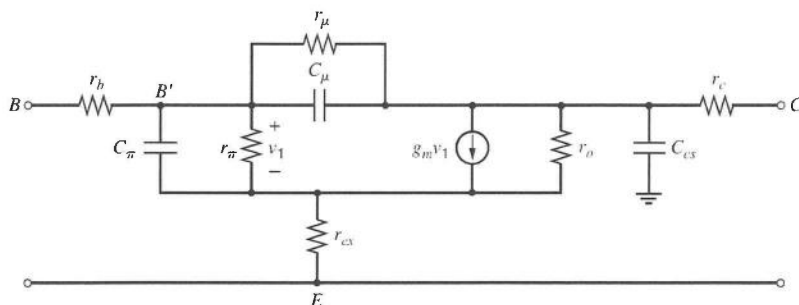


Figure 1.20 Complete bipolar transistor small-signal equivalent circuit.

is labeled B' to distinguish it from the external base contact B . The capacitance C_π contains the base-charging capacitance C_b and the emitter-base depletion layer capacitance C_{je} .

$$C_\pi = C_b + C_{je} \quad (1.118)$$

Note that the representation of parasitics in Fig. 1.20 is an approximation in that lumped elements have been used. In practice, as suggested by Fig. 1.19, C_μ is distributed across r_b and C_{cs} is distributed across r_c . This lumped representation is adequate for most purposes but can introduce errors at very high frequencies. It should also be noted that while the parasitic resistances of Fig. 1.20 can be very important at high bias currents or for high-frequency operation, they are usually omitted from the equivalent circuit for low-frequency calculations, particularly for collector bias currents less than 1 mA.

■ EXAMPLE

Derive the complete small-signal equivalent circuit for a bipolar transistor at $I_C = 1$ mA, $V_{CB} = 3$ V, and $V_{CS} = 5$ V. Device parameters are $C_{je0} = 10$ fF, $n_e = 0.5$, $\psi_{0e} = 0.9$ V, $C_{\mu0} = 10$ fF, $n_c = 0.3$, $\psi_{0c} = 0.5$ V, $C_{cs0} = 20$ fF, $n_s = 0.3$, $\psi_{0s} = 0.65$ V, $\beta_0 = 100$, $\tau_F = 10$ ps, $V_A = 20$ V, $r_b = 300 \Omega$, $r_c = 50 \Omega$, $r_{ex} = 5 \Omega$, $r_\mu = 10 \beta_0 r_o$.

Since the base-emitter junction is forward biased, the value of C_{je} is difficult to determine for reasons described in Section 1.2.1. Either a value can be determined by computer or a reasonable estimation is to double C_{je0} . Using the latter approach, we estimate

$$C_{je} = 20 \text{ fF}$$

Using (1.117a) gives, for the collector-base capacitance,

$$C_\mu = \frac{C_{\mu0}}{\left(1 + \frac{V_{CB}}{\psi_{0c}}\right)^{n_c}} = \frac{10}{\left(1 + \frac{3}{0.5}\right)^{0.3}} = 5.6 \text{ fF}$$

The collector-substrate capacitance can also be calculated using (1.117a)

$$C_{cs} = \frac{C_{cs0}}{\left(1 + \frac{V_{CS}}{\psi_{0s}}\right)^{n_s}} = \frac{20}{\left(1 + \frac{5}{0.65}\right)^{0.3}} = 10.5 \text{ fF}$$

From (1.91) the transconductance is

$$g_m = \frac{qI_C}{kT} = \frac{10^{-3}}{26 \times 10^{-3}} \text{ A/V} = 38 \text{ mA/V}$$

From (1.104) the base-charging capacitance is

$$C_b = \tau_F g_m = 10 \times 10^{-12} \times 38 \times 10^{-3} \text{ F} = 0.38 \text{ pF}$$

The value of C_π from (1.118) is

$$C_\pi = 0.38 + 0.02 \text{ pF} = 0.4 \text{ pF}$$

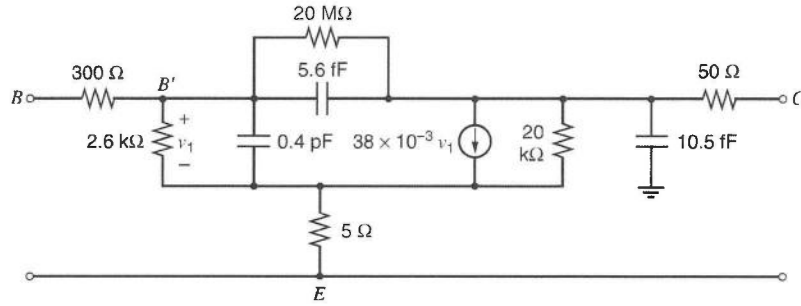


Figure 1.21 Complete small-signal equivalent circuit for a bipolar transistor at $I_C = 1$ mA, $V_{CB} = 3$ V, and $V_{CS} = 5$ V. Device parameters are $C_{je0} = 10$ fF, $n_e = 0.5$, $\psi_{0e} = 0.9$ V, $C_{\mu0} = 10$ fF, $n_c = 0.3$, $\psi_{0c} = 0.5$ V, $C_{cs0} = 20$ fF, $n_s = 0.3$, $\psi_{0s} = 0.65$ V, $\beta_0 = 100$, $\tau_F = 10$ ps, $V_A = 20$ V, $r_b = 300 \Omega$, $r_c = 50 \Omega$, $r_{ex} = 5 \Omega$, $r_\mu = 10\beta_0 r_o$.

The input resistance from (1.110) is

$$r_\pi = \frac{\beta_0}{g_m} = 100 \times 26 \Omega = 2.6 \text{ k}\Omega$$

The output resistance from (1.112) is

$$r_o = \frac{20}{10^{-3}} \Omega = 20 \text{ k}\Omega$$

and thus the collector-base resistance is

$$r_\mu = 10\beta_0 r_o = 10 \times 100 \times 20 \text{ k}\Omega = 20 \text{ M}\Omega$$

- The equivalent circuit with these parameter values is shown in Fig. 1.21.

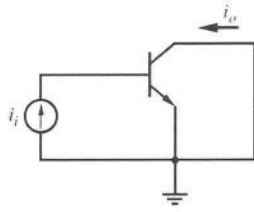
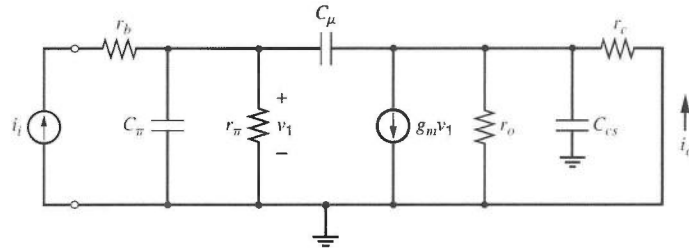
1.4.8 Specification of Transistor Frequency Response

The high-frequency gain of the transistor is controlled by the capacitive elements in the equivalent circuit of Fig. 1.20. The frequency capability of the transistor is most often specified in practice by determining the frequency where the magnitude of the short-circuit, common-emitter current gain falls to unity. This is called the *transition frequency*, f_T , and is a measure of the maximum useful frequency of the transistor when it is used as an amplifier. The value of f_T can be measured as well as calculated, using the ac circuit of Fig. 1.22. A small-signal current i_i is applied to the base, and the output current i_o is measured with the collector short-circuited for ac signals. A small-signal equivalent circuit can be formed for this situation by using the equivalent circuit of Fig. 1.20 as shown in Fig. 1.23, where r_{ex} and r_μ have been neglected. If r_c is assumed small, then r_o and C_{cs} have no influence, and we have

$$v_1 \simeq \frac{r_\pi}{1 + r_\pi(C_\pi + C_\mu)s} i_i \quad (1.119)$$

If the current fed forward through C_μ is neglected,

$$i_o \simeq g_m v_1 \quad (1.120)$$

Figure 1.22 Schematic of ac circuit for measurement of f_T .Figure 1.23 Small-signal equivalent circuit for the calculation of f_T .

Substitution of (1.119) in (1.120) gives

$$i_o \simeq i_i \frac{g_m r_\pi}{1 + r_\pi (C_\pi + C_\mu)s}$$

and thus

$$\frac{i_o}{i_i}(j\omega) = \frac{\beta_0}{1 + \beta_0 \frac{C_\pi + C_\mu}{g_m} j\omega} \quad (1.121)$$

using (1.110).

Now if $i_o/i_i(j\omega)$ is written as $\beta(j\omega)$ (the high-frequency, small-signal current gain), then

$$\beta(j\omega) = \frac{\beta_0}{1 + \beta_0 \frac{C_\pi + C_\mu}{g_m} j\omega} \quad (1.122)$$

At high frequencies the imaginary part of the denominator of (1.122) is dominant, and we can write

$$\beta(j\omega) \simeq \frac{g_m}{j\omega (C_\pi + C_\mu)} \quad (1.123)$$

From (1.123), $|\beta(j\omega)| = 1$ when

$$\omega = \omega_T = \frac{g_m}{C_\pi + C_\mu} \quad (1.124)$$

and thus

$$f_T = \frac{1}{2\pi} \frac{g_m}{C_\pi + C_\mu} \quad (1.125)$$

The transistor behavior can be illustrated by plotting $|\beta(j\omega)|$ using (1.122) as shown in Fig. 1.24. The frequency ω_β is defined as the frequency where $|\beta(j\omega)|$ is equal to $\beta_0/\sqrt{2}$

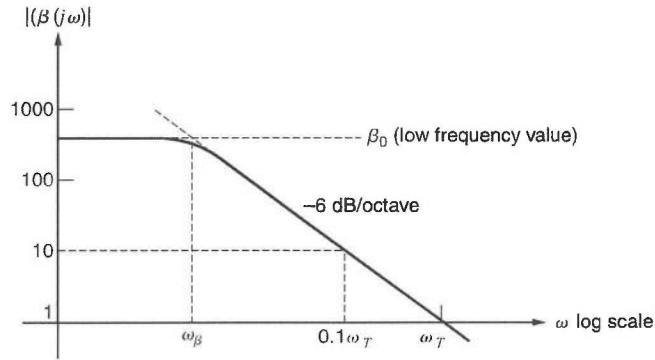


Figure 1.24 Magnitude of small-signal ac current gain $|\beta(j\omega)|$ versus frequency for a typical bipolar transistor.

(3 dB down from the low-frequency value). From (1.122) we have

$$\omega_\beta = \frac{1}{\beta_0} \frac{g_m}{C_\pi + C_\mu} = \frac{\omega_T}{\beta_0} \quad (1.126)$$

From Fig. 1.24 it can be seen that ω_T can be determined by measuring $|\beta(j\omega)|$ at some frequency ω_x where $|\beta(j\omega)|$ is falling at 6 dB/octave and using

$$\omega_T = \omega_x |\beta(j\omega_x)| \quad (1.127)$$

This is the method used in practice, since deviations from ideal behavior tend to occur as $|\beta(j\omega)|$ approaches unity. Thus $|\beta(j\omega)|$ is typically measured at some frequency where its magnitude is about 5 or 10, and (1.127) is used to determine ω_T .

It is interesting to examine the time constant, τ_T , associated with ω_T . This is defined as

$$\tau_T = \frac{1}{\omega_T} \quad (1.128)$$

and use of (1.124) in (1.128) gives

$$\tau_T = \frac{C_\pi}{g_m} + \frac{C_\mu}{g_m} \quad (1.129)$$

Substitution of (1.118) and (1.104) in (1.129) gives

$$\tau_T = \frac{C_b}{g_m} + \frac{C_{je}}{g_m} + \frac{C_\mu}{g_m} = \tau_F + \frac{C_{je}}{g_m} + \frac{C_\mu}{g_m} \quad (1.130)$$

Equation 1.130 indicates that τ_T is dependent on I_C (through g_m) and approaches a constant value of τ_F at high collector bias currents. At low values of I_C , the terms involving C_{je} and C_μ dominate, and they cause τ_T to rise and f_T to fall as I_C is decreased. This behavior is illustrated in Fig. 1.25, which is a typical plot of f_T versus I_C for an integrated-circuit npn transistor. The decline in f_T at high collector currents is not predicted by this simple theory and is due to an increase in τ_F caused by high-level injection and Kirk effect at high currents. These are the same mechanisms that cause a decrease in β_F at high currents as described in Section 1.3.5.

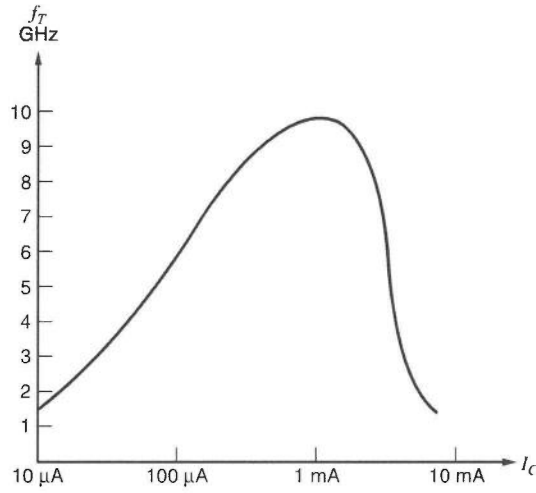


Figure 1.25 Typical curve of f_T versus I_C for an *npn* integrated-circuit transistor with $6 \mu\text{m}^2$ emitter area in a high-speed process.

■ EXAMPLE

A bipolar transistor has a short-circuit, common-emitter current gain at 1 GHz of 8 with $I_C = 0.25 \text{ mA}$ and 9 with $I_C = 1 \text{ mA}$. Assuming that high-level injection effects are negligible, calculate C_{je} and τ_F , assuming both are constant. The measured value of C_μ is 10 fF.

From the data, values of f_T are

$$f_{T1} = 8 \times 1 = 8 \text{ GHz at } I_C = 0.25 \text{ mA}$$

$$f_{T2} = 9 \times 1 = 9 \text{ GHz at } I_C = 1 \text{ mA}$$

Corresponding values of τ_T are

$$\tau_{T1} = \frac{1}{2\pi f_{T1}} = 19.9 \text{ ps}$$

$$\tau_{T2} = \frac{1}{2\pi f_{T2}} = 17.7 \text{ ps}$$

Using these data in (1.130), we have

$$19.9 \times 10^{-12} = \tau_F + 104(C_\mu + C_{je}) \quad (1.131)$$

at $I_C = 0.25 \text{ mA}$. At $I_C = 1 \text{ mA}$ we have

$$17.7 \times 10^{-12} = \tau_F + 26(C_\mu + C_{je}) \quad (1.132)$$

Subtraction of (1.132) from (1.131) yields

$$C_\mu + C_{je} = 28.2 \text{ fF}$$

Since C_μ was measured as 10 fF, the value of C_{je} is given by

$$C_{je} \simeq 18.2 \text{ fF}$$

Substitution in (1.131) gives

$$\tau_F = 17 \text{ ps}$$

This is an example of how basic device parameters can be determined from high-frequency current-gain measurements. Note that the assumption that C_{je} is constant is a useful approximation in practice because V_{BE} changes by only 36 mV as I_C increases from 0.25 mA to 1 mA.

1.5 Large-Signal Behavior of Metal-Oxide-Semiconductor Field-Effect Transistors

Metal-oxide-semiconductor field-effect transistors (MOSFETs) have become dominant in the area of digital integrated circuits because they allow high density and low power dissipation. In contrast, bipolar transistors still provide many advantages in stand-alone analog integrated circuits. For example, the transconductance per unit bias current in bipolar transistors is usually much higher than in MOS transistors. So in systems where analog techniques are used on some integrated circuits and digital techniques on others, bipolar technologies are often preferred for the analog integrated circuits and MOS technologies for the digital. To reduce system cost and increase portability, both increased levels of integration and reduced power dissipation are required, forcing the associated analog circuits to use MOS-compatible technologies. One way to achieve these goals is to use a processing technology that provides both bipolar and MOS transistors, allowing great design flexibility. However, all-MOS processes are less expensive than combined bipolar and MOS processes. Therefore, economic considerations drive integrated-circuit manufacturers to use all-MOS processes in many practical cases. As a result, the study of the characteristics of MOS transistors that affect analog integrated-circuit design is important.

