

THE  
ART  
OF  
ELECTRONICS

PAUL HOROWITZ  
WINFIELD HILL

**THIRD  
EDITION**

# The Art of Electronics

## Third Edition

At long last, here is the thoroughly revised and updated, and long-anticipated, third edition of the hugely successful *The Art of Electronics*. Widely accepted as the best single authoritative text and reference on electronic circuit design, both analog and digital, the first two editions were translated into eight languages, and sold more than a million copies worldwide. The art of electronics is explained by stressing the methods actually used by circuit designers – a combination of some basic laws, rules of thumb, and a nonmathematical treatment that encourages understanding why and how a circuit works.

Paul Horowitz is a Research Professor of Physics and of Electrical Engineering at Harvard University, where in 1974 he originated the Laboratory Electronics course from which emerged *The Art of Electronics*. In addition to his work in circuit design and electronic instrumentation, his research interests have included observational astrophysics, x-ray and particle microscopy, and optical interferometry. He is one of the pioneers of the search for intelligent life beyond Earth (SETI). He is the author of some 200 scientific articles and reports, has consulted widely for industry and government, and is the designer of numerous scientific and photographic instruments.

Winfield Hill is by inclination an electronics circuit-design guru. After dropping out of the Chemical Physics graduate program at Harvard University, and obtaining an E.E. degree, he began his engineering career at Harvard's Electronics Design Center. After 7 years of learning electronics at Harvard he founded Sea Data Corporation, where he spent 16 years designing instruments for Physical Oceanography. In 1988 he was recruited by Edwin Land to join the Rowland Institute for Science. The institute subsequently merged with Harvard University in 2003. As director of the institute's Electronics Engineering Lab he has designed some 500 scientific instruments. Recent interests include high-voltage RF (to 15 kV), high-current pulsed electronics (to 1200 A), low-noise amplifiers (to sub-nV and pA), and MOSFET pulse generators.

*This page intentionally left blank*

# THE ART OF ELECTRONICS

Third Edition

**Paul Horowitz** HARVARD UNIVERSITY

**Winfield Hill** ROWLAND INSTITUTE AT HARVARD

**CAMBRIDGE**  
UNIVERSITY PRESS

32 Avenue of the Americas, New York, NY 10013-2473, USA

Cambridge University Press is part of the University of Cambridge.

It furthers the University's mission by disseminating knowledge in the pursuit of education, learning, and research at the highest international levels of excellence.

[www.cambridge.org](http://www.cambridge.org)

Information on this title: [www.cambridge.org/9780521809269](http://www.cambridge.org/9780521809269)

© Cambridge University Press, 1980, 1989, 2015

This publication is in copyright. Subject to statutory exception and to the provisions of relevant collective licensing agreements, no reproduction of any part may take place without the written permission of Cambridge University Press.

First published 1980

Second edition 1989

Third edition 2015

7th printing 2016 with corrections

Printed in the United States of America

*A catalog record for this publication is available from the British Library.*

ISBN 978-0-521-80926-9 Hardback

Cambridge University Press has no responsibility for the persistence or accuracy of URLs for external or third-party Internet websites referred to in this publication and does not guarantee that any content on such websites is, or will remain, accurate or appropriate.

*To Vida and Ava*

*This page intentionally left blank*

*In Memoriam: Jim Williams, 1948–2011*

*This page intentionally left blank*

# CONTENTS

<b>List of Tables</b>	<b>xxii</b>	1.6.5	Regulators	34
		1.6.6	Circuit applications of diodes	35
<b>Preface to the First Edition</b>	<b>xxv</b>	1.6.7	Inductive loads and diode protection	38
<b>Preface to the Second Edition</b>	<b>xxvii</b>	1.6.8	Interlude: inductors as friends	39
	1.7		Impedance and reactance	40
<b>Preface to the Third Edition</b>	<b>xxix</b>	1.7.1	Frequency analysis of reactive circuits	41
<b>ONE: Foundations</b>	<b>1</b>	1.7.2	Reactance of inductors	44
1.1	1	1.7.3	Voltages and currents as complex numbers	44
1.2	1	1.7.4	Reactance of capacitors and inductors	45
1.2.1	1	1.7.5	Ohm's law generalized	46
1.2.2	3	1.7.6	Power in reactive circuits	47
1.2.3	7	1.7.7	Voltage dividers generalized	48
1.2.4	8	1.7.8	<i>RC</i> highpass filters	48
1.2.5	9	1.7.9	<i>RC</i> lowpass filters	50
1.2.6	12	1.7.10	<i>RC</i> differentiators and integrators in the frequency domain	51
1.2.7	13			
1.3	13	1.7.11	Inductors versus capacitors	51
1.3.1	14	1.7.12	Phasor diagrams	51
1.3.2	14	1.7.13	"Poles" and decibels per octave	52
1.3.3	15	1.7.14	Resonant circuits	52
1.3.4	17	1.7.15	<i>LC</i> filters	54
1.3.5	17	1.7.16	Other capacitor applications	54
1.4	18	1.7.17	Thévenin's theorem generalized	55
1.4.1	18	1.8	Putting it all together – an AM radio	55
1.4.2	21	1.9	Other passive components	56
1.4.3	25			
1.4.4	26	1.9.1	Electromechanical devices: switches	56
1.4.5	28	1.9.2	Electromechanical devices: relays	59
1.5	28	1.9.3	Connectors	59
1.5.1	28	1.9.4	Indicators	61
1.5.2	30	1.9.5	Variable components	63
1.6	31	1.10	A parting shot: confusing markings and itty-bitty components	64
1.6.1	31			
1.6.2	31	1.10.1	Surface-mount technology: the joy and the pain	65
1.6.3	32			
1.6.4	33			

Additional Exercises for Chapter 1	66	2.6.1	Regulated power supply	123
Review of Chapter 1	68	2.6.2	Temperature controller	123
<b>TWO: Bipolar Transistors</b>	<b>71</b>	2.6.3	Simple logic with transistors and diodes	123
2.1 Introduction	71		Additional Exercises for Chapter 2	124
2.1.1 First transistor model: current amplifier	72		Review of Chapter 2	126
2.2 Some basic transistor circuits	73	<b>THREE: Field-Effect Transistors</b>	<b>131</b>	
2.2.1 Transistor switch	73	3.1	Introduction	131
2.2.2 Switching circuit examples	75	3.1.1	FET characteristics	131
2.2.3 Emitter follower	79	3.1.2	FET types	134
2.2.4 Emitter followers as voltage regulators	82	3.1.3	Universal FET characteristics	136
2.2.5 Emitter follower biasing	83	3.1.4	FET drain characteristics	137
2.2.6 Current source	85	3.1.5	Manufacturing spread of FET characteristics	138
2.2.7 Common-emitter amplifier	87	3.1.6	Basic FET circuits	140
2.2.8 Unity-gain phase splitter	88	3.2	FET linear circuits	141
2.2.9 Transconductance	89	3.2.1	Some representative JFETs: a brief tour	141
2.3 Ebers–Moll model applied to basic transistor circuits	90	3.2.2	JFET current sources	142
2.3.1 Improved transistor model: transconductance amplifier	90	3.2.3	FET amplifiers	146
2.3.2 Consequences of the Ebers–Moll model: rules of thumb for transistor design	91	3.2.4	Differential amplifiers	152
2.3.3 The emitter follower revisited	93	3.2.5	Oscillators	155
2.3.4 The common-emitter amplifier revisited	93	3.2.6	Source followers	156
2.3.5 Biasing the common-emitter amplifier	96	3.2.7	FETs as variable resistors	161
2.3.6 An aside: the perfect transistor	99	3.2.8	FET gate current	163
2.3.7 Current mirrors	101	3.3	A closer look at JFETs	165
2.3.8 Differential amplifiers	102	3.3.1	Drain current versus gate voltage	165
2.4 Some amplifier building blocks	105	3.3.2	Drain current versus drain-source voltage: output conductance	166
2.4.1 Push–pull output stages	106	3.3.3	Transconductance versus drain current	168
2.4.2 Darlington connection	109	3.3.4	Transconductance versus drain voltage	170
2.4.3 Bootstrapping	111	3.3.5	JFET capacitance	170
2.4.4 Current sharing in paralleled BJTs	112	3.3.6	Why JFET (versus MOSFET) amplifiers?	170
2.4.5 Capacitance and Miller effect	113	3.4	FET switches	171
2.4.6 Field-effect transistors	115	3.4.1	FET analog switches	171
2.5 Negative feedback	115	3.4.2	Limitations of FET switches	174
2.5.1 Introduction to feedback	116	3.4.3	Some FET analog switch examples	182
2.5.2 Gain equation	116	3.4.4	MOSFET logic switches	184
2.5.3 Effects of feedback on amplifier circuits	117	3.5	Power MOSFETs	187
2.5.4 Two important details	120	3.5.1	High impedance, thermal stability	187
2.5.5 Two examples of transistor amplifiers with feedback	121	3.5.2	Power MOSFET switching parameters	192
2.6 Some typical transistor circuits	123			

3.5.3	Power switching from logic levels	192	4.5	A detailed look at selected op-amp circuits	254
3.5.4	Power switching cautions	196	4.5.1	Active peak detector	254
3.5.5	MOSFETs versus BJTs as high-current switches	201	4.5.2	Sample-and-hold	256
3.5.6	Some power MOSFET circuit examples	202	4.5.3	Active clamp	257
3.5.7	IGBTs and other power semiconductors	207	4.5.4	Absolute-value circuit	257
3.6	MOSFETs in linear applications	208	4.5.5	A closer look at the integrator	257
3.6.1	High-voltage piezo amplifier	208	4.5.6	A circuit cure for FET leakage	259
3.6.2	Some depletion-mode circuits	209	4.5.7	Differentiators	260
3.6.3	Paralleling MOSFETs	212	4.6	Op-amp operation with a single power supply	261
3.6.4	Thermal runaway	214	4.6.1	Biassing single-supply ac amplifiers	261
	Review of Chapter 3	219	4.6.2	Capacitive loads	264
	<b>FOUR: Operational Amplifiers</b>	<b>223</b>	4.6.3	“Single-supply” op-amps	265
4.1	Introduction to op-amps – the “perfect component”	223	4.6.4	Example: voltage-controlled oscillator	267
4.1.1	Feedback and op-amps	223	4.6.5	VCO implementation: through-hole versus surface-mount	268
4.1.2	Operational amplifiers	224	4.6.6	Zero-crossing detector	269
4.1.3	The golden rules	225	4.6.7	An op-amp table	270
4.2	Basic op-amp circuits	225	4.7	Other amplifiers and op-amp types	270
4.2.1	Inverting amplifier	225	4.8	Some typical op-amp circuits	274
4.2.2	Noninverting amplifier	226	4.8.1	General-purpose lab amplifier	274
4.2.3	Follower	227	4.8.2	Stuck-node tracer	276
4.2.4	Difference amplifier	227	4.8.3	Load-current-sensing circuit	277
4.2.5	Current sources	228	4.8.4	Integrating suntan monitor	278
4.2.6	Integrators	230	4.9	Feedback amplifier frequency compensation	280
4.2.7	Basic cautions for op-amp circuits	231	4.9.1	Gain and phase shift versus frequency	281
4.3	An op-amp smorgasbord	232	4.9.2	Amplifier compensation methods	282
4.3.1	Linear circuits	232	4.9.3	Frequency response of the feedback network	284
4.3.2	Nonlinear circuits	236		Additional Exercises for Chapter 4	287
4.3.3	Op-amp application: triangle-wave oscillator	239		Review of Chapter 4	288
4.3.4	Op-amp application: pinch-off voltage tester	240			
4.3.5	Programmable pulse-width generator	241	<b>FIVE: Precision Circuits</b>		<b>292</b>
4.3.6	Active lowpass filter	241	5.1	Precision op-amp design techniques	292
4.4	A detailed look at op-amp behavior	242	5.1.1	Precision versus dynamic range	292
4.4.1	Departure from ideal op-amp performance	243	5.1.2	Error budget	293
4.4.2	Effects of op-amp limitations on circuit behavior	249	5.2	An example: the millivoltmeter, revisited	293
4.4.3	Example: sensitive millivoltmeter	253	5.2.1	The challenge: 10 mV, 1%, 10 MΩ, 1.8 V single supply	293
4.4.4	Bandwidth and the op-amp current source	254	5.2.2	The solution: precision RRIO current source	294
			5.3	The lessons: error budget, unspecified parameters	295

5.4	Another example: precision amplifier with null offset	297	5.11.3	Selecting an auto-zero op-amp	338
	5.4.1 Circuit description	297	5.11.4	Auto-zero miscellany	340
5.5	A precision-design error budget	298	5.12	Designs by the masters: Agilent's accurate DMMs	342
	5.5.1 Error budget	299	5.12.1	It's <i>impossible!</i>	342
5.6	Component errors	299	5.12.2	Wrong – it <i>is</i> possible!	342
	5.6.1 Gain-setting resistors	300	5.12.3	Block diagram: a simple plan	343
	5.6.2 The holding capacitor	300	5.12.4	The 34401A 6.5-digit front end	343
	5.6.3 Nulling switch	300	5.12.5	The 34420A 7.5-digit frontend	344
5.7	Amplifier input errors	301	5.13	Difference, differential, and instrumentation amplifiers: introduction	347
	5.7.1 Input impedance	302	5.14	Difference amplifier	348
	5.7.2 Input bias current	302	5.14.1	Basic circuit operation	348
	5.7.3 Voltage offset	304	5.14.2	Some applications	349
	5.7.4 Common-mode rejection	305	5.14.3	Performance parameters	352
	5.7.5 Power-supply rejection	306	5.14.4	Circuit variations	355
	5.7.6 Nulling amplifier: input errors	306	5.15	Instrumentation amplifier	356
5.8	Amplifier output errors	307	5.15.1	A first (but naive) guess	357
	5.8.1 Slew rate: general considerations	307	5.15.2	Classic three-op-amp instrumentation amplifier	357
	5.8.2 Bandwidth and settling time	308	5.15.3	Input-stage considerations	358
	5.8.3 Crossover distortion and output impedance	309	5.15.4	A “roll-your-own” instrumentation amplifier	359
	5.8.4 Unity-gain power buffers	311	5.15.5	A riff on robust input protection	362
	5.8.5 Gain error	312	5.16	Instrumentation amplifier miscellany	362
	5.8.6 Gain nonlinearity	312	5.16.1	Input current and noise	362
	5.8.7 Phase error and “active compensation”	314	5.16.2	Common-mode rejection	364
5.9	RRIO op-amps: the good, the bad, and the ugly	315	5.16.3	Source impedance and CMRR	365
	5.9.1 Input issues	316	5.16.4	EMI and input protection	365
	5.9.2 Output issues	316	5.16.5	Offset and CMRR trimming	366
5.10	Choosing a precision op-amp	319	5.16.6	Sensing at the load	366
	5.10.1 “Seven precision op-amps”	319	5.16.7	Input bias path	366
	5.10.2 Number per package	322	5.16.8	Output voltage range	366
	5.10.3 Supply voltage, signal range	322	5.16.9	Application example: current source	367
	5.10.4 Single-supply operation	322	5.16.10	Other configurations	368
	5.10.5 Offset voltage	323	5.16.11	Chopper and auto-zero instrumentation amplifiers	370
	5.10.6 Voltage noise	323	5.16.12	Programmable gain instrumentation amplifiers	370
	5.10.7 Bias current	325	5.16.13	Generating a differential output	372
	5.10.8 Current noise	326	5.17	Fully differential amplifiers	373
	5.10.9 CMRR and PSRR	328	5.17.1	Differential amplifiers: basic concepts	374
	5.10.10 GBW, $f_T$ , slew rate and “ $m$ ,” and settling time	328	5.17.2	Differential amplifier application example: wideband analog link	380
	5.10.11 Distortion	329	5.17.3	Differential-input ADCs	380
	5.10.12 “Two out of three isn't bad”: creating a perfect op-amp	332	5.17.4	Impedance matching	382
5.11	Auto-zeroing (chopper-stabilized) amplifiers	333			
	5.11.1 Auto-zero op-amp properties	334			
	5.11.2 When to use auto-zero op-amps	338			