1.5.1 Transfer Characteristics of MOS Devices

A cross section of a typical enhancement-mode n -channel MOS (NMOS) transistor is shown in Fig. 1.26. Heavily doped n -type source and drain regions are fabricated in a p -type substrate (often called the body). A thin layer of silicon dioxide is grown over the substrate material and

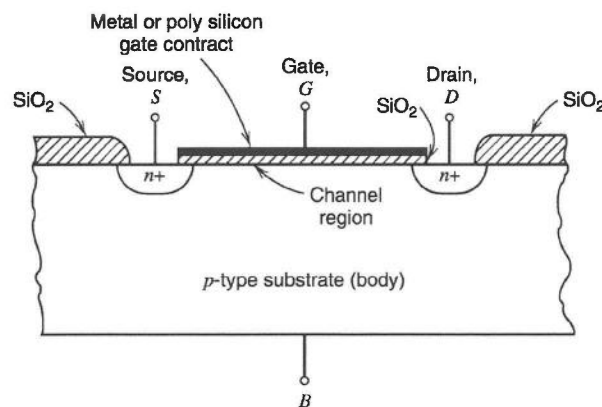


Figure 1.26 Typical enhancement-mode NMOS structure.

Current Mirrors, Active Loads, and References

4.1 Introduction

Current mirrors made by using active devices have come to be widely used in analog integrated circuits both as biasing elements and as load devices for amplifier stages. The use of current mirrors in biasing can result in superior insensitivity of circuit performance to variations in power supply and temperature. Current mirrors are frequently more economical than resistors in terms of the die area required to provide bias current of a certain value, particularly when the required value of bias current is small. When used as a load element in transistor amplifiers, the high incremental resistance of the current mirror results in high voltage gain at low power-supply voltages.

The first section of this chapter describes the general properties of current mirrors and compares various bipolar and MOS mirrors to each other using these properties. The next section deals with the use of current mirrors as load elements in amplifier stages. The last section shows how current mirrors are used to construct references that are insensitive to variations in supply and temperature. Finally, the appendix analyzes the effects of device mismatch.

4.2 Current Mirrors

4.2.1 General Properties

A current mirror is an element with at least three terminals, as shown in Fig. 4.1. The common terminal is connected to a power supply, and the input current source is connected to the input terminal. Ideally, the output current is equal to the input current multiplied by a desired current gain. If the gain is unity, the input current is reflected to the output, leading to the name *current mirror*. Under ideal conditions, the current-mirror gain is independent of input frequency, and the output current is independent of the voltage between the output and common terminals. Furthermore, the voltage between the input and common terminals is ideally zero because this condition allows the entire supply voltage to appear across the input current source, simplifying its transistor-level design. More than one input and/or output terminals are sometimes used.

In practice, real transistor-level current mirrors suffer many deviations from this ideal behavior. For example, the gain of a real current mirror is never independent of the input frequency. The topic of frequency response is covered in Chapter 7, and mainly dc and low-frequency ac signals are considered in the rest of this chapter. Deviations from ideality that will be considered in this chapter are listed below.

1. One of the most important deviations from ideality is the variation of the current-mirror output current with changes in voltage at the output terminal. This effect is characterized

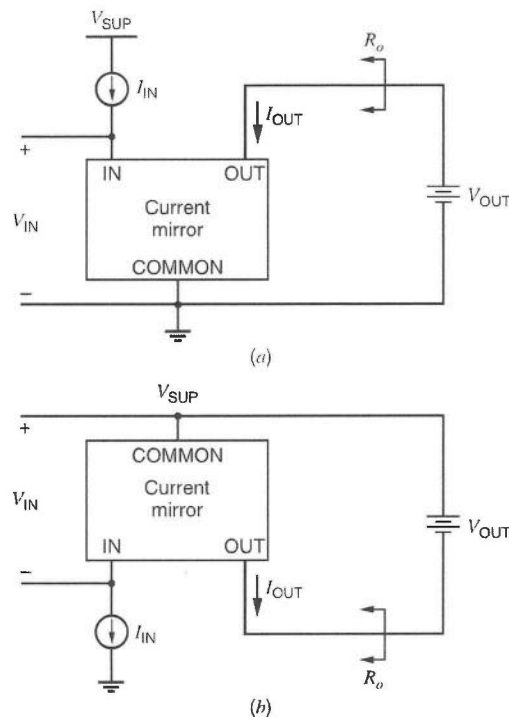


Figure 4.1 Current-mirror block diagrams referenced to (a) ground and (b) the positive supply.

by the small-signal output resistance, R_o , of the current mirror. A Norton-equivalent model of the output of the current mirror includes R_o in parallel with a current source controlled by the input current. The output resistance directly affects the performance of many circuits that use current mirrors. For example, the common-mode rejection ratio of the differential amplifier depends directly on this resistance, as does the gain of the active-load circuits. Increasing the output resistance reduces the dependence of the output current on the output voltage and is therefore desirable. Generally speaking, the output resistance increases in practical circuits when the output current decreases. Unfortunately, decreasing the output current also decreases the maximum operating speed. Therefore, when comparing the output resistance of two current mirrors, they should be compared at identical output currents.

2. Another important error source is the gain error, which is the deviation of the gain of a current mirror from its ideal value. The gain error is separated into two parts: (1) the systematic gain error and (2) the random gain error. The systematic gain error, ϵ , is the gain error that arises even when all matched elements in the mirror are perfectly matched and will be calculated for each of the current mirrors presented in this section. The random gain error is the gain error caused by unintended mismatches between matched elements.
3. When the input current source is connected to the input terminal of a real current mirror, it creates a positive voltage drop, V_{IN} , that reduces the voltage available across the input current source. Minimizing V_{IN} is important because it simplifies the design of the input current source, especially in low-supply applications. To reduce V_{IN} , current mirrors sometimes have more than one input terminal. In that case, we will calculate an input voltage for each input terminal. An example is the MOS high-swing cascode current mirror considered in Section 4.2.5.

4. A positive output voltage, V_{OUT} , is required in practice to make the output current depend mainly on the input current. This characteristic is summarized by the minimum voltage across the output branch, $V_{OUT(min)}$, that allows the output device(s) to operate in the active region. Minimizing $V_{OUT(min)}$ maximizes the range of output voltages for which the current-mirror output resistance is almost constant, which is important in applications where current mirrors are used as active loads in amplifiers (especially with low power-supply voltages). This topic is covered in Section 4.3. When current mirrors have more than one output terminal, each output must be biased above its $V_{OUT(min)}$ to make the corresponding output current depend mainly on the input current.

In later sections, the performance of various current mirrors will be compared to each other through these four parameters: R_o , ϵ , V_{IN} , and $V_{OUT(min)}$.

4.2.2 Simple Current Mirror

4.2.2.1 Bipolar

The simplest form of a current mirror consists of two transistors. Fig. 4.2 shows a bipolar version of this mirror. Transistor Q_1 is diode connected, forcing its collector-base voltage to zero. In this mode, the collector-base junction is off in the sense that no injection takes place there, and Q_1 operates in the forward-active region. Assume that Q_2 also operates in the forward-active region and that both transistors have infinite output resistance. Then I_{OUT} is controlled by V_{BE2} , which is equal to V_{BE1} by KVL. A KVL equation is at the heart of the operation of all current mirrors. Neglecting junction leakage currents,

$$V_{BE2} = V_T \ln \frac{I_{C2}}{I_{S2}} = V_{BE1} = V_T \ln \frac{I_{C1}}{I_{S1}} \quad (4.1)$$

where $V_T = kT/q$ is the thermal voltage and I_{S1} and I_{S2} are the transistor saturation currents. From (4.1),

$$I_{C2} = \frac{I_{S2}}{I_{S1}} I_{C1} \quad (4.2)$$

If the transistors are identical, $I_{S1} = I_{S2}$ and (4.2) shows that the current flowing in the collector of Q_1 is mirrored to the collector of Q_2 . KCL at the collector of Q_1 yields

$$I_{IN} - I_{C1} - \frac{I_{C1}}{\beta_F} - \frac{I_{C2}}{\beta_F} = 0 \quad (4.3)$$

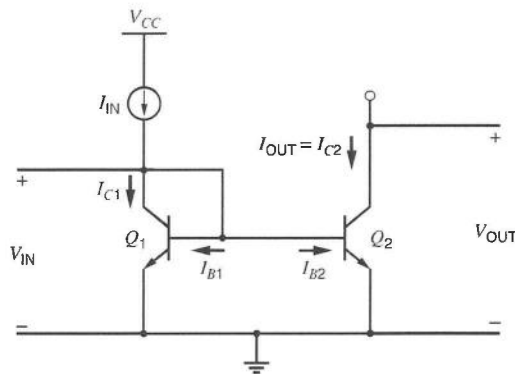


Figure 4.2 A simple bipolar current mirror.

Therefore, with identical transistors,

$$I_{OUT} = I_{C2} = I_{C1} = \frac{I_{IN}}{1 + \frac{2}{\beta_F}} \quad (4.4)$$

If β_F is large, the base currents are small and

$$I_{OUT} = I_{C1} \simeq I_{IN} \quad (4.5)$$

Thus for identical devices Q_1 and Q_2 , the gain of the current mirror is approximately unity. This result holds for both dc and low-frequency ac currents. Above the 3-dB frequency of the mirror, however, the base current increases noticeably because the impedance of the base-emitter capacitance decreases, reducing the gain of the current mirror. Frequency response is studied in Chapter 7. The rest of this section considers dc currents only.

In practice, the devices need not be identical. Then from (4.2) and (4.3),

$$I_{OUT} = \frac{I_{S2}}{I_{S1}} I_{C1} = \left(\frac{I_{S2}}{I_{S1}} I_{IN} \right) \left(\frac{1}{1 + \frac{1 + (I_{S2}/I_{S1})}{\beta_F}} \right) \quad (4.6)$$

When $I_{S2} = I_{S1}$, (4.6) is the same as (4.4). Since the saturation current of a bipolar transistor is proportional to its emitter area, the first term in (4.6) shows that the gain of the current mirror can be larger or smaller than unity because the emitter areas can be ratioed. If the desired current-mirror gain is a rational number, M/N , the area ratio is usually set by connecting M identical devices called *units* in parallel to form Q_2 and N units in parallel to form Q_1 to minimize mismatch arising from lithographic effects in forming the emitter regions. However, area ratios greater than about five to one consume a large die area dominated by the area of the larger of the two devices. Thus other methods described in later sections are preferred for the generation of large current ratios. The last term in (4.6) accounts for error introduced by finite β_F . Increasing I_{S2}/I_{S1} increases the magnitude of this error by increasing the base current of Q_2 compared to that of Q_1 .

In writing (4.1) and (4.2), we assumed that the collector currents of the transistors are independent of their collector-emitter voltages. If a transistor is biased in the forward-active region, its collector current actually increases slowly with increasing collector-emitter voltage. Fig. 4.3 shows an output characteristic for Q_2 . The output resistance of the current mirror at any given operating point is the reciprocal of the slope of the output characteristic at that point. In the forward-active region,

$$R_o = r_{o2} = \frac{V_A}{I_{C2}} \quad (4.7)$$

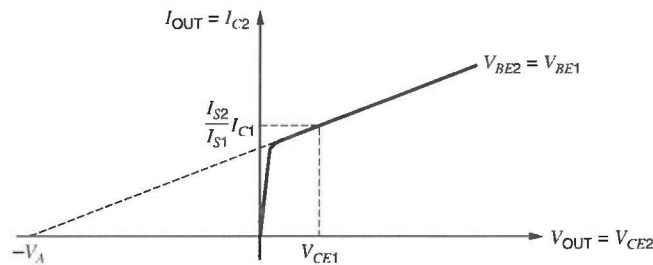


Figure 4.3 npn output characteristic.

The point where $V_{CE2} = V_{CE1}$ and $V_{BE2} = V_{BE1}$ is labeled on the characteristic. Because the collector current is controlled by the base-emitter and collector-emitter voltages, $I_{C2} = (I_{S2}/I_{S1})I_{C1}$ at this point. If the slope of the characteristic in saturation is constant, the variation in I_{C2} for changes in V_{CE2} can be predicted by a straight line that goes through the labeled point. As described in Chapter 1, extrapolation of the output characteristic in the forward-active region back to the V_{CE2} axis gives an intercept at $-V_A$, where V_A is the Early voltage. If $V_A \gg V_{CE1}$, the slope of the straight line is about equal to $(I_{S2}/I_{S1})(I_{C1}/V_A)$. Therefore,

$$I_{OUT} = \frac{I_{S2}}{I_{S1}} I_{C1} \left(1 + \frac{V_{CE2} - V_{CE1}}{V_A} \right) = \frac{\frac{I_{S2}}{I_{S1}} I_{IN} \left(1 + \frac{V_{CE2} - V_{CE1}}{V_A} \right)}{1 + \frac{1 + (I_{S2}/I_{S1})}{\beta_F}} \quad (4.8)$$

Since the ideal gain of the current mirror is I_{S2}/I_{S1} , the systematic gain error, ϵ , of the current mirror can be calculated from (4.8).

$$\epsilon = \left(\frac{1 + \frac{V_{CE2} - V_{CE1}}{V_A}}{1 + \frac{1 + (I_{S2}/I_{S1})}{\beta_F}} \right) - 1 \approx \frac{V_{CE2} - V_{CE1}}{V_A} - \frac{1 + (I_{S2}/I_{S1})}{\beta_F} \quad (4.9)$$

The first term in (4.9) stems from finite output resistance and the second term from finite β_F . If $V_{CE2} > V_{CE1}$, the polarities of the two terms are opposite. Since the two terms are independent, however, cancellation is unlikely in practice. The first term dominates when the difference in the collector-emitter voltages and β_F are large. For example, with identical transistors and $V_A = 130$ V, if the collector-emitter voltage of Q_1 is held at $V_{BE(on)}$, and if the collector-emitter voltage of Q_2 is 30 V, then the systematic gain error $(30 - 0.6)/130 - 2/200 \approx 0.22$. Thus for a circuit operating at a power-supply voltage of 30 V, the current-mirror currents can differ by more than 20 percent from those values calculated by assuming that the transistor output resistance and β_F are infinite. Although the first term in (4.9) stems from finite output resistance, it does not depend on r_{o2} directly but instead on the collector-emitter and Early voltages. The Early voltage is independent of the bias current, and

$$V_{IN} = V_{CE1} = V_{BE1} = V_{BE(on)} \quad (4.10)$$

Since $V_{BE(on)}$ is proportional to the natural logarithm of the collector current, V_{IN} changes little with changes in bias current. Therefore, changing the bias current in a current mirror changes systematic gain error mainly through changes in V_{CE2} .