5.17.5 Differential amplifier selection criteria	383	7.2.4 Timing with digital counters	465
Review of Chapter 5	388	Review of Chapter 7	470
<b>SIX: Filters</b>			
6.1 Introduction	391	391	<b>EIGHT: Low-Noise Techniques</b>
6.2 Passive filters	391	8.1 "Noise"	473
6.2.1 Frequency response with $RC$ filters	391	8.1.1 Johnson (Nyquist) noise	474
6.2.2 Ideal performance with $LC$ filters	393	8.1.2 Shot noise	475
6.2.3 Several simple examples	393	8.1.3 $1/f$ noise (flicker noise)	476
6.2.4 Enter active filters: an overview	396	8.1.4 Burst noise	477
6.2.5 Key filter performance criteria	399	8.1.5 Band-limited noise	477
6.2.6 Filter types	400	8.1.6 Interference	478
6.2.7 Filter implementation	405	8.2 Signal-to-noise ratio and noise figure	478
6.3 Active-filter circuits	406	8.2.1 Noise power density and bandwidth	479
6.3.1 VCVS circuits	407	8.2.2 Signal-to-noise ratio	479
6.3.2 VCVS filter design using our simplified table	407	8.2.3 Noise figure	479
6.3.3 State-variable filters	410	8.2.4 Noise temperature	480
6.3.4 Twin-T notch filters	414	8.3 Bipolar transistor amplifier noise	481
6.3.5 Allpass filters	415	8.3.1 Voltage noise, $e_n$	481
6.3.6 Switched-capacitor filters	415	8.3.2 Current noise $i_n$	483
6.3.7 Digital signal processing	418	8.3.3 BJT voltage noise, revisited	484
6.3.8 Filter miscellany	422	8.3.4 A simple design example: loudspeaker as microphone	486
Additional Exercises for Chapter 6	422	8.3.5 Shot noise in current sources and emitter followers	487
Review of Chapter 6	423	8.4 Finding $e_n$ from noise-figure specifications	489
<b>SEVEN: Oscillators and Timers</b>			
7.1 Oscillators	425	8.4.1 Step 1: NF versus $I_C$	489
7.1.1 Introduction to oscillators	425	8.4.2 Step 2: NF versus $R_s$	489
7.1.2 Relaxation oscillators	425	8.4.3 Step 3: getting to $e_n$	490
7.1.3 The classic oscillator-timer chip: the 555	428	8.4.4 Step 4: the spectrum of $e_n$	491
7.1.4 Other relaxation-oscillator ICs	432	8.4.5 The spectrum of $i_n$	491
7.1.5 Sinewave oscillators	435	8.4.6 When operating current is not your choice	491
7.1.6 Quartz-crystal oscillators	443	8.5 Low-noise design with bipolar transistors	492
7.1.7 Higher stability: TCXO, OCXO, and beyond	450	8.5.1 Noise-figure example	492
7.1.8 Frequency synthesis: DDS and PLL	451	8.5.2 Charting amplifier noise with $e_n$ and $i_n$	493
7.1.9 Quadrature oscillators	453	8.5.3 Noise resistance	494
7.1.10 Oscillator "jitter"	457	8.5.4 Charting comparative noise	495
7.2 Timers	457	8.5.5 Low-noise design with BJTs: two examples	495
7.2.1 Step-triggered pulses	458	8.5.6 Minimizing noise: BJTs, FETs, and transformers	496
7.2.2 Monostable multivibrators	461	8.5.7 A design example: 40¢ "lightning detector" preamp	497
7.2.3 A monostable application: limiting pulse width and duty cycle	465	8.5.8 Selecting a low-noise bipolar transistor	500
		8.5.9 An extreme low-noise design challenge	505

8.6	Low-noise design with JFETs	509	8.11.13 Test fixture for compensation and calibration	554
8.6.1	Voltage noise of JFETs	509	8.11.14 A final remark	555
8.6.2	Current noise of JFETs	511	8.12 Noise measurements and noise sources	555
8.6.3	Design example: low-noise wideband JFET “hybrid” amplifiers	512	8.12.1 Measurement without a noise source	555
8.6.4	Designs by the masters: SR560 low-noise preamplifier	512	8.12.2 An example: transistor-noise test circuit	556
8.6.5	Selecting low-noise JFETs	515	8.12.3 Measurement with a noise source	556
8.7	Charting the bipolar–FET shootout	517	8.12.4 Noise and signal sources	558
8.7.1	What about MOSFETs?	519	8.13 Bandwidth limiting and rms voltage measurement	561
8.8	Noise in differential and feedback amplifiers	520	8.13.1 Limiting the bandwidth	561
8.9	Noise in operational amplifier circuits	521	8.13.2 Calculating the integrated noise	563
8.9.1	Guide to Table 8.3: choosing low-noise op-amps	525	8.13.3 Op-amp “low-frequency noise” with asymmetric filter	564
8.9.2	Power-supply rejection ratio	533	8.13.4 Finding the $1/f$ corner frequency	566
8.9.3	Wrapup: choosing a low-noise op-amp	533	8.13.5 Measuring the noise voltage	567
8.9.4	Low-noise instrumentation amplifiers and video amplifiers	533	8.13.6 Measuring the noise current	569
8.9.5	Low-noise hybrid op-amps	534	8.13.7 Another way: roll-your-own $fA/\sqrt{Hz}$ instrument	571
8.10	Signal transformers	535	8.13.8 Noise potpourri	574
8.10.1	A low-noise wideband amplifier with transformer feedback	536	8.14 Signal-to-noise improvement by bandwidth narrowing	574
8.11	Noise in transimpedance amplifiers	537	8.14.1 Lock-in detection	575
8.11.1	Summary of the stability problem	537	8.15 Power-supply noise	578
8.11.2	Amplifier input noise	538	8.15.1 Capacitance multiplier	578
8.11.3	The $e_nC$ noise problem	538	8.16 Interference, shielding, and grounding	579
8.11.4	Noise in the transresistance amplifier	539	8.16.1 Interfering signals	579
8.11.5	An example: wideband JFET photodiode amplifier	540	8.16.2 Signal grounds	582
8.11.6	Noise versus gain in the transimpedance amplifier	540	8.16.3 Grounding between instruments	583
8.11.7	Output bandwidth limiting in the transimpedance amplifier	542	Additional Exercises for Chapter 8	588
8.11.8	Composite transimpedance amplifiers	543	Review of Chapter 8	590
8.11.9	Reducing input capacitance: bootstrapping the transimpedance amplifier	547	<b>NINE: Voltage Regulation and Power Conversion</b>	<b>594</b>
8.11.10	Isolating input capacitance: cascoding the transimpedance amplifier	547	9.1 Tutorial: from zener to series-pass linear regulator	595
8.11.11	Transimpedance amplifiers with capacitive feedback	548	9.1.1 Adding feedback	596
8.11.12	Scanning tunneling microscope preamplifier	552	9.2 Basic linear regulator circuits with the classic 723	598
		553	9.2.1 The 723 regulator	598
			9.2.2 In defense of the beleaguered 723	600
			9.3 Fully integrated linear regulators	600
			9.3.1 Taxonomy of linear regulator ICs	601
			9.3.2 Three-terminal fixed regulators	601

9.3.3	Three-terminal adjustable regulators	602	9.7.1	The ac-to-dc input stage	660
9.3.4	317-style regulator: application hints	604	9.7.2	The dc-to-dc converter	662
9.3.5	317-style regulator: circuit examples	608	9.8	A real-world switcher example	665
9.3.6	Lower-dropout regulators	610	9.8.1	Switchers: top-level view	665
9.3.7	True low-dropout regulators	611	9.8.2	Switchers: basic operation	665
9.3.8	Current-reference 3-terminal regulator	611	9.8.3	Switchers: looking more closely	668
9.3.9	Dropout voltages compared	612	9.8.4	The “reference design”	671
9.3.10	Dual-voltage regulator circuit example	613	9.8.5	Wrapup: general comments on line-powered switching power supplies	672
9.3.11	Linear regulator choices	613	9.8.6	When to use switchers	672
9.3.12	Linear regulator idiosyncrasies	613	9.9	Inverters and switching amplifiers	673
9.3.13	Noise and ripple filtering	619	9.10	Voltage references	674
9.3.14	Current sources	620	9.10.1	Zener diode	674
9.4	Heat and power design	623	9.10.2	Bandgap ( $V_{BE}$ ) reference	679
9.4.1	Power transistors and heatsinking	624	9.10.3	JFET pinch-off ( $V_P$ ) reference	680
9.4.2	Safe operating area	627	9.10.4	Floating-gate reference	681
9.5	From ac line to unregulated supply	628	9.10.5	Three-terminal precision references	681
9.5.1	ac-line components	629	9.10.6	Voltage reference noise	682
9.5.2	Transformer	632	9.10.7	Voltage references: additional Comments	683
9.5.3	dc components	633	9.11	Commercial power-supply modules	684
9.5.4	Unregulated split supply – on the bench!	634	9.12	Energy storage: batteries and capacitors	686
9.5.5	Linear versus switcher: ripple and noise	635	9.12.1	Battery characteristics	687
9.6	Switching regulators and dc–dc converters	636	9.12.2	Choosing a battery	688
9.6.1	Linear versus switching	636	9.12.3	Energy storage in capacitors	688
9.6.2	Switching converter topologies	638	9.13	Additional topics in power regulation	690
9.6.3	Inductorless switching converters	638	9.13.1	Overshoot crowbars	690
9.6.4	Converters with inductors: the basic non-isolated topologies	641	9.13.2	Extending input-voltage range	693
9.6.5	Step-down (buck) converter	642	9.13.3	Foldback current limiting	693
9.6.6	Step-up (boost) converter	647	9.13.4	Outboard pass transistor	695
9.6.7	Inverting converter	648	9.13.5	High-voltage regulators	695
9.6.8	Comments on the non-isolated converters	649		Review of Chapter 9	699
9.6.9	Voltage mode and current mode	651		<b>TEN: Digital Logic</b>	<b>703</b>
9.6.10	Converters with transformers: the basic designs	653	10.1	Basic logic concepts	703
9.6.11	The flyback converter	655	10.1.1	Digital versus analog	703
9.6.12	Forward converters	656	10.1.2	Logic states	704
9.6.13	Bridge converters	659	10.1.3	Number codes	705
9.7	Ac-line-powered (“offline”) switching converters	660	10.1.4	Gates and truth tables	708
			10.1.5	Discrete circuits for gates	711
			10.1.6	Gate-logic example	712
			10.1.7	Assertion-level logic notation	713
			10.2	Digital integrated circuits: CMOS and Bipolar (TTL)	714
			10.2.1	Catalog of common gates	715
			10.2.2	IC gate circuits	717
			10.2.3	CMOS and bipolar (“TTL”) characteristics	718

10.2.4 Three-state and open-collector devices	720	11.3 An example: pseudorandom byte generator	770
10.3 Combinational logic	722	11.3.1 How to make pseudorandom bytes	771
10.3.1 Logic identities	722	11.3.2 Implementation in standard logic	772
10.3.2 Minimization and Karnaugh maps	723	11.3.3 Implementation with programmable logic	772
10.3.3 Combinational functions available as ICs	724	11.3.4 Programmable logic – HDL entry	775
10.4 Sequential logic	728	11.3.5 Implementation with a microcontroller	777
10.4.1 Devices with memory: flip-flops	728	11.4 Advice	782
10.4.2 Clocked flip-flops	730	11.4.1 By Technologies	782
10.4.3 Combining memory and gates: sequential logic	734	11.4.2 By User Communities	785
10.4.4 Synchronizer	737	Review of Chapter 11	787
10.4.5 Monostable multivibrator	739		
10.4.6 Single-pulse generation with flip-flops and counters	739		
10.5 Sequential functions available as integrated circuits	740	<b>TWELVE: Logic Interfacing</b>	<b>790</b>
10.5.1 Latches and registers	740	12.1 CMOS and TTL logic interfacing	790
10.5.2 Counters	741	12.1.1 Logic family chronology – a brief history	790
10.5.3 Shift registers	744	12.1.2 Input and output characteristics	794
10.5.4 Programmable logic devices	745	12.1.3 Interfacing between logic families	798
10.5.5 Miscellaneous sequential functions	746	12.1.4 Driving digital logic inputs	802
10.6 Some typical digital circuits	748	12.1.5 Input protection	804
10.6.1 Modulo- $n$ counter: a timing example	748	12.1.6 Some comments about logic inputs	805
10.6.2 Multiplexed LED digital display	751	12.1.7 Driving digital logic from comparators or op-amps	806
10.6.3 An $n$ -pulse generator	752	12.2 An aside: probing digital signals	808
10.7 Micropower digital design	753	12.3 Comparators	809
10.7.1 Keeping CMOS low power	754	12.3.1 Outputs	810
10.8 Logic pathology	755	12.3.2 Inputs	812
10.8.1 dc problems	755	12.3.3 Other parameters	815
10.8.2 Switching problems	756	12.3.4 Other cautions	816
10.8.3 Congenital weaknesses of TTL and CMOS	758	12.4 Driving external digital loads from logic levels	817
Additional Exercises for Chapter 10	760	12.4.1 Positive loads: direct drive	817
Review of Chapter 10	762	12.4.2 Positive loads: transistor assisted	820
<b>ELEVEN: Programmable Logic Devices</b>	<b>764</b>	12.4.3 Negative or ac loads	821
11.1 A brief history	764	12.4.4 Protecting power switches	823
11.2 The hardware	765	12.4.5 nMOS LSI interfacing	826
11.2.1 The basic PAL	765	12.5 Optoelectronics: emitters	829
11.2.2 The PLA	768	12.5.1 Indicators and LEDs	829
11.2.3 The FPGA	768	12.5.2 Laser diodes	834
11.2.4 The configuration memory	769	12.5.3 Displays	836
11.2.5 Other programmable logic devices	769	12.6 Optoelectronics: detectors	840
11.2.6 The software	769		

12.6.1	Photodiodes and phototransistors	841	13.2.8	PWM as digital-to-analog converter	888
12.6.2	Photomultipliers	842	13.2.9	Frequency-to-voltage converters	890
12.7	Optocouplers and relays	843	13.2.10	Rate multiplier	890
12.7.1	I: Phototransistor output optocouplers	844	13.2.11	Choosing a DAC	891
12.7.2	II: Logic-output optocouplers	844	13.3	Some DAC application examples	891
12.7.3	III: Gate driver optocouplers	846	13.3.1	General-purpose laboratory source	891
12.7.4	IV: Analog-oriented optocouplers	847	13.3.2	Eight-channel source	893
12.7.5	V: Solid-state relays (transistor output)	848	13.3.3	Nanoamp wide-compliance bipolarity current source	894
12.7.6	VI: Solid-state relays (triac/SCR output)	849	13.3.4	Precision coil driver	897
12.7.7	VII: ac-input optocouplers	851	13.4	Converter linearity – a closer look	899
12.7.8	Interrupters	851	13.5	Analog-to-digital converters	900
12.8	Optoelectronics: fiber-optic digital links	852	13.5.1	Digitizing: aliasing, sampling rate, and sampling depth	900
12.8.1	TOSLINK	852	13.5.2	ADC Technologies	902
12.8.2	Versatile Link	854	13.6	ADCs I: Parallel (“flash”) encoder	903
12.8.3	ST/SC glass-fiber modules	855	13.6.1	Modified flash encoders	903
12.8.4	Fully integrated high-speed fiber-transceiver modules	855	13.6.2	Driving flash, folding, and RF ADCs	904
12.9	Digital signals and long wires	856	13.6.3	Undersampling flash-converter example	907
12.9.1	On-board interconnections	856	13.7	ADCs II: Successive approximation	908
12.9.2	Intercard connections	858	13.7.1	A simple SAR example	909
12.10	Driving Cables	858	13.7.2	Variations on successive approximation	909
12.10.1	Coaxial cable	858	13.7.3	An A/D conversion example	910
12.10.2	The right way – I: Far-end termination	860	13.8	ADCs III: integrating	912
12.10.3	Differential-pair cable	864	13.8.1	Voltage-to-frequency conversion	912
12.10.4	RS-232	871	13.8.2	Single-slope integration	914
	12.10.5 Wrapup	874	13.8.3	Integrating converters	914
	Review of Chapter 12	875	13.8.4	Dual-slope integration	914
<b>THIRTEEN : Digital meets Analog</b>		<b>879</b>	13.8.5	Analog switches in conversion applications (a detour)	916
13.1	Some preliminaries	879	13.8.6	Designs by the masters: Agilent’s world-class “multislope” converters	918
13.1.1	The basic performance parameters	879	13.9	ADCs IV: delta–sigma	922
13.1.2	Codes	880	13.9.1	A simple delta–sigma for our suntan monitor	922
13.1.3	Converter errors	880	13.9.2	Demystifying the delta–sigma converter	923
13.1.4	Stand-alone versus integrated	880	13.9.3	$\Delta\Sigma$ ADC and DAC	923
13.2	Digital-to-analog converters	881	13.9.4	The $\Delta\Sigma$ process	924
13.2.1	Resistor-string DACs	881	13.9.5	An aside: “noise shaping”	927
13.2.2	R–2R ladder DACs	882	13.9.6	The bottom line	928
13.2.3	Current-steering DACs	883	13.9.7	A simulation	928
13.2.4	Multiplying DACs	884	13.9.8	What about DACs?	930
13.2.5	Generating a voltage output	885			
13.2.6	Six DACs	886			
13.2.7	Delta–sigma DACs	888			

13.9.9 Pros and Cons of $\Delta\Sigma$ oversampling converters	931	13.14.8 A “hybrid digital filter”	983
13.9.10 Idle tones	932	Additional Exercises for Chapter 13	984
13.9.11 Some delta-sigma application examples	932	Review of Chapter 13	985
13.10 ADCs: choices and tradeoffs	938	<b>FOURTEEN: Computers, Controllers, and Data Links</b>	<b>989</b>
13.10.1 Delta-sigma and the competition	938	14.1 Computer architecture: CPU and data bus	990
13.10.2 Sampling versus averaging ADCs: noise	940	14.1.1 CPU	990
13.10.3 Micropower A/D converters	941	14.1.2 Memory	991
13.11 Some unusual A/D and D/A converters	942	14.1.3 Mass memory	991
13.11.1 ADE7753 multifunction ac power metering IC	943	14.1.4 Graphics, network, parallel, and serial ports	992
13.11.2 AD7873 touchscreen digitizer	944	14.1.5 Real-time I/O	992
13.11.3 AD7927 ADC with sequencer	945	14.1.6 Data bus	992
13.11.4 AD7730 precision bridge-measurement subsystem	945	14.2 A computer instruction set	993
13.12 Some A/D conversion system examples	946	14.2.1 Assembly language and machine language	993
13.12.1 Multiplexed 16-channel data-acquisition system	946	14.2.2 Simplified “x86” instruction set	993
13.12.2 Parallel multichannel successive-approximation data-acquisition system	950	14.2.3 A programming example	996
13.12.3 Parallel multichannel delta-sigma data-acquisition system	952	14.3 Bus signals and interfacing	997
13.13 Phase-locked loops	955	14.3.1 Fundamental bus signals: data, address, strobe	997
13.13.1 Introduction to phase-locked loops	955	14.3.2 Programmed I/O: data out	998
13.13.2 PLL components	957	14.3.3 Programming the XY vector display	1000
13.13.3 PLL design	960	14.3.4 Programmed I/O: data in	1001
13.13.4 Design example: frequency multiplier	961	14.3.5 Programmed I/O: status registers	1002
13.13.5 PLL capture and lock	964	14.3.6 Programmed I/O: command registers	1004
13.13.6 Some PLL applications	966	14.3.7 Interrupts	1005
13.13.7 Wrapup: noise and jitter rejection in PLLs	974	14.3.8 Interrupt handling	1006
13.14 Pseudorandom bit sequences and noise generation	974	14.3.9 Interrupts in general	1008
13.14.1 Digital-noise generation	974	14.3.10 Direct memory access	1010
13.14.2 Feedback shift register sequences	975	14.3.11 Summary of PC104/ISA 8-bit bus signals	1012
13.14.3 Analog noise generation from maximal-length sequences	977	14.3.12 The PC104 as an embedded single-board computer	1013
13.14.4 Power spectrum of shift-register sequences	977	14.4 Memory types	1014
13.14.5 Low-pass filtering	979	14.4.1 Volatile and non-volatile memory	1014
13.14.6 Wrapup	981	14.4.2 Static versus dynamic RAM	1015
13.14.7 “True” random noise generators	982	14.4.3 Static RAM	1016
		14.4.4 Dynamic RAM	1018
		14.4.5 Nonvolatile memory	1021
		14.4.6 Memory wrapup	1026
		14.5 Other buses and data links: overview	1027
		14.6 Parallel buses and data links	1028
		14.6.1 Parallel chip “bus” interface – an example	1028