Finally, the minimum output voltage required to keep Q_2 in the forward-active region is

$$V_{OUT(min)} = V_{CE2(sat)} \quad (4.11)$$

4.2.2.2 MOS

Figure 4.4 shows an MOS version of the simple current mirror. The drain-gate voltage of M_1 is zero; therefore, the channel does not exist at the drain, and the transistor operates in the saturation or active region if the threshold is positive. Although the principle of operation for MOS transistors does not involve forward biasing any diodes, M_1 is said to be *diode connected* in an analogy to the bipolar case. Assume that M_2 also operates in the active region and that both transistors have infinite output resistance. Then I_{D2} is controlled by V_{GS2} , which is equal to V_{GS1} by KVL. A KVL equation is at the heart of the operation of all current mirrors. As described in Section 1.5.3, the gate-source voltage of a given MOS transistor is usually

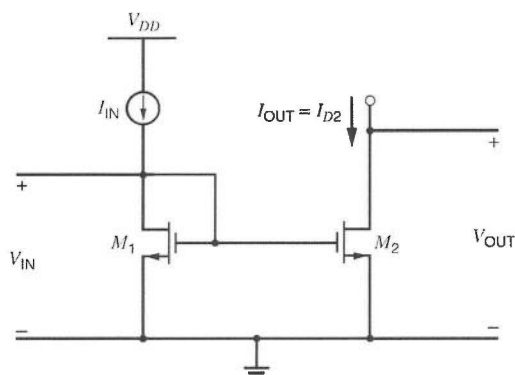


Figure 4.4 A simple MOS current mirror.

separated into two parts: the threshold V_t and the overdrive V_{ov} . Assuming square-law behavior as in (1.157), the overdrive for M_2 is

$$V_{ov2} = V_{GS2} - V_t = \sqrt{\frac{2I_{D2}}{k'(W/L)_2}} \quad (4.12)$$

Since the transconductance parameter k' is proportional to mobility, and since mobility falls with increasing temperature, the overdrive rises with temperature. In contrast, Section 1.5.4 shows that the threshold falls with increasing temperature. From KVL and (1.157),

$$V_{GS2} = V_t + \sqrt{\frac{2I_{D2}}{k'(W/L)_2}} = V_{GS1} = V_t + \sqrt{\frac{2I_{D1}}{k'(W/L)_1}} \quad (4.13)$$

Equation 4.13 shows that the overdrive of M_2 is equal to that of M_1 .

$$V_{ov2} = V_{ov1} = V_{ov} \quad (4.14)$$

If the transistors are identical, $(W/L)_2 = (W/L)_1$, and therefore

$$I_{OUT} = I_{D2} = I_{D1} \quad (4.15)$$

Equation 4.15 shows that the current that flows in the drain of M_1 is mirrored to the drain of M_2 . Since $\beta_F \rightarrow \infty$ for MOS transistors, (4.15) and KCL at the drain of M_1 yield

$$I_{OUT} = I_{D1} = I_{IN} \quad (4.16)$$

Thus for identical devices operating in the active region with infinite output resistance, the gain of the current mirror is unity. This result holds when the gate currents are zero; that is, (4.16) is at least approximately correct for dc and low-frequency ac currents. As the input frequency increases, however, the gate currents of M_1 and M_2 increase because each transistor has a nonzero gate-source capacitance. The part of the input current that flows into the gate leads does not flow into the drain of M_1 and is not mirrored to M_2 ; therefore, the gain of the current mirror decreases as the frequency of the input current increases. The rest of this section considers dc currents only.

In practice, the devices need not be identical. Then from (4.13) and (4.16),

$$I_{OUT} = \frac{(W/L)_2}{(W/L)_1} I_{D1} = \frac{(W/L)_2}{(W/L)_1} I_{IN} \quad (4.17)$$

Equation 4.17 shows that the gain of the current mirror can be larger or smaller than unity because the transistor sizes can be ratioed. To ratio the transistor sizes, either the widths or the lengths can be made unequal in principle. In practice, however, the lengths of M_1 and M_2 are rarely made unequal. The lengths that enter into (4.17) are the effective channel lengths given by (2.35). Equation 2.35 shows that the effective channel length of a given transistor differs from its drawn length by offset terms stemming from the depletion region at the drain and lateral diffusion at the drain and source. Since the offset terms are independent of the drawn length, a ratio of two effective channel lengths is equal to the drawn ratio only if the drawn lengths are identical. As a result, a ratio of unequal channel lengths depends on process parameters that may not be well controlled in practice. Similarly, Section 2.9.1 shows that the effective width of a given transistor differs from the drawn width because of lateral oxidation resulting in a *bird's beak*. Therefore, a ratio of unequal channel widths will also be process dependent. In many applications, however, the shortest channel length allowed in a given technology is selected for most transistors to maximize speed and minimize area. In contrast, the drawn channel widths are usually many times larger than the minimum dimensions allowed in a given technology. Therefore, to minimize the effect of the offset terms when the current-mirror gain is designed to differ from unity, the widths are ratioed rather than the lengths in most practical cases. If the desired current-mirror gain is a rational number, M/N , the ratio is usually set by connecting M identical devices called *units* in parallel to form M_2 and N units in parallel to form M_1 to minimize mismatch arising from lithographic effects in forming the gate regions. As in the bipolar case, ratios greater than about five to one consume a large die area dominated by the area of the larger of the two devices. Thus other methods described in later sections are preferred for the generation of large current ratios.

In writing (4.13) and (4.15), we assumed that the drain currents of the transistors are independent of their drain-source voltages. If a transistor is biased in the active region, its drain current actually increases slowly with increasing drain-source voltage. Figure 4.5 shows an output characteristic for M_2 . The output resistance of the current mirror at any given operating point is the reciprocal of the slope of the output characteristic at that point. In the active region,

$$R_o = r_{o2} = \frac{V_A}{I_{D2}} = \frac{1}{\lambda I_{D2}} \quad (4.18)$$

The point where $V_{DS2} = V_{DS1}$ and $V_{GS2} = V_{GS1}$ is labeled on the characteristic. Because the drain current is controlled by the gate-source and drain-source voltages, $I_{D2} = [(W/L)_2/(W/L)_1]I_{D1}$ at this point. If the slope of the characteristic in saturation is constant, the variation in I_{D2} for changes in V_{DS2} can be predicted by a straight line that goes through the

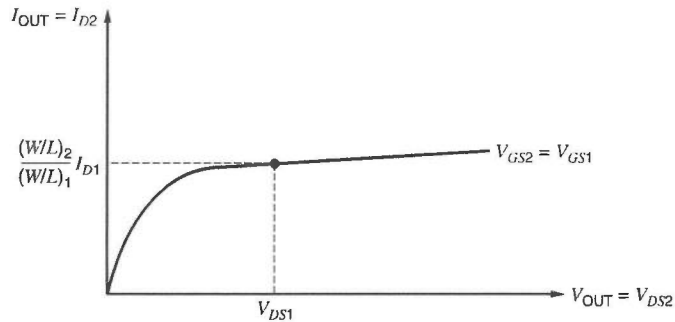


Figure 4.5 Output characteristic of simple MOS current mirror.

labeled point. As described in Chapter 1, extrapolation of the output characteristic in the active region back to the V_{DS2} axis gives an intercept at $-V_A = -1/\lambda$, where V_A is the Early voltage. If $V_A \gg V_{DS1}$, the slope of the straight line is about equal to $[(W/L)_2/(W/L)_1][I_{D1}/V_A]$. Therefore,

$$I_{OUT} = \frac{(W/L)_2}{(W/L)_1} I_{IN} \left(1 + \frac{V_{DS2} - V_{DS1}}{V_A} \right) \quad (4.19)$$

Since the ideal gain of the current mirror is $(W/L)_2/(W/L)_1$, the systematic gain error, ϵ , of the current mirror can be calculated from (4.19).

$$\epsilon = \frac{V_{DS2} - V_{DS1}}{V_A} \quad (4.20)$$

For example, if the drain-source voltage of M_1 is held at 1.2 V, and if the drain-source voltage of M_2 is 5 V, then the systematic gain error is $(5 - 1.2)/10 \simeq 0.38$ with $V_A = 10$ V. Thus for a circuit operating at a power-supply voltage of 5 V, the current-mirror currents can differ by more than 35 percent from those values calculated by assuming that the transistor output resistance is infinite. Although ϵ stems from finite output resistance, it does not depend on r_{o2} directly but instead on the drain-source and Early voltages. Since the Early voltage is independent of the bias current, this observation shows that changing the input bias current in a current mirror changes systematic gain error mainly through changes to the drain-source voltages.

For the simple MOS current mirror, the input voltage is

$$V_{IN} = V_{GS1} = V_t + V_{ov1} = V_t + V_{ov} \quad (4.21)$$

With square-law behavior, the overdrive in (4.21) is proportional to the square root of the input current. In contrast, (4.10) shows that the entire V_{IN} in a simple bipolar mirror is proportional to the natural logarithm of the input current. Therefore, for a given change in the input current, the variation in V_{IN} in a simple MOS current mirror is generally larger than in its bipolar counterpart.

Finally, the minimum output voltage required to keep M_2 in the active region is

$$V_{OUT(min)} = V_{ov2} = V_{ov} = \sqrt{\frac{2I_{OUT}}{k'(W/L)_2}} \quad (4.22)$$

Equation 4.22 predicts that $V_{OUT(min)}$ depends on the transistor geometry and can be made arbitrarily small in a simple MOS mirror, unlike in the bipolar case. However, if the overdrive predicted by (4.22) is less than $2nV_T$, where n is defined in (1.247) and V_T is a thermal voltage, the result is invalid except to indicate that the transistors operate in weak inversion. At room temperature with $n = 1.5$, $2nV_T \simeq 78$ mV. If the transistors operate in weak inversion,

$$V_{OUT(min)} \simeq 3V_T \quad (4.23)$$

as shown in Fig. 1.43.¹

4.2.3 Simple Current Mirror with Beta Helper

4.2.3.1 Bipolar

In addition to the variation in output current due to finite output resistance, the second term in (4.9) shows that the collector current I_{C2} differs from the input current because of finite β_F . To reduce this source of error, an additional transistor can be added, as shown in Fig. 4.6. If Q_1 and Q_3 are identical, the emitter current of transistor Q_2 is

$$I_{E2} = -\frac{I_{C1}}{\beta_F} - \frac{I_{C3}}{\beta_F} = -\frac{2}{\beta_F} I_{C1} \quad (4.24)$$

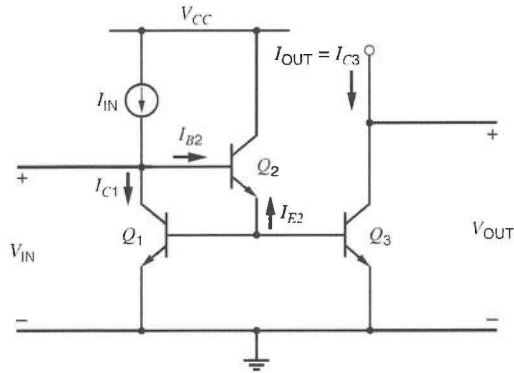


Figure 4.6 Simple current mirror with beta helper.

where I_E , I_C , and I_B are defined as positive when flowing into the transistor, and where we have neglected the effects of finite output resistance. The base current of transistor Q_2 is equal to

$$I_{B2} = -\frac{I_{E2}}{\beta_F + 1} = \frac{2}{\beta_F(\beta_F + 1)} I_{C1} \quad (4.25)$$

Finally, KCL at the collector of Q_1 gives

$$I_{IN} - I_{C1} - \frac{2}{\beta_F(\beta_F + 1)} I_{C1} = 0 \quad (4.26)$$

Since I_{C1} and I_{C3} are equal when Q_1 and Q_3 are identical,

$$I_{OUT} = I_{C3} = \frac{I_{IN}}{1 + \frac{2}{\beta_F(\beta_F + 1)}} \simeq I_{IN} \left(1 - \frac{2}{\beta_F(\beta_F + 1)} \right) \quad (4.27)$$

Equation 4.27 shows that the systematic gain error from finite β_F has been reduced by a factor of $[\beta_F + 1]$, which is the current gain of emitter follower Q_2 . As a result, Q_2 is often referred to as a *beta helper*.

Although the beta helper has little effect on the output resistance and the minimum output voltage of the current mirror, it increases the input voltage by the base-emitter voltage of Q_2 :

$$V_{IN} = V_{BE1(on)} + V_{BE2(on)} \quad (4.28)$$

If multiple emitter followers are cascaded to further reduce the gain error arising from finite β_F , V_{IN} increases by an extra base-emitter voltage for each additional emitter follower, posing one limit to the use of cascaded emitter followers.

Current mirrors often use a beta helper when they are constructed with *pnp* transistors because the value of β_F for *pnp* transistors is usually less than for *nnp* transistors. Another application of the beta-helper configuration is in current mirrors with multiple outputs. An example with two independent outputs is shown in Fig. 4.7. At first, ignore Q_2 and imagine that Q_1 is simply diode connected. Also, let $R_1 = R_3 = R_4 = 0$ here. (The effects of nonzero resistances will be considered in Section 4.2.4.) Then the gain from the input to each output is primarily determined by the area ratios I_{S3}/I_{S1} and I_{S4}/I_{S1} . Because the bases of three instead of two transistors are connected together, the total base current is increased here, which increases the gain error from the input to either output arising from finite β_F . Furthermore, the

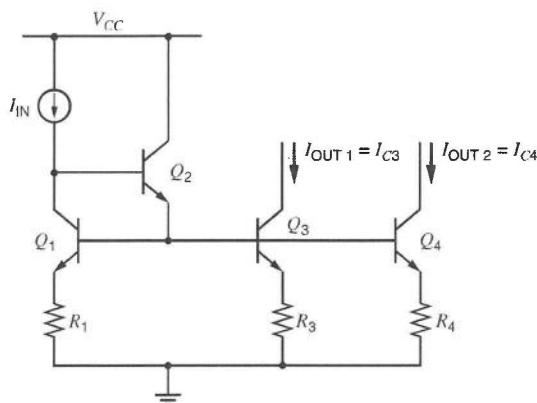


Figure 4.7 Simple current mirror with beta helper, multiple outputs, and emitter degeneration.

gain errors worsen as the number of independent outputs increases. Since the beta helper, Q_2 , reduces the gain error from the input to each output by a factor of $[\beta_F + 1]$, it is often used in bipolar current mirrors with multiple outputs.

4.2.3.2 MOS

Since $\beta_F \rightarrow \infty$ for an MOS transistor, beta helpers are not used in simple MOS current mirrors to reduce the systematic gain error. However, a beta-helper configuration can increase the bandwidth of MOS and bipolar current mirrors.