14.6.2	Parallel chip data links – two high-speed examples	1030	15.7	Design example 5: stabilized mechanical platform	1077
14.6.3	Other parallel computer buses	1030	15.8	Peripheral ICs for microcontrollers	1078
14.6.4	Parallel peripheral buses and data links	1031	15.8.1	Peripherals with direct connection	1079
14.7	Serial buses and data links	1032	15.8.2	Peripherals with SPI connection	1082
14.7.1	SPI	1032	15.8.3	Peripherals with I <sup>2</sup> C connection	1084
14.7.2	I <sup>2</sup> C 2-wire interface (“TWI”)	1034	15.8.4	Some important hardware constraints	1086
14.7.3	Dallas–Maxim “1-wire” serial interface	1035	15.9	Development environment	1086
14.7.4	JTAG	1036	15.9.1	Software	1086
14.7.5	Clock-be-gone: clock recovery	1037	15.9.2	Real-time programming constraints	1088
14.7.6	SATA, eSATA, and SAS	1037	15.9.3	Hardware	1089
14.7.7	PCI Express	1037	15.9.4	The Arduino Project	1092
14.7.8	Asynchronous serial (RS-232, RS-485)	1038	15.10	Wrapup	1092
14.7.9	Manchester coding	1039	15.10.1	How expensive are the tools?	1092
14.7.10	Biphase coding	1041	15.10.2	When to use microcontrollers	1093
14.7.11	RLL binary: bit stuffing	1041	15.10.3	How to select a microcontroller	1094
14.7.12	RLL coding: 8b/10b and others	1041	15.10.4	A parting shot	1094
14.7.13	USB	1042		Review of Chapter 15	1095
14.7.14	FireWire	1042			
14.7.15	Controller Area Network (CAN)	1043	<b>APPENDIX A: Math Review</b>		<b>1097</b>
14.7.16	Ethernet	1045	A.1	Trigonometry, exponentials, and logarithms	1097
14.8	Number formats	1046	A.2	Complex numbers	1097
14.8.1	Integers	1046	A.3	Differentiation (Calculus)	1099
14.8.2	Floating-point numbers	1047	A.3.1	Derivatives of some common functions	1099
	Review of Chapter 14	1049	A.3.2	Some rules for combining derivatives	1100
			A.3.3	Some examples of differentiation	1100
<b>FIFTEEN: Microcontrollers</b>		<b>1053</b>	<b>APPENDIX B: How to Draw Schematic Diagrams</b>		<b>1101</b>
15.1	Introduction	1053	B.1	General principles	1101
15.2	Design example 1: suntan monitor (V)	1054	B.2	Rules	1101
15.2.1	Implementation with a microcontroller	1054	B.3	Hints	1103
15.2.2	Microcontroller code (“firmware”)	1056	B.4	A humble example	1103
15.3	Overview of popular microcontroller families	1059	<b>APPENDIX C: Resistor Types</b>		<b>1104</b>
15.3.1	On-chip peripherals	1061	C.1	Some history	1104
15.4	Design example 2: ac power control	1062	C.2	Available resistance values	1104
15.4.1	Microcontroller implementation	1062	C.3	Resistance marking	1105
15.4.2	Microcontroller code	1064	C.4	Resistor types	1105
15.5	Design example 3: frequency synthesizer	1065	C.5	Confusion derby	1105
15.5.1	Microcontroller code	1067	<b>APPENDIX D: Thévenin’s Theorem</b>		<b>1107</b>
15.6	Design example 4: thermal controller	1069	D.1	The proof	1107
15.6.1	The hardware	1070			
15.6.2	The control loop	1074			
15.6.3	Microcontroller code	1075			

D.1.1	Two examples – voltage dividers	1107	I.3	Recording analog-format broadcast or cable television	1135
D.2	Norton's theorem	1108	I.4	Digital television: what is it?	1136
D.3	Another example	1108	I.5	Digital television: broadcast and cable delivery	1138
D.4	Millman's theorem	1108	I.6	Direct satellite television	1139
<b>APPENDIX E: LC Butterworth Filters</b>		<b>1109</b>	I.7	Digital video streaming over internet	1140
E.1	Lowpass filter	1109	I.8	Digital cable: premium services and conditional access	1141
E.2	Highpass filter	1109	I.8.1	Digital cable: video-on-demand	1141
E.3	Filter examples	1109	I.8.2	Digital cable: switched broadcast	1142
<b>APPENDIX F: Load Lines</b>		<b>1112</b>	I.9	Recording digital television	1142
F.1	An example	1112	I.10	Display technology	1142
F.2	Three-terminal devices	1112	I.11	Video connections: analog and digital	1143
<b>APPENDIX G: The Curve Tracer</b>		<b>1115</b>	<b>APPENDIX J: SPICE Primer</b>		<b>1146</b>
<b>APPENDIX H: Transmission Lines and Impedance Matching</b>		<b>1116</b>	J.1	Setting up ICAP SPICE	1146
H.1	Some properties of transmission lines	1116	J.2	Entering a Diagram	1146
H.1.1	Characteristic impedance	1116	J.3	Running a simulation	1146
H.1.2	Termination: pulses	1117	J.3.1	Schematic entry	1146
H.1.3	Termination: sinusoidal signals	1120	J.3.2	Simulation: frequency sweep	1147
H.1.4	Loss in transmission lines	1121	J.3.3	Simulation: input and output waveforms	1147
H.2	Impedance matching	1122	J.4	Some final points	1148
H.2.1	Resistive (lossy) broadband matching network	1123	J.5	A detailed example: exploring amplifier distortion	1148
H.2.2	Resistive attenuator	1123	J.6	Expanding the parts database	1149
H.2.3	Transformer (lossless) broadband matching network	1124	<b>APPENDIX K: "Where Do I Go to Buy Electronic Goodies?"</b>		<b>1150</b>
H.2.4	Reactive (lossless) narrowband matching networks	1125	<b>APPENDIX L: Workbench Instruments and Tools</b>		<b>1152</b>
H.3	Lumped-element delay lines and pulse-forming networks	1126	<b>APPENDIX M: Catalogs, Magazines, Data-books</b>		<b>1153</b>
H.4	Epilogue: ladder derivation of characteristic impedance	1127	<b>APPENDIX N: Further Reading and References</b>		<b>1154</b>
H.4.1	First method: terminated line	1127	<b>APPENDIX O: The Oscilloscope</b>		<b>1158</b>
H.4.2	Second method: semi-infinite line	1127	O.1	The analog oscilloscope	1158
H.4.3	Postscript: lumped-element delay lines	1128	O.1.1	Vertical	1158
<b>APPENDIX I: Television: A Compact Tutorial</b>		<b>1131</b>	O.1.2	Horizontal	1158
I.1	Television: video plus audio	1131	O.1.3	Triggering	1159
I.1.1	The audio	1131	O.1.4	Hints for beginners	1160
I.1.2	The video	1132	O.1.5	Probes	1160
I.2	Combining and sending the audio + video: modulation	1133	O.1.6	Grounds	1161
			O.1.7	Other analog scope features	1161
			O.2	The digital oscilloscope	1162

O.2.1 What's different?	1162	<b>APPENDIX P: Acronyms and Abbreviations</b>	<b>1166</b>
O.2.2 Some cautions	1164	<b>Index</b>	<b>1171</b>