4.2.4 Simple Current Mirror with Degeneration

4.2.4.1 Bipolar

The performance of the simple bipolar transistor current mirror of Fig. 4.6 can be improved by the addition of emitter degeneration as shown in Fig. 4.7 for a current mirror with two independent outputs. The purpose of the emitter resistors is twofold. First, Appendix A.4.1 shows that the matching between I_{IN} and outputs I_{C3} and I_{C4} can be greatly improved by using emitter degeneration. Second, as shown in Section 3.3.8, the use of emitter degeneration boosts the output resistance of each output of the current mirror. Transistors Q_1 and Q_2 combine to present a very low resistance at the bases of Q_3 and Q_4 . Therefore, from (3.99), the small-signal output resistance seen at the collectors of Q_3 and Q_4 is

$$R_o \simeq r_o(1 + g_m R_E) \quad (4.29)$$

if $r_\pi \gg R_E$. Taking Q_3 as an example and using $g_{m3} = I_{C3}/V_T$, we find

$$R_o \simeq r_{o3} \left(1 + \frac{I_{C3} R_3}{V_T} \right) \quad (4.30)$$

This increase in the output resistance for a given output current also decreases the component of systematic gain error that stems from finite output resistance by the same factor. From (4.9) and (4.30) with infinite β_F ,

$$\epsilon \simeq \frac{V_{CE2} - V_{CE1}}{V_A \left(1 + \frac{I_{C3} R_3}{V_T} \right)} \quad (4.31)$$

The quantity $I_{C3}R_3$ is just the dc voltage drop across R_3 . If this quantity is 260 mV, for example, then R_o is about $10r_o$ at room temperature, and ϵ is reduced by a factor of about eleven. Unfortunately, this improvement in R_o is limited by corresponding increases in the input and minimum output voltages of the mirror:

$$V_{IN} \simeq V_{BE1(on)} + V_{BE2(on)} + I_{IN}R_1 \quad (4.32)$$

and

$$V_{OUT(min)} = V_{CE3(sat)} + I_{C3}R_3 \quad (4.33)$$

The emitter areas of Q_1 , Q_3 , and Q_4 may be matched or ratioed. For example, if we want $I_{OUT1} = I_{IN}$ and $I_{OUT2} = 2I_{IN}$, we would make Q_3 identical to Q_1 , and Q_4 consist of two copies of Q_1 connected in parallel so that $I_{S4} = 2I_{S1}$. In addition, we could make $R_3 = R_1$, and R_4 consist of two copies of R_1 connected in parallel so that $R_4 = R_1/2$. Note that all the dc voltage drops across R_1 , R_3 , and R_4 would then be equal. Using KVL around the loop including Q_1 and Q_4 and neglecting base currents, we find

$$I_{C1}R_1 + V_T \ln \frac{I_{C1}}{I_{S1}} = I_{C4}R_4 + V_T \ln \frac{I_{C4}}{I_{S4}} \quad (4.34)$$

from which

$$I_{OUT2} = I_{C4} = \frac{1}{R_4} \left(I_{IN}R_1 + V_T \ln \frac{I_{IN}}{I_{C4}} \frac{I_{S4}}{I_{S1}} \right) \quad (4.35)$$

Since $I_{S4} = 2I_{S1}$, the solution to (4.35) is

$$I_{OUT2} = \frac{R_1}{R_4} I_{IN} = 2I_{IN} \quad (4.36)$$

because the last term in (4.35) goes to zero. If we make the voltage drops $I_{IN}R_1$ and $I_{C4}R_4$ much greater than V_T , the current-mirror gain to the Q_4 output is determined primarily by the resistor ratio R_4/R_1 , and only to a secondary extent by the emitter area ratio, because the natural log term in (4.35) varies slowly with its argument.

4.2.4.2 MOS

Source degeneration is rarely used in MOS current mirrors because, in effect, MOS transistors are inherently controlled resistors. Thus, matching in MOS current mirrors is improved simply by increasing the gate areas of the transistors.^{2,3,4} Furthermore, the output resistance can be increased by increasing the channel length. To increase the output resistance while keeping the current and $V_{GS} - V_t$ constant, the W/L ratio must be held constant. Therefore, the channel width must be increased as much as the length, and the price paid for the improved output resistance is that increased chip area is consumed by the current mirror.

4.2.5 Cascode Current Mirror

4.2.5.1 Bipolar

Section 3.4.2 shows that the cascode connection achieves a very high output resistance. Since this is a desirable characteristic for a current mirror, exploring the use of cascodes for high-performance current mirrors is natural. A bipolar-transistor current mirror based on the cascode connection is shown in Fig. 4.8. Transistors Q_3 and Q_1 form a simple current mirror, and emitter resistances can be added to improve the matching. Transistor Q_2 acts as the common-base part of the cascode and transfers the collector current of Q_1 to the output while presenting a high

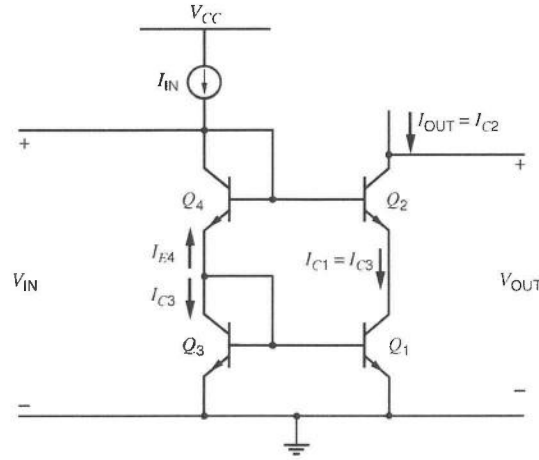


Figure 4.8 Cascode current mirror with bipolar transistors.

output resistance. Transistor Q_4 acts as a diode level shifter and biases the base of Q_2 so that Q_1 operates in the forward-active region with $V_{CE1} \simeq V_{CE3} = V_{BE3(\text{on})}$. If we assume that the small-signal resistances of diodes Q_3 and Q_4 are small, a direct application of (3.98) with $R_E = r_{o1}$ concludes that

$$R_o = r_{o2} \left(1 + \frac{g_{m2}r_{o1}}{1 + \frac{g_{m2}r_{o1}}{\beta_0}} \right) \simeq \beta_0 r_{o2} \quad (4.37)$$

because $g_{m2}r_{o1} \simeq g_{m1}r_{o1} \gg \beta_0$. This calculation assumes that almost all of the small-signal current that flows into the collector of Q_2 flows out its base because the small-signal resistance connected to the emitter of Q_2 is much greater than that connected to its base. A key problem with this calculation, however, is that it ignores the effect of the simple current mirror formed by Q_3 and Q_1 . Let i_{b2} and i_{e2} represent increases in the base and emitter currents flowing out of Q_2 caused by increasing output voltage. Then the simple mirror forces $i_{e2} \simeq i_{b2}$. As a result, the variation in the collector current of Q_2 splits into two equal parts and half flows in $r_{\pi2}$. A small-signal analysis shows that R_o in (4.37) is reduced by half to

$$R_o \simeq \frac{\beta_0 r_{o2}}{2} \quad (4.38)$$

Thus, the cascode configuration boosts the output resistance by approximately $\beta_0/2$. For $\beta_0 = 100$, $V_A = 130$ V, and $I_{C2} = 1$ mA,

$$R_o = \frac{\beta_0 V_A}{2I_{C2}} = \frac{100(130)}{2 \text{ mA}} = 6.5 \text{ M}\Omega \quad (4.39)$$

In this calculation of output resistance, we have neglected the effects of r_μ . Although this assumption is easy to justify in the case of the simple current mirror, it must be reexamined here because the output resistance is so high. The collector-base resistance r_μ results from modulation of the base-recombination current as a consequence of the Early effect, as described in Chapter 1. For a transistor whose base current is composed entirely of base-recombination current, the percentage change in base current when V_{CE} is changed at a constant V_{BE} would equal that of the collector current, and r_μ would be equal to $\beta_0 r_o$. In this case, the effect of

r_μ would be to reduce the output resistance of the cascode current mirror given in (4.38) by a factor of 1.5.

In actual integrated-circuit *nnp* transistors, however, only a small percentage of the base current results from recombination in the base. Since only this component is modulated by the Early effect, the observed values of r_μ are a factor of 10 or more larger than $\beta_0 r_o$. Therefore, r_μ has a negligible effect here with *nnp* transistors. On the other hand, for lateral *pnp* transistors, the feedback resistance r_μ is much smaller than for *nnp* transistors because most of the base current results from base-region recombination. The actual value of this resistance depends on a number of process and device-geometry variables, but observed values range from 2 to 5 times $\beta_0 r_o$. Therefore, for a cascode current mirror constructed with lateral *pnp* transistors, the effect of r_μ on the output resistance can be significant. Furthermore, when considering current mirrors that give output resistances higher than $\beta_0 r_o$, the effects of r_μ must be considered.

In the cascode current mirror, the base of Q_1 is connected to a low-resistance point because Q_3 is diode connected. As a result, feedback from $r_{\mu 1}$ is greatly attenuated and has negligible effect on the output resistance. On the other hand, if the resistance from the base of Q_1 to ground is increased while all other parameters are held constant, local feedback from $r_{\mu 1}$ significantly affects the base-emitter voltage of Q_1 and reduces the output resistance. In the limit where the resistance from the base of Q_1 to ground becomes infinite, Q_1 acts as if it were diode connected. Local feedback is considered in Chapter 8.

The input voltage of the cascode current mirror is

$$V_{IN} = V_{BE3} + V_{BE4} = 2V_{BE(on)} \quad (4.40)$$

Although V_{IN} is higher here than in (4.10) for a simple current mirror, the increase becomes a limitation only if the power-supply voltage is reduced to nearly two diode drops.

The minimum output voltage for which the output resistance is given by (4.38) must allow both Q_1 and Q_2 to be biased in the forward-active region. Since $V_{CE1} \simeq V_{CE3} = V_{BE(on)}$,

$$V_{OUT(min)} = V_{CE1} + V_{CE2(sat)} \simeq V_{BE(on)} + V_{CE2(sat)} \quad (4.41)$$

Comparing (4.41) and (4.11) shows that the minimum output voltage for a cascode current mirror is higher than for a simple current mirror by a diode drop. This increase poses an important limitation on the minimum supply voltage when the current mirror is used as an active load for an amplifier.

Since $V_{CE1} \simeq V_{CE3}$, $I_{C1} \simeq I_{C3}$, and the systematic gain error arising from finite transistor output resistance is almost zero. A key limitation of the cascode current mirror, however, is that the systematic gain error arising from finite β_F is worse than for a simple current mirror. From KCL at the collector of Q_3 ,

$$-I_{E4} = I_{C3} + \frac{2I_{C3}}{\beta_F} \quad (4.42)$$

From KCL at the collector of Q_4 ,

$$I_{IN} = -I_{E4} + \frac{I_{C2}}{\beta_F} \quad (4.43)$$

The collector current of Q_2 is

$$I_{C2} = \frac{\beta_F}{\beta_F + 1} I_{C3} \quad (4.44)$$

Substituting (4.42) and (4.44) into (4.43) gives

$$I_{IN} = I_{C3} + \frac{2I_{C3}}{\beta_F} + \frac{I_{C3}}{\beta_F + 1} \quad (4.45)$$

Rearranging (4.45) to find I_{C3} and substituting back into (4.44) gives

$$I_{OUT} = I_{C2} = \left(\frac{\beta_F}{\beta_F + 1} \right) \left(\frac{I_{IN}}{1 + \frac{2}{\beta_F} + \frac{1}{\beta_F + 1}} \right) \quad (4.46)$$

Equation 4.46 can be rearranged to give

$$I_{OUT} = I_{IN} \left(1 - \frac{4\beta_F + 2}{\beta_F^2 + 4\beta_F + 2} \right) \quad (4.47)$$

Equation 4.47 shows that the systematic gain error is

$$\epsilon = -\frac{4\beta_F + 2}{\beta_F^2 + 4\beta_F + 2} \quad (4.48)$$

When $\beta_F \gg 1$, (4.48) simplifies to

$$\epsilon \simeq -\frac{4}{\beta_F + 4} \quad (4.49)$$

In contrast, the systematic gain error stemming from finite β_F in a simple current mirror with identical transistors is about $-2/\beta_F$, which is less in magnitude than (4.49) predicts for a cascode current mirror if $\beta_F > 4$. This limitation of a cascode current mirror is overcome by the Wilson current mirror described in Section 4.2.6.

4.2.5.2 MOS

The cascode current mirror is widely used in MOS technology, where it does not suffer from finite β_F effects. Figure 4.9 shows the simplest form. From (3.107), the small-signal output resistance is

$$R_o = r_{o2}[1 + (g_{m2} + g_{mb2})r_{o1}] + r_{o1} \quad (4.50)$$

As shown in the previous section, the bipolar cascode current mirror cannot realize an output resistance larger than $\beta_0 r_o/2$ because β_0 is finite and nonzero small-signal base current flows in the cascode transistor. In contrast, the MOS cascode is capable of realizing arbitrarily high output resistance by increasing the number of stacked cascode devices because $\beta_0 \rightarrow \infty$ for MOS transistors. However, the MOS substrate leakage current described in Section 1.9 can create a resistive shunt to ground from the output node, which can dominate the output resistance for $V_{OUT} > V_{OUT(min)}$.⁵

■ EXAMPLE

Find the output resistance of the double-cascode current mirror shown in Fig. 4.10. Assume all the transistors operate in the active region with $I_D = 10 \mu\text{A}$, $V_A = 50 \text{ V}$, and $g_m r_o = 50$. Neglect body effect.

The output resistance of each transistor is

$$r_o = \frac{V_A}{I_D} = \frac{50 \text{ V}}{10 \mu\text{A}} = 5 \text{ M}\Omega$$

From (4.50), looking into the drain of M_2 :

$$R_{o2} = r_{o2}(1 + g_{m2}r_{o1}) + r_{o1} \quad (4.51)$$

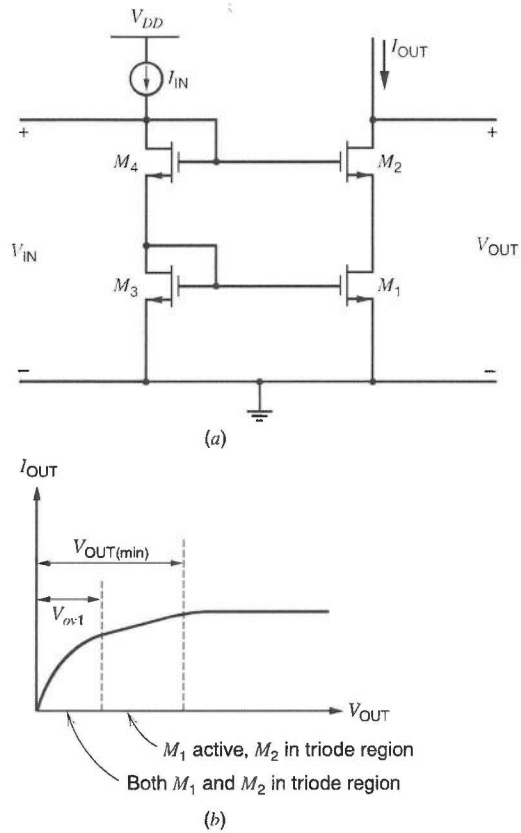


Figure 4.9 (a) Cascode current mirror using MOS transistors. (b) I-V characteristic.

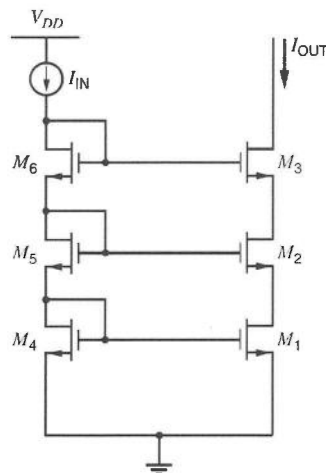


Figure 4.10 Example of a double-cascode current mirror.

Similarly, looking into the drain of M_3 :

$$R_o = r_{o3}[1 + g_{m3}R_{o2}] + R_{o2} \quad (4.52)$$

Each cascode stage increases the output resistance by a factor of about $(1 + g_m r_o)$. Therefore,

$$R_o \simeq r_o(1 + g_m r_o)^2 \simeq 5(51)^2 \text{ M}\Omega \simeq 13 \text{ G}\Omega \quad (4.53)$$

With such a large output resistance, other parasitic leakage paths, such as the substrate leakage path, could be comparable to this resistance in practice.

From KVL in Fig. 4.9,

$$V_{DS1} = V_{GS3} + V_{GS4} - V_{GS2} \quad (4.54)$$

Since $V_{DS3} = V_{GS3}$, (4.54) shows that $V_{DS1} = V_{DS3}$ when $V_{GS2} = V_{GS4}$. Under this condition, the systematic gain error of the cascode current mirror is zero because M_1 and M_3 are identically biased, and because $\beta_F \rightarrow \infty$ for MOS transistors. In practice, V_{GS2} is not exactly equal to V_{GS4} even with perfect matching unless $V_{OUT} = V_{IN}$ because of channel-length modulation. As a result, $V_{DS1} \simeq V_{DS3}$ and

$$\epsilon \simeq 0 \quad (4.55)$$

The input voltage of the MOS cascode current mirror in Fig. 4.9 is

$$\begin{aligned} V_{IN} &= V_{GS3} + V_{GS4} \\ &= V_{t3} + V_{ov3} + V_{t4} + V_{ov4} \end{aligned} \quad (4.56)$$

The input voltage here includes two gate-source drops, each composed of threshold and overdrive components. Ignoring the body effect and assuming the transistors all have equal overdrives,

$$V_{IN} = 2V_t + 2V_{ov} \quad (4.57)$$

Also, adding extra cascode levels increases the input voltage by another threshold and another overdrive component for each additional cascode. Furthermore, the body effect increases the threshold of all transistors with $V_{SB} > 0$. Together, these facts increase the difficulty of designing the input current source for low power-supply voltages.

When M_1 and M_2 both operate in the active region, $V_{DS1} \simeq V_{DS3} = V_{GS3}$. For M_2 to operate in the active region, $V_{DS2} > V_{ov2}$ is required. Therefore, the minimum output voltage for which M_1 and M_2 operate in the active region is

$$\begin{aligned} V_{OUT(\min)} &= V_{DS1} + V_{ov2} \\ &\simeq V_{GS3} + V_{ov2} = V_t + V_{ov3} + V_{ov2} \end{aligned} \quad (4.58)$$

If the transistors all have equal overdrives,

$$V_{OUT(\min)} \simeq V_t + 2V_{ov} \quad (4.59)$$

On the other hand, M_2 operates in the triode region if $V_{OUT} < V_{OUT(\min)}$, and both M_1 and M_2 operate in the triode region if $V_{OUT} < V_{ov1}$. These results are shown graphically in Fig. 4.9b.

Although the overdrive term in (4.59) can be made small by using large values of W for a given current, the threshold term represents a significant loss of voltage swing when the current mirror is used as an active load in an amplifier. The threshold term in (4.59) stems from the biasing of the drain-source voltage of M_1 so that

$$V_{DS1} = V_{IN} - V_{GS2} \quad (4.60)$$

Ignoring the body effect and assuming that M_1 - M_4 all operate in the active region with equal overdrives,

$$V_{DS1} = V_t + V_{ov} \quad (4.61)$$

Therefore, the drain-source voltage of M_1 is a threshold larger than necessary to operate M_1 in the active region. To reduce V_{DS1} , the voltage from the gate of M_2 to ground can be level shifted down by a threshold as shown in Fig. 4.11a. In practice, a source follower is used to implement the level shift, as shown in Fig. 4.11b.⁶ Transistor M_5 acts as the source follower and is biased by the output of the simple current mirror M_3 and M_6 . Because the gate-source voltage of M_5 is greater than its threshold by the overdrive, however, the drain-source voltage of M_1 would be zero with equal thresholds and overdrives on all transistors. To bias M_1 at the

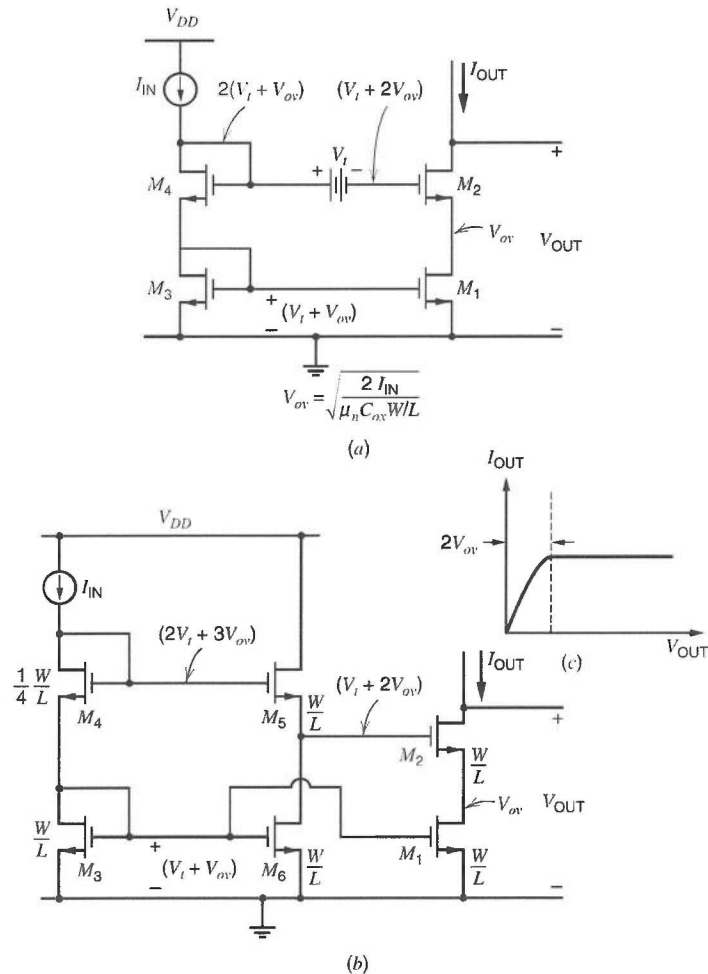


Figure 4.11 (a) MOS cascode current mirror with improved biasing for maximum voltage swing. (b) Practical implementation. (c) I-V characteristic.

boundary between the active and triode regions,

$$V_{DS1} = V_{ov} \quad (4.62)$$

is required. Therefore, the overdrive on M_4 is doubled by reducing its W/L by a factor of four to satisfy (4.62). As a result, the threshold term in (4.59) is eliminated and

$$V_{OUT(min)} \simeq 2V_{ov} \quad (4.63)$$

Because the minimum output voltage does not contain a threshold component, the range of output voltages for which M_1 and M_2 both operate in the active region is significantly improved. Therefore, the current mirror in Fig. 4.11 places much less restriction on the range of output voltages that can be achieved in an amplifier using this current mirror as an active load than the mirror in Fig. 4.9. For this reason, the mirror in Fig. 4.11 is called a *high-swing* cascode current mirror. This type of level shifting to reduce $V_{OUT(min)}$ can also be applied to bipolar circuits.

The output resistance of the high-swing cascode current mirror is the same as in (4.50) when both M_1 and M_2 operate in the active region. However, the input voltage and the systematic gain error are worsened compared to the cascode current mirror without level shift. The input voltage is still given by (4.56), but the overdrive component of the gate-source voltage of M_4 has increased by a factor of two because its W/L has been reduced by a factor of four. Therefore,

$$V_{IN} = 2V_t + 3V_{ov} \quad (4.64)$$

Since M_3 and M_1 form a simple current mirror with unequal drain-source voltages, the systematic gain error is

$$\epsilon = \frac{V_{DS1} - V_{DS3}}{V_A} \simeq \frac{V_{ov1} - (V_t + V_{ov1})}{V_A} = -\frac{V_t}{V_A} \quad (4.65)$$

The negative sign in (4.65) shows that $I_{OUT} < I_{IN}$. For example, if $I_{IN} = 100 \mu\text{A}$, $V_t = 1 \text{ V}$, and $V_A = 10 \text{ V}$, $\epsilon \simeq -0.1$, which means that $I_{OUT} \simeq 90 \mu\text{A}$.

In practice, $(W/L)_4 < (1/4)(W/L)$ is usually selected for two reasons. First, MOS transistors display an indistinct transition from the triode to active regions. Therefore, increasing the drain-source voltage of M_1 by a few hundred millivolts above V_{ov1} is usually required to realize the incremental output resistance predicted by (4.50). Second, although the body effect was not considered in this analysis, it tends to reduce the drain-source voltage on M_1 , which is determined by the following KVL loop

$$V_{DS1} = V_{GS3} + V_{GS4} - V_{GS5} - V_{GS2} \quad (4.66)$$

Each of the gate-source voltage terms in (4.66) contains a threshold component. Since the source-body voltage of M_5 is higher than that of M_4 , $V_{t5} > V_{t4}$. Also, $V_{t2} > V_{t3}$ because the source-body voltage of M_2 is higher than that of M_3 . Simulations with high-accuracy models are usually required to find the optimum $(W/L)_4$.

One drawback of the current mirror in Fig. 4.11 is that the input current is mirrored to a new branch to do the level shift. Combining the input branches eliminates the possibility of mismatch between the two branch currents and may reduce the power dissipation. In a single combined input branch, some element must provide a voltage drop equal to the desired difference between the gate voltages of M_1 and M_2 . To bias M_1 at the edge of the active region, the required voltages from the gates M_1 and M_2 to ground are $V_t + V_{ov}$ and $V_t + 2V_{ov}$, respectively. Therefore, the desired difference in the gate voltages is V_{ov} . This voltage difference can be developed across the drain to the source of a transistor deliberately operated in the triode region, as shown in Fig. 4.12a.⁷ Since M_6 is diode connected, it operates

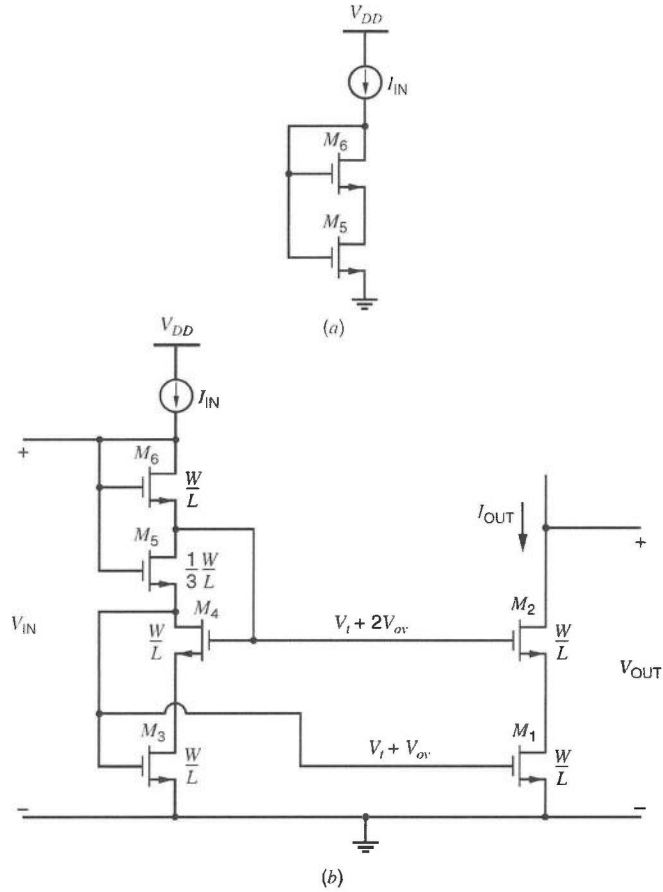


Figure 4.12 (a) Circuit that forces M_5 to operate in the triode region. (b) Sooch cascode current mirror using the circuit in (a).

in the active region as long as the input current and threshold are positive. However, since the gate-source voltage of M_6 is equal to the gate-drain voltage of M_5 , a channel exists at the drain of M_5 when it exists at the source of M_6 . In other words, M_6 forces M_5 to operate in the triode region.

To use the circuit in Fig. 4.12a in a current mirror, we would like to choose the aspect ratios of the transistors so that the drain-source voltage of M_5 is V_{ov} . Since M_6 operates in the active region,

$$I_{IN} = \frac{k'}{2} \left(\frac{W}{L} \right)_6 (V_{GS6} - V_t)^2 \quad (4.67)$$

Since M_5 operates in the triode region,

$$I_{IN} = \frac{k'}{2} \left(\frac{W}{L} \right)_5 \left(2(V_{GS5} - V_t)V_{DS5} - (V_{DS5})^2 \right) \quad (4.68)$$

The goal is to set

$$V_{DS5} = V_{ov} \quad (4.69)$$

when

$$V_{GS6} = V_t + V_{ov} \quad (4.70)$$

From (4.69) and (4.70),

$$V_{GS5} = V_{GS6} + V_{DS5} = V_t + 2V_{ov} \quad (4.71)$$

Substituting (4.68) - (4.71) into (4.67) gives

$$\frac{k'}{2} \left(\frac{W}{L} \right)_6 (V_{ov})^2 = \frac{k'}{2} \left(\frac{W}{L} \right)_5 (2(2V_{ov})V_{ov} - (V_{ov})^2) \quad (4.72)$$

Equation 4.72 can be simplified to

$$\left(\frac{W}{L} \right)_5 = \frac{1}{3} \left(\frac{W}{L} \right)_6 \quad (4.73)$$

The circuit of Fig. 4.12a is used in the current mirror of Fig. 4.12b,⁷ which is called the *Sooch* cascode current mirror after its inventor. At first, ignore transistor M_4 and assume that M_3 is simply diode connected. The difference between the voltages to ground from the gates of M_1 and M_2 is set by the drain-source voltage of M_5 . By choosing equal aspect ratios for all devices except M_5 , whose aspect ratio is given by (4.73), the drain-source voltage of M_5 is V_{ov} and M_1 is biased at the edge of the active region. The output resistance, minimum output voltage, input voltage, and systematic gain error are the same as in (4.50), (4.63), (4.64), and (4.65) respectively.

Now we will consider the effect of transistor M_4 . The purpose of M_4 is to set the drain-source voltage of M_3 equal to that of M_1 . Without M_4 , these drain-source voltages differ by a threshold, causing nonzero systematic gain error. With M_4 ,

$$V_{DS3} = V_{G2} - V_{GS4} \quad (4.74)$$

where

$$V_{G2} = V_{GS3} + V_{DS5} \quad (4.75)$$

Ignoring channel-length modulation,

$$V_{G2} = (V_t + V_{ov}) + V_{ov} = V_t + 2V_{ov} \quad (4.76)$$

Ignoring the body effect and assuming that M_4 operates in the active region,

$$V_{GS4} = V_t + V_{ov} \quad (4.77)$$

Then substituting (4.76) and (4.77) into (4.74) gives

$$V_{DS3} = V_{ov} \quad (4.78)$$

If M_2 also operates in the active region under these conditions, $V_{DS3} = V_{DS1}$. As a result, the systematic gain error is

$$\epsilon = 0 \quad (4.79)$$

Therefore, the purpose of M_4 is to equalize the drain-source voltages of M_3 and M_1 to reduce the systematic gain error.

For M_4 to operate in the active region, $V_{DS4} > V_{ov}$ is required. Since

$$V_{DS4} = V_{GS3} - V_{DS3} = (V_t + V_{ov}) - V_{ov} = V_t \quad (4.80)$$

Equation 4.80 shows that M_4 operates in the active region if $V_t > V_{ov}$. Although this condition is usually satisfied, a low threshold and/or high overdrive may cause M_4 to operate in the triode region. If this happens, the gate-source voltage of M_4 depends strongly on its drain-source voltage, increasing the systematic gain error. Since increasing temperature causes the threshold to decrease, but the overdrive to increase, checking the region of operation of M_4 in simulation at the maximum expected operating temperature is important in practice.

The main limitation of the high-swing cascode current mirrors just presented, is that the input voltage is large. In Fig. 4.11, the input voltage is the sum of the gate-source voltages of M_3 and M_4 and is given by (4.64) ignoring body effect. In Fig. 4.12, the input voltage is

$$\begin{aligned} V_{IN} &= V_{GS3} + V_{DS5} + V_{GS6} \\ &= V_t + V_{ov} + V_{ov} + V_t + V_{ov} \\ &= 2V_t + 3V_{ov} \end{aligned} \quad (4.81)$$

Equation 4.81 shows that the input voltage of the high-swing cascode current mirror in Fig. 4.12 is the same as in (4.64) for Fig. 4.11. The large input voltages may limit the minimum power-supply voltage because a transistor-level implementation of the input current source requires some nonzero drop for proper operation. With threshold voltages of about 1 V, the cascode current mirrors in Figs. 4.11 and 4.12 can operate properly for power-supply voltages greater than about 3 V. Below about 2 V, however, reduced thresholds or a new configuration is required. Reducing the magnitude of the threshold for all transistors increases the difficulty in turning off transistors that are used as switches. This problem can be overcome by using low-threshold devices in the current mirror and high-threshold devices as switches, but this solution increases process complexity and cost. Therefore, circuit techniques to reduce the input voltage are important to minimize cost.