# LIST OF TABLES

1.1. Representative Diodes.	32	8.4. Noise Integrals.	564
2.1. Representative Bipolar Transistors.	74	8.5. Auto-zero Noise Measurements.	569
2.2. Bipolar Power Transistors.	106	9.1. 7800-style Fixed Regulators.	602
3.1. JFET Mini-table.	141	9.2. Three-terminal Adjustable Voltage Regulators	
3.2. Selected Fast JFET-input Op-amps.	155	(LM317-style).	605
3.3. Analog Switches.	176	9.3. Low-dropout Linear Voltage Regulators.	614
3.4a. MOSFETs – Small <i>n</i> -channel (to 250 V), and <i>p</i> -channel (to 100 V).	188	9.4. Selected Charge-pump Converters.	640
3.4b. <i>n</i> -channel Power MOSFETs, 55 V to 4500 V.	189	9.5a. Voltage-mode Integrated Switching	
3.5. MOSFET Switch Candidates.	206	Regulators.	653
3.6. Depletion-mode <i>n</i> -channel MOSFETs.	210	9.5b. Selected Current-mode Integrated Switching	
3.7. Junction Field-Effect Transistors (JFETs).	217	Regulators.	654
3.8. Low-side MOSFET Gate Drivers.	218	9.6. External-switch Controllers.	658
4.1. Op-amp Parameters.	245	9.7. Shunt (2-terminal) Voltage References.	677
4.2a. Representative Operational Amplifiers.	271	9.8. Series (3-terminal) Voltage References.	678
4.2b. Monolithic Power and High-voltage Op-amps.	272	9.9. Battery Choices.	689
5.1. Millivoltmeter Candidate Op-amps.	296	9.10. Energy Storage: Capacitor Versus Battery.	690
5.2. Representative Precision Op-amps.	302	10.1. Selected Logic Families.	706
5.3. Nine Low-input-current Op-amps.	303	10.2. 4-bit Signed Integers in Three Systems of	
5.4. Representative High-speed Op-amps.	310	Representation.	707
5.5. “Seven” Precision Op-amps: High Voltage.	320	10.3. Standard Logic Gates.	716
5.6. Chopper and Auto-zero Op-amps.	335	10.4. Logic Identities.	722
5.7. Selected Difference Amplifiers.	353	10.5. Selected Counter ICs.	742
5.8. Selected Instrumentation Amplifiers	363	10.6. Selected Reset/Supervisors.	756
5.9. Selected Programmable-gain Instrumentation Amplifiers.	370	12.1. Representative Comparators.	812
5.10. Selected Differential Amplifiers.	375	12.2. Comparators.	813
6.1. Time-domain Performance Comparison for Lowpass Filters.	383	12.3. Power Logic Registers.	819
6.2. VCVS Lowpass Filters.	406	12.4. A Few Protected MOSFETs.	825
7.1. 555-type Oscillators.	408	12.5. Selected High-side Switches.	826
7.2. Oscillator Types.	430	12.6. Selected Panel-mount LEDs.	832
7.3. Monostable Multivibrators.	452	13.1. Six Digital-to-analog Converters.	889
7.4. “Type 123” Monostable Timing.	462	13.2. Selected Digital-to-analog Converters.	893
8.1a. Low-noise Bipolar Transistors (BJTs).	480	13.3. Multiplying DACs.	894
8.1b. Dual Low-noise BJTs.	501	13.4. Selected Fast ADCs.	905
8.2. Low-noise Junction FETs (JFETs).	502	13.5. Successive-approximation ADCs.	910
8.3a. Low-noise BJT-input Op-amps.	516	13.6. Selected Micropower ADCs.	916
8.3b. Low-noise FET-input Op-amps.	522	13.7. 4053-style SPDT Switches.	917
8.3c. High-speed Low-noise Op-amps.	523	13.8. Agilent’s Multislope III ADCs.	921
	524	13.9. Selected Delta–sigma ADCs.	935
		13.10. Audio Delta–sigma ADCs.	937
		13.11. Audio ADCs.	939
		13.12. Speciality ADCs.	942

---

13.13. Phase-locked Loop ICs.	972	14.3. Common Buses and Data Links.	1029
13.14. Single-tap LFSRs.	976	14.4. RS-232 Signals.	1039
13.15. LFSRs with Length a Multiple of 8.	976	14.5. ASCII Codes.	1040
14.1. Simplified x86 Instruction Set.	994	C.1. Selected Resistor Types.	1106
14.2. PC104/ISA Bus Signals.	1013	E.1. Butterworth Lowpass Filters.	1110
		H.1. Pi and T Attenuators.	1124

*This page intentionally left blank*

# PREFACE TO THE FIRST EDITION

This volume is intended as an electronic circuit design textbook and reference book; it begins at a level suitable for those with no previous exposure to electronics and carries the reader through to a reasonable degree of proficiency in electronic circuit design. We have used a straightforward approach to the essential ideas of circuit design, coupled with an in-depth selection of topics. We have attempted to combine the pragmatic approach of the practicing physicist with the quantitative approach of the engineer, who wants a thoroughly evaluated circuit design.

This book evolved from a set of notes written to accompany a one-semester course in laboratory electronics at Harvard. That course has a varied enrollment – undergraduates picking up skills for their eventual work in science or industry, graduate students with a field of research clearly in mind, and advanced graduate students and post-doctoral researchers who suddenly find themselves hampered by their inability to “do electronics.”

It soon became clear that existing textbooks were inadequate for such a course. Although there are excellent treatments of each electronics specialty, written for the planned sequence of a four-year engineering curriculum or for the practicing engineer, those books that attempt to address the whole field of electronics seem to suffer from excessive detail (the handbook syndrome), from oversimplification (the cookbook syndrome), or from poor balance of material. Much of the favorite pedagogy of beginning textbooks is quite unnecessary and, in fact, is not used by practicing engineers, while useful circuitry and methods of analysis in daily use by circuit designers lie hidden in application notes, engineering journals, and hard-to-get data books. In other words, there is a tendency among textbook writers to represent the theory, rather than the art, of electronics.

We collaborated in writing this book with the specific intention of combining the discipline of a circuit design engineer with the perspective of a practicing experimental physicist and teacher of electronics. Thus, the treatment in this book reflects our philosophy that electronics, as currently practiced, is basically a simple art, a combination of some basic laws, rules of thumb, and a large bag of tricks. For these reasons we have omitted entirely the

usual discussions of solid-state physics, the *h*-parameter model of transistors, and complicated network theory, and reduced to a bare minimum the mention of load lines and the *s*-plane. The treatment is largely nonmathematical, with strong encouragement of circuit brainstorming and mental (or, at most, back-of-the-envelope) calculation of circuit values and performance.

In addition to the subjects usually treated in electronics books, we have included the following:

- an easy-to-use transistor model;
- extensive discussion of useful subcircuits, such as current sources and current mirrors;
- single-supply op-amp design;
- easy-to-understand discussions of topics on which practical design information is often difficult to find: op-amp frequency compensation, low-noise circuits, phase-locked loops, and precision linear design;
- simplified design of active filters, with tables and graphs;
- a section on noise, shielding, and grounding;
- a unique graphical method for streamlined low-noise amplifier analysis;
- a chapter on voltage references and regulators, including constant current supplies;
- a discussion of monostable multivibrators and their idiosyncrasies;
- a collection of digital logic pathology, and what to do about it;
- an extensive discussion of interfacing to logic, with emphasis on the new NMOS and PMOS LSI;
- a detailed discussion of A/D and D/A conversion techniques;
- a section on digital noise generation;
- a discussion of minicomputers and interfacing to data buses, with an introduction to assembly language;
- a chapter on microprocessors, with actual design examples and discussion – how to design them into instruments, and how to make them do what you want;
- a chapter on construction techniques: prototyping, printed circuit boards, instrument design;

- a simplified way to evaluate high-speed switching circuits;
- a chapter on scientific measurement and data processing: what you can measure and how accurately, and what to do with the data;
- bandwidth narrowing methods made clear: signal averaging, multichannel scaling, lock-in amplifiers, and pulse-height analysis;
- amusing collections of “bad circuits,” and collections of “circuit ideas”;
- useful appendixes on how to draw schematic diagrams, IC generic types, *LC* filter design, resistor values, oscilloscopes, mathematics review, and others;
- tables of diodes, transistors, FETs, op-amps, comparators, regulators, voltage references, microprocessors, and other devices, generally listing the characteristics of both the most popular and the best types.

Throughout we have adopted a philosophy of naming names, often comparing the characteristics of competing devices for use in any circuit, and the advantages of alternative circuit configurations. Example circuits are drawn with real device types, not black boxes. The overall intent is to bring the reader to the point of understanding clearly the choices one makes in designing a circuit – how to choose circuit configurations, device types, and parts values. The use of largely nonmathematical circuit design techniques does not result in circuits that cut corners or compromise performance or reliability. On the contrary, such techniques enhance one’s understanding of the real choices and compromises faced in engineering a circuit and represent the best approach to good circuit design.

This book can be used for a full-year electronic circuit design course at the college level, with only a minimum mathematical prerequisite; namely, some acquaintance with trigonometric and exponential functions, and preferably a bit of differential calculus. (A short review of complex numbers and derivatives is included as an appendix.) If the less essential sections are omitted, it can serve as the text for a one-semester course (as it does at Harvard).

A separately available laboratory manual, *Laboratory Manual for the Art of Electronics* (Horowitz and Robinson, 1981), contains twenty-three lab exercises, together with reading and problem assignments keyed to the text.

To assist the reader in navigation, we have designated with open boxes in the margin those sections within each chapter that we feel can be safely passed over in an abbreviated reading. For a one-semester course it would probably be wise to omit, in addition, the materials of Chapter 5 (first half), 7, 12, 13, 14, and possibly 15, as explained in the introductory paragraphs of those chapters.

We would like to thank our colleagues for their thoughtful comments and assistance in the preparation of the manuscript, particularly Mike Aronson, Howard Berg, Dennis Crouse, Carol Davis, David Griesinger, John Hagen, Tom Hayes, Peter Horowitz, Bob Kline, Costas Papaliolios, Jay Sage, and Bill Vetterling. We are indebted to Eric Hieber and Jim Mobley, and to Rhona Johnson and Ken Werner of Cambridge University Press, for their imaginative and highly professional work.

Paul Horowitz  
Winfield Hill  
April 1980

# PREFACE TO THE SECOND EDITION

Electronics, perhaps more than any other field of technology, has enjoyed an explosive development in the last four decades. Thus it was with some trepidation that we attempted, in 1980, to bring out a definitive volume teaching the art of the subject. By “art” we meant the kind of mastery that comes from an intimate familiarity with real circuits, actual devices, and the like, rather than the more abstract approach often favored in textbooks on electronics. Of course, in a rapidly evolving field, such a nuts-and-bolts approach has its hazards – most notably a frighteningly quick obsolescence.

The pace of electronics technology did not disappoint us! Hardly was the ink dry on the first edition before we felt foolish reading our words about “the classic [2Kbyte] 2716 EPROM... with a price tag of about \$25.” They’re so classic you can’t even get them anymore, having been replaced by EEPROMs 64 times as large, and costing less than half the price! Thus a major element of this revision responds to improved devices and methods – completely rewritten chapters on microcomputers and microprocessors (using the IBM PC and the 68008) and substantially revised chapters on digital electronics (including PLDs, and the new HC and AC logic families), on op-amps and precision design (reflecting the availability of excellent FET-input op-amps), and on construction techniques (including CAD/CAM). Every table has been revised, some substantially; for example, in Table 4.1 (operational amplifiers) only 65% of the original 120 entries survived, with 135 new op-amps added.

We have used this opportunity to respond to readers’ suggestions and to our own experiences using and teaching from the first edition. Thus we have rewritten the chapter on FETs (it was too complicated) and repositioned it before the chapter on op-amps (which are increasingly of FET construction). We have added a new chapter on low-power and micropower design (both analog and digital), a field both important and neglected. Most of the remaining chapters have been extensively revised. We have added many new tables, including A/D and D/A converters, digital logic components, and low-power devices, and throughout the book we have expanded the number of figures. The

book now contains 78 tables (available separately as *The Horowitz and Hill Component Selection Tables*) and over 1000 figures.

Throughout the revision we have strived to retain the feeling of informality and easy access that made the first edition so successful and popular, both as reference and text. We are aware of the difficulty students often experience when approaching electronics for the first time: the field is densely interwoven, and there is no path of learning that takes you, by logical steps, from neophyte to broadly competent designer. Thus we have added extensive cross-referencing throughout the text; in addition, we have expanded the separate *Laboratory Manual* into a *Student Manual (Student Manual for The Art of Electronics)*, by Thomas C. Hayes and Paul Horowitz, complete with additional worked examples of circuit designs, explanatory material, reading assignments, laboratory exercises, and solutions to selected problems. By offering a student supplement, we have been able to keep this volume concise and rich with detail, as requested by our many readers who use the volume primarily as a reference work.

We hope this new edition responds to all our readers’ needs – both students and practicing engineers. We welcome suggestions and corrections, which should be addressed directly to Paul Horowitz, Physics Department, Harvard University, Cambridge, MA 02138.

In preparing this new edition, we are appreciative of the help we received from Mike Aronson and Brian Matthews (AOX, Inc.), John Greene (University of Cape Town), Jeremy Avigad and Tom Hayes (Harvard University), Peter Horowitz (EVI, Inc.), Don Stern, and Owen Walker. We thank Jim Mobley for his excellent copyediting, Sophia Prybylski and David Tranah of Cambridge University Press for their encouragement and professional dedication, and the never-sleeping typesetters at Rosenlau Publishing Services, Inc. for their masterful composition in *T<sub>E</sub>X*.

Finally, in the spirit of modern jurisprudence, we remind you to read the legal notice here appended.

Paul Horowitz  
Winfield Hill  
March 1989

***Legal notice***

In this book we have attempted to teach the techniques of electronic design, using circuit examples and data that we believe to be accurate. However, the examples, data, and other information are intended solely as teaching aids and should not be used in any particular application without independent testing and verification by the person making the application. Independent testing and verification are especially important in any application in which incorrect functioning could result in personal injury or damage to property.

For these reasons, we make no warranties, express or implied, that the examples, data, or other infor-

mation in this volume are free of error, that they are consistent with industry standards, or that they will meet the requirements for any particular application. THE AUTHORS AND PUBLISHER EXPRESSLY DISCLAIM THE IMPLIED WARRANTIES OF MERCHANTABILITY AND OF FITNESS FOR ANY PARTICULAR PURPOSE, even if the authors have been advised of a particular purpose, and even if a particular purpose is indicated in the book. The authors and publisher also disclaim all liability for direct, indirect, incidental, or consequential damages that result from any use of the examples, data, or other information in this book.

# PREFACE TO THE THIRD EDITION

Moore's Law continues to assert itself, unabated, since the publication of the second edition a quarter century ago. In this new third (and final!) edition we have responded to this upheaval with major enhancements:

- an emphasis on devices and circuits for *A/D* and *D/A conversion* (Chapter 13), because embedded microcontrollers are everywhere
- illustration of specialized peripheral ICs for use with microcontrollers (Chapter 15)
- detailed discussions of logic family choices, and of interfacing logic signals to the real world (Chapters 10 and 12)
- greatly expanded treatment of important topics in the essential analog portion of instrument design:
  - precision circuit design (Chapter 5)
  - low-noise design (Chapter 8)
  - power switching (Chapters 3, 9, and 12)
  - power conversion (Chapter 9)

And we have added many entirely new topics, including:

- digital audio and video (including cable and satellite TV)
- transmission lines
- circuit simulation with SPICE
- transimpedance amplifiers
- depletion-mode MOSFETs
- protected MOSFETs
- high-side drivers
- quartz crystal properties and oscillators
- a full exploration of JFETs
- high-voltage regulators
- optoelectronics
- power logic registers
- delta-sigma converters
- precision multislope conversion
- memory technologies
- serial buses
- illustrative “Designs by the Masters”

In this new edition we have responded, also, to the reality that previous editions have been enthusiastically embraced by the community of practicing circuit designers, even though *The Art of Electronics* (now 35 years in print) originated as a course textbook. So we've continued the “how we do it” approach to circuit design; and we've ex-

panded the depth of treatment, while (we hope) retaining the easy access and explanation of basics. At the same time we have split off some of the specifically course-related teaching and lab material into a separate *Learning the Art of Electronics* volume, a substantial expansion of the previous edition's companion *Student Manual for The Art of Electronics*.<sup>1</sup>

Digital oscilloscopes have made it easy to capture, annotate, and combine measured waveforms, a capability we have exploited by including some 90 'scope screenshots illustrating the behavior of working circuits. Along with those doses of reality, we have included (in tables and graphs) substantial quantities of highly useful measured data – such as transistor noise and gain characteristics ( $e_n$ ,  $i_n$ ,  $r_{bb'}$ ;  $h_{fe}$ ,  $g_m$ ,  $g_{oss}$ ), analog switch characteristics ( $R_{ON}$ ,  $Q_{inj}$ , capacitance), op-amp input and output characteristics ( $e_n$  and  $i_n$  over frequency, input common-mode range, output swing, auto-zero recovery, distortion, available packages), and approximate prices (!) – the sort of data often buried or omitted in datasheets but which you need (and don't have the time to measure) when designing circuits.

We've worked diligently, over the 20 years it has taken to prepare this edition, to include important circuit design information, in the form of some 350 graphs, 50 photographs, and 87 tables (listing more than 1900 active components), the last enabling intelligent choice of circuit components by listing essential characteristics (both specified and measured) of available parts.

Because of the significant expansion of topics and depth of detail, we've had to leave behind some topics that were treated in the second edition,<sup>2</sup> notwithstanding the use of larger pages, more compact fonts, and most figures sized to fit in a single column. Some additional related material that we had hoped to include in this volume (on real-world properties of components, and advanced topics in BJTs, FETs, op-amps, and power control) will instead be published in a forthcoming companion volume, *The Art*

<sup>1</sup> Both by Hayes, T. and Horowitz, P., Cambridge University Press, 1989 and 2016.

<sup>2</sup> Which, however, will continue to be available as an e-book.

*of Electronics: The x-Chapters.* References in this volume to those x-chapter sections and figures are set in italics. A newly updated [artofelectronics.com](http://artofelectronics.com) website will provide a home for a continuation of the previous edition's collections of *Circuit ideas* and *Bad circuits*; it is our hope that it will become a community, also, for a lively electronic circuit forum.