To reduce the input voltage, the input branch can be split into two branches, as shown in Fig. 4.13. If M_1 and M_2 are biased in the active region, the output resistance is still given by (4.50). Also, the minimum output voltage for which (4.50) applies is still given by (4.63). Furthermore, if M_4 operates in the active region, the drain-source voltage of M_3 is equal to that of M_1 , and the systematic gain error is still zero as in (4.79).

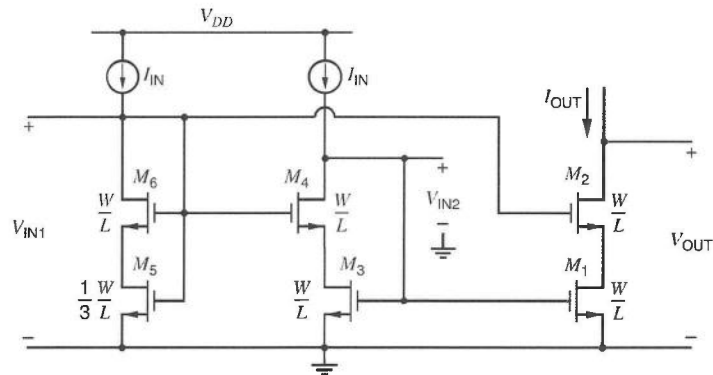


Figure 4.13 MOS high-swing current mirror with two input branches.

Since the mirror in Fig. 4.13 has two input branches, an input voltage can be calculated for each:

$$V_{IN1} = V_{DS5} + V_{GS6} = V_I + 2V_{ov} \quad (4.82)$$

$$V_{IN2} = V_{GS3} = V_I + V_{ov} \quad (4.83)$$

Both V_{IN1} and V_{IN2} are less than the input voltage given in (4.64) for Fig. 4.12b by more than a threshold, allowing the input current sources to operate properly with power-supply voltages greater than about 2 V, assuming thresholds of about 1 V.

Finally, in Fig. 4.13, the drain-source voltage of M_5 is only used to bias the source of M_6 . Therefore, M_5 and M_6 can be collapsed into one diode-connected transistor whose source is grounded. Call this replacement transistor M_7 . The aspect ratio of M_7 should be a factor of four smaller than the aspect ratios of M_1 - M_4 to maintain the bias conditions as in Fig. 4.13. In practice, the aspect ratio of M_7 is further reduced to bias M_1 past the edge of the active region and to overcome a mismatch in the thresholds of M_7 and M_2 caused by body effect.

4.2.6 Wilson Current Mirror

4.2.6.1 Bipolar

The main limitation of the bipolar cascode current mirror is that the systematic gain error stemming from finite β_F was large, as given in (4.49). To overcome this limitation, the Wilson current mirror can be used as shown in Fig. 4.14a.⁸ This circuit uses negative feedback through Q_1 , activating Q_3 to reduce the base-current error and raise the output resistance. (See Chapter 8.)

From a qualitative standpoint, the difference between the input current and I_{C3} flows into the base of Q_2 . This base current is multiplied by $(\beta_F + 1)$ and flows in the diode-connected transistor Q_1 , which causes current of the same magnitude to flow in Q_3 . A feedback path is thus formed that regulates I_{C3} so that it is nearly equal to the input current, reducing the systematic gain error caused by finite β_F . Similarly, when the output voltage increases, the collector current of Q_2 also increases, in turn increasing the collector current of Q_1 . As a result, the collector current of Q_3 increases, which reduces the base current of Q_2 . The decrease in the base current of Q_2 caused by negative feedback reduces the original change in the collector current of Q_2 and increases the output resistance.

To find the output resistance of the Wilson current mirror when all transistors operate in the active region, we will analyze the small-signal model shown in Fig. 4.14b, in which a

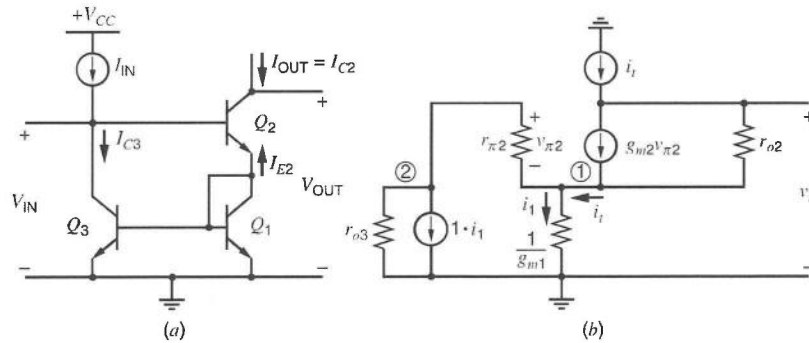


Figure 4.14 (a) Bipolar Wilson current mirror. (b) Small-signal model.

test current source i_t is applied at the output. Transistors Q_1 and Q_3 form a simple current mirror. Since Q_1 is diode connected, the small-signal resistance from the base of Q_1 to ground is $(1/g_{m1}) || r_{\pi1} || r_{\pi3} || r_{o1}$. Assume that an unknown current i_1 flows in this resistance. When $g_{m1}r_{\pi1} \gg 1$, $g_{m1}r_{\pi3} \gg 1$, and $g_{m1}r_{o1} \gg 1$, this resistance is approximately equal to $1/g_{m1}$. Transistor Q_3 could be modeled as a voltage-controlled current source of value $g_{m3}v_{\pi3}$ in parallel with r_{o3} . Since $v_{\pi3} = v_{\pi1} \simeq i_1/g_{m1}$, the voltage-controlled current source in the model for Q_3 can be replaced by a current-controlled current source of value $(g_{m3}/g_{m1})(i_1) = \beta(i_1)$, as shown in Fig. 4.14b. This model represents the behavior of the simple current mirror directly: The input current i_1 is mirrored to the output by the current-controlled current source.

Using this model, the resulting voltage v_t is

$$v_t = \frac{i_1}{g_{m1}} + (i_t - g_{m2}v_{\pi2})r_{o2} \quad (4.84)$$

To find the relationship between i_1 and $v_{\pi2}$, note that the voltage across r_{o3} is $(i_1/g_{m1} + v_{\pi2})$ and use KCL at node ② in Fig. 4.14b to show that

$$\frac{v_{\pi2}}{r_{\pi2}} + i_1 + \frac{\frac{i_1}{g_{m1}} + v_{\pi2}}{r_{o3}} = 0 \quad (4.85)$$

Rearranging (4.85) gives

$$v_{\pi2} = -i_1 r_{\pi2} \left(\frac{1 + \frac{1}{g_{m1}r_{o3}}}{1 + \frac{r_{\pi2}}{r_{o3}}} \right) \quad (4.86)$$

To find the relationship between i_1 and i_t , use KCL at node ① in Fig. 4.14b to show that

$$i_t = i_1 - \frac{v_{\pi2}}{r_{\pi2}} \quad (4.87)$$

Substituting (4.86) into (4.87) and rearranging gives

$$i_1 = \frac{i_t}{1 + \left(\frac{1 + \frac{1}{g_{m1}r_{o3}}}{1 + \frac{r_{\pi2}}{r_{o3}}} \right)} \quad (4.88)$$

Substituting (4.88) into (4.86) and rearranging gives

$$v_{\pi2} = -i_t r_{\pi2} \left(\frac{1 + \frac{1}{g_{m1}r_{o3}}}{2 + \frac{r_{\pi2}}{r_{o3}} + \frac{1}{g_{m1}r_{o3}}} \right) \quad (4.89)$$

Substituting (4.88) and (4.89) into (4.84) and rearranging gives

$$R_o = \frac{v_t}{i_t} = \frac{1}{g_{m1} \left[1 + \left(\frac{1 + \frac{1}{g_{m1}r_{o3}}}{1 + \frac{r_{\pi2}}{r_{o3}}} \right) \right]} + r_{o2} + \frac{g_{m2}r_{\pi2}r_{o2} \left(1 + \frac{1}{g_{m1}r_{o3}} \right)}{2 + \frac{r_{\pi2}}{r_{o3}} + \frac{1}{g_{m1}r_{o3}}} \quad (4.90)$$

If $r_{o3} \rightarrow \infty$, the small-signal current that flows in the collector of Q_3 is equal to i_1 and (4.90) reduces to

$$R_o = \frac{1}{g_{m1(2)}} + r_{o2} + \frac{g_{m2}r_{\pi2}r_{o2}}{2} \simeq \frac{\beta_0 r_{o2}}{2} \quad (4.91)$$

This result is the same as (4.38) for the cascode current mirror. In the cascode current mirror, the small-signal current that flows in the base of Q_2 is mirrored through Q_3 to Q_1 so that the small-signal base and emitter currents leaving Q_2 are approximately equal. On the other hand, in the Wilson current mirror, the small-signal current that flows in the emitter of Q_2 is mirrored through Q_1 to Q_3 and then flows in the base of Q_2 . Although the cause and effect relationship here is opposite of that in a cascode current mirror, the output resistance is unchanged because the small-signal base and emitter currents leaving Q_2 are still forced to be equal. Therefore, the small-signal collector current of Q_2 that flows because of changes in the output voltage still splits into two equal parts with half flowing in $r_{\pi2}$.

For the purpose of dc analysis, we assume that $V_A \rightarrow \infty$ and that the transistors are identical. Then the input voltage is

$$V_{IN} = V_{CE3} = V_{BE1} + V_{BE2} = 2V_{BE(on)} \quad (4.92)$$

which is the same as in (4.40) for a cascode current mirror. Also, the minimum output voltage for which both transistors in the output branch operate in the forward-active region is

$$V_{OUT(min)} = V_{CE1} + V_{CE2(sat)} = V_{BE(on)} + V_{CE2(sat)} \quad (4.93)$$

The result in (4.93) is the same as in (4.41) for a cascode current mirror.

To find the systematic gain error, start with KCL at the collector of Q_1 to show that

$$-I_{E2} = I_{C1} + I_{B1} + I_{B3} = I_{C1} \left(1 + \frac{1}{\beta_F} \right) + \frac{I_{C3}}{\beta_F} \quad (4.94)$$

Since we assumed that the transistors are identical and $V_A \rightarrow \infty$,

$$I_{C3} = I_{C1} \quad (4.95)$$

Substituting (4.95) into (4.94) gives

$$-I_{E2} = I_{C1} \left(1 + \frac{2}{\beta_F} \right) \quad (4.96)$$

Using (4.96), the collector current of Q_2 is then

$$I_{C2} = -I_{E2} \left(\frac{\beta_F}{1 + \beta_F} \right) = I_{C1} \left(1 + \frac{2}{\beta_F} \right) \left(\frac{\beta_F}{1 + \beta_F} \right) \quad (4.97)$$

Rearranging (4.97) we obtain

$$I_{C1} = I_{C2} \left[\frac{1}{\left(1 + \frac{2}{\beta_F} \right) \left(\frac{\beta_F}{1 + \beta_F} \right)} \right] \quad (4.98)$$

From KCL at the base of Q_2 ,

$$I_{C3} = I_{IN} - \frac{I_{C2}}{\beta_F} \quad (4.99)$$

Inserting (4.98) and (4.99) into (4.95), we find that

$$I_{OUT} = I_{C2} = I_{IN} \left(1 - \frac{2}{\beta_F^2 + 2\beta_F + 2} \right) = \frac{I_{IN}}{1 + \frac{2}{\beta_F(\beta_F + 2)}} \quad (4.100)$$

In the configuration shown in Fig. 4.14a, the systematic gain error arising from finite output resistance is not zero because Q_3 and Q_1 operate with collector-emitter voltages that differ by the base-emitter voltage of Q_2 . With finite V_A and finite β_F ,

$$\begin{aligned} I_{OUT} &\simeq I_{IN} \left(1 - \frac{2}{\beta_F^2 + 2\beta_F + 2} \right) \left(1 + \frac{V_{CE1} - V_{CE3}}{V_A} \right) \\ &\simeq I_{IN} \left(1 - \frac{2}{\beta_F^2 + 2\beta_F + 2} \right) \left(1 - \frac{V_{BE2}}{V_A} \right) \end{aligned} \quad (4.101)$$

Therefore, the systematic gain error is

$$\epsilon \simeq - \left(\frac{2}{\beta_F^2 + 2\beta_F + 2} + \frac{V_{BE2}}{V_A} \right) \quad (4.102)$$

Comparing (4.102) to (4.49) shows two key points. First, the systematic gain error arising from finite β_F in a Wilson current mirror is much less than in a cascode current mirror. Second, the systematic gain error arising from finite output resistance is worse in the Wilson current mirror shown in Fig. 4.14a than in the cascode current mirror shown in Fig. 4.9. However, this limitation is not fundamental because it can be overcome by introducing a new diode-connected transistor between the collector of Q_3 and the base of Q_2 to equalize the collector-emitter voltages of Q_3 and Q_1 .

4.2.6.2 MOS

Wilson current mirrors are also used in MOS technology, as shown in Fig. 4.15. Ignoring M_4 , the circuit operation is essentially identical to the bipolar case with $\beta_F \rightarrow \infty$. One way to calculate the output resistance is to let $r_{\pi 2} \rightarrow \infty$ in (4.90), which gives

$$R_o = \frac{1}{g_{m1}} + r_{o2} + g_{m2}r_{o2} \left(1 + \frac{1}{g_{m1}r_{o3}} \right) r_{o3} \simeq (1 + g_{m2}r_{o3})r_{o2} \quad (4.103)$$

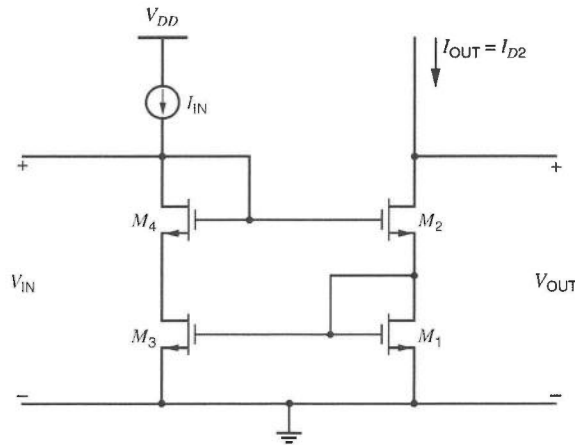


Figure 4.15 Improved MOS Wilson current mirror with an additional device such that the drain voltages of M_1 and M_3 are equal.

Since the calculation in (4.103) is based on the small-signal model for the bipolar Wilson current mirror in Fig. 4.14b, it ignores the body effect in transistor M_2 . Repeating the analysis with a body-effect generator in parallel with r_{o2} gives

$$R_o \simeq (2 + g_{m2}r_{o3})r_{o2} \quad (4.104)$$

The body effect on M_2 has little effect on (4.104) because M_1 is diode connected and therefore the voltage from the source of M_2 to ground is almost constant.

Although $\beta_F \rightarrow \infty$ for MOS transistors, the systematic gain error is not zero without M_4 because the drain-source voltage of M_3 differs from that of M_1 by the gate-source voltage of M_2 . Therefore, without M_4 ,

$$\epsilon = \frac{V_{DS1} - V_{DS3}}{V_A} = -\frac{V_{GS2}}{V_A} \quad (4.105)$$

Transistor M_4 is inserted in series with M_3 to equalize the drain-source voltages of M_3 and M_1 so that

$$\epsilon \simeq 0 \quad (4.106)$$

With M_4 , the output resistance is still given by (4.104) if all transistors operate in the active region. Also, insertion of M_4 does not change either the minimum output voltage for which (4.104) applies or the input voltage. Ignoring body effect and assuming equal overdrives on all transistors, the minimum output voltage is

$$V_{OUT(min)} = V_{GS1} + V_{ov2} = V_t + 2V_{ov} \quad (4.107)$$

Under the same conditions, the input voltage is

$$V_{IN} = V_{GS1} + V_{GS2} = 2V_t + 2V_{ov} \quad (4.108)$$

4.3 Active Loads

4.3.1 Motivation

In differential amplifiers of the type described in Chapter 3, resistors are used as the load elements. For example, consider the differential amplifier shown in Fig. 3.45. For this circuit, the differential-mode (dm) voltage gain is

$$A_{dm} = -g_m R_C \quad (4.109)$$

Large gain is often desirable because it allows negative feedback to make the gain with feedback insensitive to variations in the parameters that determine the gain without feedback. This topic is covered in Chapter 8. In Chapter 9, we will show that the required gain should be obtained in as few stages as possible to minimize potential problems with instability. Therefore, maximizing the gain of each stage is important.

Multiplying the numerator and denominator of (4.109) by I gives

$$A_{dm} = -\frac{I(R_C)}{I/g_m} \quad (4.110)$$

With bipolar transistors, let I represent the collector current I_C of each transistor in the differential pair. From (1.91), (4.110) can be rewritten as

$$A_{dm} = -\frac{I_C R_C}{V_T} \quad (4.111)$$

Index

- Abrupt junction, 4
- Active cascode, 210
- Active-cascode op amp, 446
- Active-device parameter summary, 73
- Active level shift, 439, 458, 469, 485
- Active load, 276, 421, 448, 457, 546
 - common-emitter amplifier, 277, 280, 284
 - common-mode rejection ratio, 291
 - common-source amplifier, 277, 280, 282
 - complementary, 277
 - current-mirror, 285
 - depletion, 280
 - diode-connected, 282
 - enhancement, 284
 - noise, 776
 - offset voltage, 330
- Advanced bipolar integrated-circuit
 - fabrication, 92, 106
- All-*npn* output stage, 375
- Amplifier:
 - current, 557
 - fully differential, 799, 811
 - operational, *see* Operational amplifier
 - transconductance, 562
 - transresistance, 560
 - voltage, 557
 - wideband, 514, 527, 575, 617
- Amplifier stages:
 - Class A, 349, 576
 - Class AB, 361, 461, 688
 - Class B, 362
 - multiple-transistor, 201
 - single-transistor, 173
- Analog multiplier, 708, 711
- Anneal, 87
- Avalanche breakdown, 7, 21, 49, 71
- Avalanche noise, 743
- Average power, 345

- Balanced circuit, 796
- Balanced differential amplifier, 799
- Balanced modulator, 712
- Band-gap reference, 315, 321
 - curvature compensated, 318
- Bandwidth of feedback amplifiers, 624
- Base-charging capacitance, 27
- Base-diffused resistors, 118
- Base diffusion, 86
- Base-emitter voltage temperature coefficient, 11, 315
- Base resistance, 30, 98
- Base transport factor, 13
- Base width, 12
- Beta, *see* Current gain
- Bias reference circuit:
 - band-gap, 315, 321
 - bootstrap, 307
 - current routing, 329
 - low-current, 297
 - proportional to absolute temperature, 321
 - self-biased, 307
 - start-up circuit, 308, 312
 - supply-insensitive, 303
 - temperature-insensitive, 315
 - threshold-referenced, 307
 - V_{BE} -referenced, 308, 311
 - voltage routing, 329
 - V_T -referenced, 314, 315
- BiCMOS, 152
 - amplifier, 249, 618, 794
 - cascode, 209
 - Darlington, 204, 791
 - output stage, 377, 398
 - technology, 152
- Bilateral amplifier, 172
- Bipolar transistor:
 - advanced technology, 95, 106
 - base-charging capacitance, 27
 - base resistance, 30, 98
 - base transit time, 28
 - breakdown voltage, 20, 86
 - in CMOS technology, 144, 151
 - collector-base resistance, 30
 - collector series resistance, 89, 100
 - current-mirror matching, 325
 - cutoff region, 146
 - diffusion profile, 91
 - fabrication, 88, 92
 - forward-active region, 45
 - frequency response, 34
 - heterojunction, 153

- Bipolar transistor (*Continued*)
 - input resistance, 28
 - inverse-active region, 16
 - large-signal models, 8
 - lateral, 140
 - noise, 745, 750
 - output characteristics, 15
 - output resistance, 29
 - parameter summary (functional), 73
 - parameter summary (numerical), 106, 108, 112, 114
 - parasitic capacitance, 31, 106, 111, 116
 - reverse-active region, *see* Inverse-active region
 - saturation current, 18, 97, 316
 - saturation region, 17
 - self-aligned structure, 94
 - small-signal model, 25
 - SPICE model parameters, 162
 - transconductance, 26
- Bird's beak, 141, 257, 426
- Blackman's impedance formula, 608
- Body effect in MOS transistors, 53, 139
- Boltzmann approximation, 10
- Bootstrap bias technique, 307
- Breakdown, 6
 - in bipolar transistors, 38
 - in MOSFETs, 49
 - in superbeta transistors, 124
- Breakdown voltage:
 - base-emitter, 22
 - collector-base, 20, 86
 - collector-emitter, 21
 - drain-substrate, 49
 - junction-diode, 7
 - MOS transistor, 45
 - superbeta transistors, 124
 - Zener diode, 8
- Built-in potential, 2
- Buried layer, 89
- Burst noise, 745
- Capacitance:
 - base-charging, 27
 - base-collector, 31, 97
 - base-emitter, 31, 96
 - channel-substrate, 138
 - collector-substrate, 31, 103
 - depletion region, 5, 31, 54, 103, 140, 530
 - drain-body, 54
 - gate-body, 55
 - gate-drain, 51
 - gate-source, 51
 - gate-substrate, *see* Capacitance: gate-body overlap, 55, 142
 - sidewall, 143
 - source-body, 54
- Capacitive neutralization, 835
- Capacitors:
 - in bipolar integrated circuits, 115
 - lateral, 150
 - in MOS integrated circuits, 148
 - MOS transistors, 148
 - poly-poly, 146
 - vertical, 146
- Capture range, 718, 726
- Cascode configuration, 205, 207, 439, 442, 446, 527
 - active, 210, 446
- Cascode current mirror, 261, 690
 - Bipolar, 261
 - MOS, 264
- Cascode frequency response, 527
- Channel-length modulation, 43
- Class A output stage, 349, 576
- Class AB input stage, 688
- Class AB output stage, 364
- Class B output stage, 361
- Clipping, 346, 361
- Closed-loop gain, 401, 554, 601
- Clubhead, 116
- CMFB, *see* Common-mode feedback
- CMOS operational amplifier:
 - fully differential, 834
 - with single-ended output, 421, 439, 442, 446
- CMOS output stage, 379
- CMOS technology, 46, 128
- Collector-base capacitance, 102
- Collector-base resistance, 30
- Collector-series resistance, 32, 100
- Common-base configuration, 182, 184, 185, 514, 529
- Common-base frequency response, 514
- Common-base stage noise performance, 771
- Common centroid geometry, 437
- Common-collector-common-collector configuration, 201
- Common-collector-common-emitter configuration, 201

- Common-collector configuration, 191, 505
- Common-drain configuration, 194, 511, 591
- Common-emitter-common-base configuration, *see* cascode configuration
- Common-emitter configuration, 174, 490, 495
- Common-emitter frequency response, 490, 498
- Common-emitter output stage, 576
- Common-gate configuration, 185, 187, 514
- Common-gate frequency response, 514
- Common-mode feedback, 227, 286, 458, 804
 - circuits, 796
 - using a resistive divider, 812
 - using switched capacitors, 821
 - using transistors in the triode region, 812
 - using two differential pairs, 816
- Common-mode feedback loop, 804
 - bandwidth, 844
 - stability and compensation, 810
- Common-mode gain, 222, 804
- Common-mode half-circuit, 226, 239, 804
- Common-mode input range, 854
- Common-mode input resistance, 228
- Common-mode rejection ratio, 223, 228, 234, 417, 427
 - of active-load stage, 293
 - frequency response, 501
- Common-mode-to-differential-mode gain, 222
- Common-source amplifier, 181, 490, 499
- Common-source-common-gate configuration, *see* cascode configuration
- Common-source frequency response, 490, 498, 500
- Compensation:
 - of amplifiers, 633
 - capacitor, 597, 637, 682
 - by feedback zero, 677
 - methods, 637
 - of MOS amplifiers, 643, 650
 - of NE5234 op amp, 658
 - nested Miller method, 654
 - theory, 633
- Complementary load, 277
- Complementary output stage, 360
- Composite *pnp*, 377
- Conductivity, 80
- Copper, 156
- Correlation, 747, 756, 789
- Critical field, 7, 59
- Crossover distortion, 362, 365, 368, 375, 384
- Crossunder, 117
- Crowding, 32, 99
- Current amplifier, 562
- Current crowding, 32, 99
- Current density, 425
- Current-feedback pair, 586, 677, 768
- Current gain:
 - dependence on operating conditions, 23, 110
 - forced, 18
 - forward, 12
 - inverse, 19
 - nnp* transistor, 105
 - pnp* transistor, 107
 - short-circuit, 177
 - small-signal, 29
 - small-signal high-frequency, 34
 - temperature coefficient, 23
- Current mirror, 251
 - with beta helper, 258
 - cascode, 261, 264
 - current routing, 329
 - with degeneration, 260, 261
 - gain error, 252
 - general properties, 251
 - high-swing cascode, 268
 - input voltage, 252
 - load, 285
 - matching, 325
 - minimum output voltage, 255
 - output resistance, 252, 254, 257, 258, 260, 261, 264, 266, 272, 275
 - simple, 253, 255
 - Sooch cascode, 269, 270
 - voltage routing, 329
 - Wilson, 272, 275, 595, 610
- Current routing, 329
- Current source, 297, 300, 301, 302, 303, 304, 307
 - Peaking, 301, 302
 - Widlar, 297, 300, 304
- Cutoff, 46
- Darlington configuration, 201, 203, 397, 595, 791

- dc analysis of the NE5234 op amp, 452
- Deadband, 361, 368, 375
- Depletion-mode load, 281
- Depletion region:
 - capacitance, 5, 31, 54, 103, 530
 - collector-base, 14, 89
 - of *pn* junction, 1
- Deviation from ideality in real op amps, 415
- Die, 91
- Dielectric isolation, 123
- Differential amplifier, 220, 404
 - perfectly balanced, 222
 - unbalanced, 237
- Differential-mode gain, 222, 804
- Differential-mode half-circuit, 225, 239, 804
- Differential-mode input resistance, 228
- Differential-mode-to-common-mode gain, 222, 838
- Differential pair:
 - bipolar transistor, 214
 - MOS, 217
 - noise performance, 773
 - with current-mirror load, 285, 534
- Differential signal source, 800
- Differential-to-single-ended conversion, 291
- Differentiator, 406
- Diffused-layer sheet resistance, 82
- Diffused resistors, 115
- Diffusion constant, 10
- Diffusion current, 10
- Diffusion of impurities, 80
- Diffusion profile of a bipolar transistor, 91, 92, 96, 98
- Diode:
 - Junction, 121
 - Zener, *see* Zener diode
- Diode-connected load, 282
- Diode-connected MOS transistor, 255
- Distortion, 358, 399
- Distortion in the source follower, 355
- Distortion reduction by negative feedback, 553
- Dominant pole, 495, 497, 499, 518, 522, 524
- Doping, 79
- Double-diffused *nnp* transistors, 95
- Double-diffused *pnp* transistors, 126
- Drift, emitter-coupled pair, 234
 - source-coupled pair, 238
- Driver stage, 343, 351, 366, 367, 369, 370, 371, 389, 395
- Early effect, 16, 21, 262
- Early voltage, 15, 29, 43, 52, 255, 258
- Ebers-Moll equations, 19
- Economics of IC fabrication, 156
- Effective channel length, 43, 140
- Effective channel width, 141
- Efficiency:
 - Class A, 347, 349, 350
 - Class B, 360, 362, 366, 375
- Electromigration, 156
- Emitter-coupled multivibrator, 728
- Emitter-coupled pair, 214, 704, 773
 - emitter degeneration, effect of, 216
 - frequency response, 495
 - input offset current, 230, 234
 - input offset voltage, 230, 231
 - offset voltage drift, 233
 - small-signal analysis, 220, 237
 - transfer characteristics, 214
- Emitter degeneration, 196, 216, 260, 587
- Emitter-diffused resistors, 117
- Emitter-follower, 191
 - drive requirements, 351
 - frequency response, 505
 - noise performance, 773
 - output stage, 341, 344, 576
 - small-signal properties, 192, 352
 - terminal impedances, 507
 - transfer characteristics, 341
- Emitter injection efficiency, 13
- Emitter resistance, 32
- Enhancement load, 284
- Epitaxial growth, 86
- Epitaxial resistor, 119
- Equivalent circuit of an op amp, 420
- Equivalent input noise generators, 756
- Equivalent input noise resistance, 758
- Equivalent input shot noise current, 759
- Etching, 85
- Fabrication of integrated circuits, 78
- Feedback, 403, 553
 - bandwidth, 624
 - common mode, 227, 286, 458, 804
 - configurations, 557
 - effect on distortion, 555

- effect on gain sensitivity, 555
- effect on noise, 764
- effect on terminal impedances, 558, 561, 562, 563, 607
- ideal analysis, 553
- latchup, 151
- loading, effect of, 563
- local, 227, 263
- loop gain, 384, 401, 554
- practical configurations, 563
- return ratio, 599, 612
- series-series, 562, 569
- series-shunt, 557, 579
- shunt-series, 561, 583
- shunt-shunt, 560, 563
- single-stage, 587
- table of relationships, 588
- Feedback-zero compensation, 677
- Feedforward, 389, 644–650, 660
- Field-effect transistor, *see* MOS transistor
- Field region, 128
- First-order phase-locked loop, 720
- Flicker noise, 741
- Flicker noise corner frequency, 746
- Folded-cascode op amp, 446, 652
- Forward-active region, 8, 26, 45, 112, 175
- Free-running frequency, 717, 726
- Frequency response:
 - bipolar transistor, 33, 56
 - cascode, 527
 - common-base, 514
 - common-drain, 511
 - common-emitter, 495
 - common-emitter cascode, 527
 - common-gate, 514
 - common-mode, 501
 - common-source, 499
 - common-source cascode, 527
 - of CMRR, 503
 - of current mirror, 251, 256
 - differential-mode, 495, 496, 500
 - of differential pair with current-mirror load, 290, 534
 - emitter follower, 505
 - MOS amplifier, 643, 650
 - MOS transistor, 55
 - multistage amplifier, 518
 - NE5234 op amp, 539
 - single stage, 490
 - source follower, 511
- Full-power bandwidth, 692
- Fully differential amplifiers, 796
 - cross gains, 838
- Fully differential op amps, 796, 823, 845
 - folded cascode, 834
 - telescopic cascode, 855
 - two-stage, 823
- Gain-bandwidth product, 625
- Gain margin, 630
- Gaussian distribution, 232, 244, 738
- Generation, 79
- Gilbert multiplier, 704
- g*-parameters, 583
- Graded junction, 5
- Guard ring, 152
- Half circuit:
 - of balanced amplifier, 225, 226, 799
 - coupled, 237
 - of unbalanced amplifier, 239, 240
 - independent, 238
- Harmonic distortion, 357, 399
- Heterojunction, 153
- High-frequency equivalent circuit of the NE5234 op amp, 540, 659
- High-level injection, 36, 110
- High-voltage integrated-circuit fabrication, 88
- Homojunction, 53
- Hot carriers, 49
- h*-parameter, 579
- Hybrid- π model, 30, 52
- Ideal feedback analysis, 553
- Impact ionization, 71, 209
- Impurity concentration, 79
- Instability in amplifiers, 626, 666
- Instantaneous power, 345–351, 365
- Integrated circuit:
 - advanced bipolar fabrication, 92, 107
 - biasing, 251
 - capacitors, 120, 146
 - cost considerations, 156, 159
 - current mirrors, 251
 - device models, 1
 - economics, 156
 - fabrication, 78
 - fabrication yield, 157
 - frequency response, 490

- Integrated circuit (*Continued*)
 - mixed-signal, 221
 - noise, 736
 - op amps, 402, 796, 823
 - output stages, 341
 - packaging, 159
 - passive components, 115, 146
 - phase-locked loop, 727
 - resistors, 115, 119, 146
- Integrator:
 - continuous-time, 406
 - switched-capacitor, 412
- Interconnect delay, 156
- Internal amplifier, 407
- Intrinsic carrier concentration, 2, 79
- Inverse-active region of bipolar transistors, 19, 45
- Inversion, 40
- Inverting amplifier, 402
- Ion implantation, 85, 87
- Isolation diffusion, 89, 90
- Junction breakdown, 6, 49
- Junction diodes, 122
- KCL, 184
- Kirk effect, 25, 36
- KVL, 184
- Large-signal model:
 - bipolar transistor, 8
 - MOS transistor, 45
- Latchup in CMOS, 151, 341
- Lateral bipolar transistor, 146
- Lateral *pnp* transistor, 108, 368, 376
- Layout, 435
- Leakage current, 7, 18, 19
 - collector base, 19
- Level shifting, 194, 196, 439, 485
- Lifetime of minority carriers, 11
- Line regulation, 595
- Load line, 345–351, 364, 365
- Load regulation, 595
- Local feedback, 587
- Local oxidation, 87
- Lock range, 717, 724
- Logarithmic amplifier, 405
- Loop gain, 384, 401, 552
- Low-current biasing, 297
- Low-level injection, 10, 110
- Low-permittivity dielectric, 156
- Matching in transistor current mirrors, 325
- Miller capacitance, 491, 498, 500, 638, 641, 831
- Miller effect, 114, 490, 492, 497, 499, 828
- Minimum detectable signal, 754
- Minority-carrier lifetime, 11
- Mismatch effects in differential amplifiers, 229
- Moat, 93
- Moat etch, 123
- Mobility, 80
- Mobility degradation, 65
- Model selection for IC analysis, 170
- Modulator, 712
- MOS transistor:
 - active region, 45
 - bird's beak, 141, 257, 426
 - body effect, 53, 139
 - breakdown voltage, 149
 - channel-length modulation, 43
 - complementary MOS (CMOS), 39, 46, 127
 - critical field, 59
 - current density, 425
 - current-mirror matching, 328
 - cutoff region, 46
 - depletion mode, 47, 144, 281
 - diode connected, 255
 - effective channel length, 43, 140
 - effective channel width, 141
 - enhancement-mode, 38, 127
 - fabrication, 127
 - field-effect transistor, 41
 - field region, 128
 - frequency response, 55
 - hot carriers, 49
 - impact ionization, 71
 - input resistance, 52
 - inversion, 40
 - junction breakdown, 49
 - large-signal model, 45
 - mobility, 42
 - mobility degradation, 65, 135
 - moderate inversion, 69
 - n*-channel, 38, 131, 144
 - noise, 750, 762
 - ohmic region, 44

- operational amplifier, 421
- output resistance, 52
- output stages, 379
- overdrive, 47, 181, 434
- oxide breakdown, 49
- parameter summary (functional), 73
- parameter summary (numerical), 132, 133, 134, 135
- parasitic elements, 54
- p*-channel, 39, 144
- punchthrough, 49
- saturation region, 44, 51
- short-channel effects, 59
- silicon gate, 129
- small-signal model, 49
- source-coupled pairs, 217
- source follower, 194
- SPICE model parameters, 162
- strong inversion, 65
- substrate current, 71
- subthreshold conduction, *see* Weak inversion
- threshold temperature dependence, 47
- threshold voltage, 40, 129, 131
- transconductance, 50, 63
- transfer characteristic, 38
- triode region, 44
- weak inversion, 65, 180
- velocity saturation, 59
- Multiplication factor, 7
- Multiplier circuits, 704
- Multistage amplifier frequency response, 518
- Negative feedback, *see* Feedback
- Neutralization, 835
- NMOS, 38
- Noise:
 - in active loads, 776
 - amplitude distribution, 738
 - avalanche, 743
 - bandwidth, 776
 - burst, 742
 - in capacitors and inductors, 747
 - circuit calculations, 750
 - common-base stage, 771
 - differential pair, 773
 - in diodes, 736
 - in direct-coupled amplifiers, 785
 - emitter follower, 773
 - equivalent input noise, 754
 - $1/f$, 741
 - feedback, effect of, 764
 - figure, 786
 - flicker noise, 741
 - generator, 738
 - models, 744
 - MOS amplifier, 746, 762, 779
 - in NE5234 op amp, 776
 - operational amplifier, 776, 855
 - popcorn noise, 742
 - in resistors, 747
 - shot, 736
 - spectral density, 737, 748
 - temperature, 786
 - thermal, 740
 - white, 737
 - in Zener diodes, 743
- Nondominant pole, 495, 497, 500, 536
 - of NE5234 op amp, 542
- Noninverting amplifier, 404
- Nonlinear analog operations, 405
- Nonlinearity, 343, 352, 353, 357, 368, 375, 393, 395, 708
- Nonoverlapping clock signals, 409, 412
 - charge-transfer phase, 409
 - input sample phase, 409
- Normal distribution, *see* Gaussian distribution
- Nyquist criterion, 626, 666
- Nyquist diagram, 627
- Offset current, 416
 - emitter-coupled pair, 231, 233
 - NE5234 op amp, 477, 481
- Offset voltage, 424
 - differential pair with active load, 330
 - emitter-coupled pair, 231, 233
 - emitter-coupled pair with active load, 330
 - NE5234 op amp, 477
 - source-coupled pair, 235
 - source-coupled pair with active load, 337
- Open-circuit time constants, 519, 536
 - also see* Zero-value time constants
- Open-loop gain, 401
- Operational amplifier:
 - active-cascode, 446
 - applications, 401
 - bipolar, 448
 - with cascodes, 438

Operational amplifier (*Continued*)

- CMOS, 421
- common-mode input range, 416, 427
- common-mode rejection ratio, 417, 427
- compensation, 420, 633, 676
- deviations from ideality, 415
- differential amplifier, 404
- effect of overdrive voltages, 434
- equivalent circuit, 420
- folded-cascode, 442, 652, 845
- frequency response, 420
- fully differential op amps, 796, 823
- input bias current, 415
- input resistance, 420, 422
- internal amplifier, 407
- inverting amplifier, 402
- layout considerations, 435
- MOS, 421, 438, 442, 446
- NE5234 op amp analysis, 452
- noise, 776
- offset current, 416
- offset drift, 416
- offset voltage, 416, 424
- open-circuit voltage gain, 422
- output buffer, 407
- output resistance, 420, 422
- output swing, 423
- power-supply rejection ratio, 418, 430, 435
- random offset voltage, 426
- settling time, 411
- with single-ended outputs, 400
- slew rate, 420, 681
- small-signal characteristics, 422
- supply capacitance, 432
- systematic offset voltage, 424
- telescopic-cascode, 439, 650
- Oscillation, 624, 666, 679
- Output buffer, 407
- Output resistance of current mirror, 252, 254, 257, 258, 260, 261, 264, 266, 272, 275
- Output stage, 341
 - all-*n**p**n*, 375
 - BiCMOS, 377, 397
 - Class A, 349, 576
 - Class AB, 361, 379
 - Class B, 360
 - CMOS, 379
 - combined common-drain common-source, 388, 390
 - common-drain, 353, 380
 - common emitter, 576
 - common-source, with error amplifiers, 381
 - complementary, 360
 - Darlington, 397
 - emitter follower, 341
 - NE5234 op amp, 458, 474
 - overload protection, 377
 - parallel common-source, 390
 - push-pull, 359
 - quasi-complementary, 376, 381, 384, 388
 - 709 op amp, 366
 - 741 op amp, 368
 - short-circuit protection, 377
 - source follower, 353, 380
- Overdrive, 47
- Overload protection, 377
- Oxidation, 84
- Oxide breakdown, 49
- Packaging considerations, 159
- Parasitic elements:
 - bipolar transistor, 31, 98
 - MOS transistor, 54, 141
- Passive components:
 - in bipolar integrated circuits, 115
 - in MOS technology, 144
- Permittivity, 3
- Phase detector, 712, 716, 725, 727
- Phase-locked loop:
 - basic concepts, 718
 - capture range, 718, 726
 - first-order loop, 720
 - integrated circuit, 721, 727
 - lock range, 717, 724
 - loop bandwidth, 719, 724
 - root locus, 719, 722
 - second-order loop, 723
- Phase margin, 629–638, 643, 648, 652, 658, 663, 676, 686, 694
- Photolithography, 84
- Photoresist, 84
- Pinch-off in MOSFETs, 43
- Pinch resistors, 118, 119
- PMOS, 39
- p**n* junction depletion region, 1
- Pole splitting, 639, 642, 654, 680

- Polysilicon, 87, 93, 129
- Popcorn noise, 742
- Power conversion efficiency, 347
- Power hyperbola, 348
- Power output:
 - Class A, 349
 - Class B, 362
- Practical Class B output stages, 368
- Predeposition, 81, 84
- Probability distribution, 244
- Protection of output stages, 377
- Punchthrough, 49, 125
- Push-pull output stage, 359
- Quasi-complementary output stages, 376, 381, 384, 388
- Reciprocity condition, 19
- Recombination, 10, 11, 79
- References, 297
- Regulated cascode, *see* Active cascode
- Regulator, 593
- Replica biasing, 822
- Resistivity, 79
- Resistor:
 - base-diffused, 116
 - base-pinch, 118
 - diffused, 146
 - emitter-diffused, 117
 - epitaxial, 119
 - MOS device, 146
 - in MOS technology, 146
 - polysilicon, 146
 - well, 144
- Return ratio, 599
 - simulation of, 621, 857
- Reverse-active region, *see* Inverse-active region of bipolar transistors
- Reverse injection, 13
- Rise time, 543
- Root locus, 626, 642, 664
 - for dominant-pole compensation, 676
 - for feedback-zero compensation, 677
 - for three-pole transfer function, 665
- Root locus rules, 667
- Saturation current, of bipolar transistor, 11, 97, 316
- Saturation region:
 - of bipolar transistor, 16, 175
 - of MOSFET, 44, 45
- Scattering-limited velocity, 59, 135
- Second-harmonic distortion, 358, 396
- Second-order phase-locked loop, 722
- Self-aligned structure, 94
- Self-biasing, 307
- Sensitivity to power-supply voltage, 304
- Series-series feedback, 562, 569
- Series-series triple, 575, 578, 583
- Series-shunt feedback, 557, 579
- Sheet resistance, 82
- Short-channel effects in MOS Transistors, 59
- Short-circuit current gain, 177
- Short-circuit protection, 377, 468
- Short-circuit time-constants, 536
- Shot noise, 736
- Shunt-series feedback, 561, 583
- Shunt-shunt feedback, 560, 563
- Signal-to-noise ratio, 798
- Silicide, 95, 147
- Silicon dioxide, 84, 87
- Silicon nitride, 87
- Single-stage feedback, 587
- Single-stage frequency response, 490
- Slew rate, 681
 - effect on sinusoidal signals, 692
 - in fully differential folded-cascode op amp, 863
 - improvement, 685, 687
- Small-signal analysis of NE5234 op amp, 469
- Small-signal model:
 - bipolar transistor, 25
 - MOS transistor, 49
- Source-coupled pair, 217
 - frequency response, 499
 - input offset voltage, 236
 - offset voltage drift, 236
 - small-signal analysis, 220, 237
 - transfer characteristics, 217
- Source degeneration, 199, 261
- Source follower, 194, 511, 592
 - distortion, 355
 - frequency response, 511
 - output stage, 353
 - small-signal properties, 194
 - super, 212, 609
 - transfer characteristics, 353

- Space-charge region, *see* Depletion region
- Spectral density of noise, 737, 748
- SPICE parameter summary, 162
- Square-law characteristic of MOSFETs, 43
- Square-law circuit, 734
- Start-up circuit, 308, 310
- Step coverage, 153
- Step response, 543, 862
- Stress migration, 156
- Substitutional impurities, 80
- Substrate contact, 152
- Substrate current in MOSFETs, 71
- Substrate *pnp* transistors, 111
- Subthreshold conduction, 135
 - also see* MOS transistor: weak inversion
- Summing node, 412, 553
- Summing-point constraints, 403, 405
- Super source follower, 212, 609
- Superbeta transistors, 124
- Supply-insensitive biasing, 303
- Sustaining voltage, 21
- Switched-capacitor amplifier, 407
 - parasitic insensitive, 411, 857
- Switched-capacitor common-mode
 - feedback, 821, 853
- Switched-capacitor filter, 412
- Switched-capacitor integrator, 412
- Symmetry:
 - mirror, 436
 - translational, 436
- T* model, 182, 185
- Tail current source, 214, 217
- Telescopic-cascode op amp, 439, 650
- Temperature coefficient:
 - of band-gap reference, 317, 319, 324
 - of base-emitter voltage, 13, 311, 315
 - of bias reference circuits, 311, 319
 - of bipolar transistor current gain, 23
 - effective, 319
 - of integrated-circuit resistors, 120
 - of threshold voltage, 48
 - of Zener diodes, 121, 122
- Temperature-insensitive biasing, 315
- Thermal noise, 747
- Thin-film resistors, 127
- Third-harmonic distortion, 358, 396
- Threshold-referenced bias circuit, 307
- Threshold voltage of MOS transistors, 40, 129, 131
- Threshold voltage temperature dependence, 47
- Time constant, 36
 - open circuit, 519, 536
 - short circuit, 536
 - zero value, 519
- Time response, 542
- Transconductance amplifier, 563
- Transfer characteristic:
 - Class B stage, 360
 - common-emitter amplifier, 176
 - emitter-coupled pair, 214
 - emitter-coupled pair with emitter degeneration, 216
 - emitter follower, 341
 - source-coupled pair, 217
 - source follower, 353
- Transit time, 28
- Transition frequency f_T , 34, 55, 64
- Transresistance amplifier, 561
- Tunneling, 8, 121
- Two-port representation, 171, 563, 584
- Unbalanced fully differential circuits, 838
- Uniform-base transistor, 10
- Unilateral amplifier, 173
- Unit devices, 254, 257, 426
- Unity-gain feedback configuration, 404, 634, 649, 660, 681
- V_{BE} -referenced bias circuit, 307, 311
- VCO, 716, 725, 734
- Velocity saturation, 59
- Virtual ground, 404
- Voltage amplifier, 559
- Voltage-controlled oscillator, 716, 728, 734
- Voltage gain, 176, 422
 - open-circuit, 177, 180
- Voltage regulator, 593
- Voltage regulator analysis, 593
- Voltage routing, 329
- V_T -referenced bias circuit, 314
- Wafer, 79
- Well, 46, 128
- Well contact, 152
- White noise, 737

- Wideband amplifier, 514, 527, 575, 618, 677, 753
- Widlar current source, 297, 300, 304
- Wilson current mirror, 272, 275, 595, 611
- Yield considerations, 157
- y-parameters, 563
- Zener diode, 8, 22, 121, 593, 595, 743
- Zener diode noise, 743
- Zener diode temperature coefficient, 120, 122
- Zener-referenced bias circuit, 595
- Zero-value time constant analysis, 519
- z-parameters, 569