As always, we welcome corrections and suggestions (and, of course, fan mail), which can be sent to [horowitz@physics.harvard.edu](mailto:horowitz@physics.harvard.edu) or to [hill@rowland.harvard.edu](mailto:hill@rowland.harvard.edu).

**With gratitude.** Where to start, in thanking our invaluable colleagues? Surely topping the list is David Tranah, our indefatigable editor at the Cambridge University Press mother-ship, our linchpin, helpful L<sup>A</sup>T<sub>E</sub>Xpert, wise advisor of all things bookish, and (would you believe?) *compositor!* This guy slogged through 1,905 pages of marked-up text, retrofitting the L<sup>A</sup>T<sub>E</sub>X source files with corrections from multiple personalities, then entering a few thousand index entries, and making it all work with its 1,500+ linked figures and tables. And then putting up with a couple of fussy authors. We are totally indebted to David. We owe him a pint of ale.

We are grateful to Jim Macarthur, circuit designer extraordinaire, for his careful reading of chapter drafts, and invariably helpful suggestions for improvement; we adopted every one. Our colleague Peter Lu taught us the delights of Adobe Illustrator, and appeared at a moment's notice when we went off the rails; the book's figures are testament to the quality of his tutoring. And our always-entertaining colleague Jason Gallicchio generously contributed his master Mathematica talents to reveal graphically the properties of delta-sigma conversion, nonlinear control, filter functions; he left his mark, also, in the microcontroller chapter, contributing both wisdom and code.

For their many helpful contributions we thank Bob Adams, Mike Burns, Steve Cerwin, Jesse Colman, Michael Covington, Doug Doskocil, Jon Hagen, Tom Hayes, Phil Hobbs, Peter Horowitz, George Kontopidis, Maggie McFee, Curtis Mead, Ali Mehmed, Angel Peterchev, Jim Phillips, Marco Sartore, Andrew Speck, Jim Thompson, Jim van Zee, GuYeon Wei, John Willison, Jonathan Wolff, John Woodgate, and Woody Yang. We thank also others whom (we're sure) we've here overlooked, with apologies for the omission. Additional contributors to the book's content (circuits, inspired web-based tools, unusual measurements, etc., from the likes of Uwe Beis, Tom Bruhns, and John Larkin) are referenced throughout the book in the relevant text.

Simon Capelin has kept us out of the doldrums with his unflagging encouragement and his apparent inability to scold us for missed deadlines (our contract called for delivery of the finished manuscript in December...of 1994! We're only 20 years late). In the production chain we are indebted to our project manager Peggy Rote, our copy editor Vicki Danahy, and a cast of unnamed graphic artists who converted our pencil circuit sketches into beautiful vector graphics.

We remember fondly our late colleague and friend Jim Williams for wonderful insider stories of circuit failures and circuit conquests, and for his take-no-prisoners approach to precision circuit design. His no-bullshit attitude is a model for us all.

And finally, we are forever indebted to our loving, supportive, and ever-tolerant spouses Vida and Ava, who suffered through decades of abandonment as we obsessed over every detail of our second encore.

**A note on the tools.** Tables were assembled in Microsoft Excel, and graphical data was plotted with Igor Pro; both were then beautified with Adobe Illustrator, with text and annotations in the sans-serif Helvetica Neue LT typeface. Oscilloscope screenshots are from our trusty Tektronix TDS3044 and 3054 "lunchboxes," taken to finishing school in Illustrator, by way of Photoshop. The photographs in the book were taken primarily with two cameras: a Calumet Horseman 6×9 cm view camera with a 105 mm Schneider Symmar f/5.6 lens and Kodak Plus-X 120 roll film (developed in Microdol-X 1:3 at 75°F and digitized with a Mamiya multiformat scanner), and a Canon 5D with a Scheimpflug<sup>3</sup>-enabling 90 mm tilt-shift lens. The authors composed the manuscript in L<sup>A</sup>T<sub>E</sub>X, using the PCT<sub>E</sub>X software from Personal TeX, Incorporated. The text is set in the Times New Roman and Helvetica typefaces, the former dating from 1931,<sup>4</sup> the latter designed in 1957 by Max Miedinger.

Paul Horowitz

Winfield Hill

January 2015

Cambridge, Massachusetts

\* \* \* \* \*

<sup>3</sup> What's that? Google it!

<sup>4</sup> Developed in response to a criticism of the antiquated typeface in *The Times* (London).

**Legal Notice Addendum**

In addition to the Legal Notice appended to the Preface to the Second Edition, we also make no representation regarding whether use of the examples, data, or other information in this volume might infringe others' intellectual property rights, including US and foreign patents. It is the reader's sole responsibility to ensure that he or she is not infringing any intellectual property rights, even for use which is considered to be exper-

imental in nature. By using any of the examples, data, or other information in this volume, the reader has agreed to assume all liability for any damages arising from or relating to such use, regardless of whether such liability is based on intellectual property or any other cause of action, and regardless of whether the damages are direct, indirect, incidental, consequential, or any other type of damage. The authors and publisher disclaim any such liability.

*This page intentionally left blank*

# FOUNDATIONS

## CHAPTER 1

### 1.1 Introduction

The field of electronics is one of the great success stories of the 20th century. From the crude spark-gap transmitters and “cat’s-whisker” detectors at its beginning, the first half-century brought an era of vacuum-tube electronics that developed considerable sophistication and found ready application in areas such as communications, navigation, instrumentation, control, and computation. The latter half-century brought “solid-state” electronics – first as discrete transistors, then as magnificent arrays of them within “integrated circuits” (ICs) – in a flood of stunning advances that shows no sign of abating. Compact and inexpensive consumer products now routinely contain many millions of transistors in VLSI (very large-scale integration) chips, combined with elegant optoelectronics (displays, lasers, and so on); they can process sounds, images, and data, and (for example) permit wireless networking and shirt-pocket access to the pooled capabilities of the Internet. Perhaps as noteworthy is the pleasant trend toward increased performance per dollar.<sup>1</sup> The cost of an electronic microcircuit routinely decreases to a fraction of its initial cost as the manufacturing process is perfected (see Figure 10.87 for an example). In fact, it is often the case that the panel controls and cabinet hardware of an instrument cost more than the electronics inside.

On reading of these exciting new developments in electronics, you may get the impression that you should be able to construct powerful, elegant, yet inexpensive, little gadgets to do almost any conceivable task – all you need to know is how all these miracle devices work. If you’ve had that feeling, this book is for you. In it we have attempted to convey the excitement and know-how of the subject of electronics.

In this chapter we begin the study of the laws, rules of thumb, and tricks that constitute the art of electronics as we see it. It is necessary to begin at the beginning – with talk of voltage, current, power, and the components that make up

electronic circuits. Because you can’t touch, see, smell, or hear electricity, there will be a certain amount of abstraction (particularly in the first chapter), as well as some dependence on such visualizing instruments as oscilloscopes and voltmeters. In many ways the first chapter is also the most mathematical, in spite of our efforts to keep mathematics to a minimum in order to foster a good intuitive understanding of circuit design and behavior.

In this new edition we’ve included some intuition-aiding approximations that our students have found helpful. And, by introducing one or two “active” components ahead of their time, we’re able to jump directly into some applications that are usually impossible in a traditional textbook “passive electronics” chapter; this will keep things interesting, and even exciting.

Once we have considered the foundations of electronics, we will quickly get into the active circuits (amplifiers, oscillators, logic circuits, etc.) that make electronics the exciting field it is. The reader with some background in electronics may wish to skip over this chapter, since it assumes no prior knowledge of electronics. Further generalizations at this time would be pointless, so let’s just dive right in.

### 1.2 Voltage, current, and resistance

#### 1.2.1 Voltage and current

There are two quantities that we like to keep track of in electronic circuits: voltage and current. These are usually changing with time; otherwise nothing interesting is happening.

**Voltage** (symbol  $V$  or sometimes  $E$ ). Officially, the voltage between two points is the cost in energy (work done) required to move a unit of positive charge from the more negative point (lower potential) to the more positive point (higher potential). Equivalently, it is the energy released when a unit charge moves “downhill” from the higher potential to the lower.<sup>2</sup> Voltage is also called

<sup>1</sup> A mid-century computer (the IBM 650) cost \$300,000, weighed 2.7 tons, and contained 126 lamps on its control panel; in an amusing reversal, a contemporary energy-efficient lamp contains a computer of greater capability *within its base*, and costs about \$10.

<sup>2</sup> These are the *definitions*, but hardly the way circuit designers think of voltage. With time, you’ll develop a good intuitive sense of what voltage really is, in an electronic circuit. Roughly (*very roughly*) speaking, voltages are what you apply to cause currents to flow.

*potential difference* or *electromotive force* (EMF). The unit of measure is the *volt*, with voltages usually expressed in volts (V), kilovolts ( $1\text{ kV} = 10^3\text{ V}$ ), millivolts ( $1\text{ mV} = 10^{-3}\text{ V}$ ), or microvolts ( $1\text{ }\mu\text{V} = 10^{-6}\text{ V}$ ) (see the box on prefixes). A joule (J) of work is done in moving a coulomb (C) of charge through a potential difference of 1 V. (The coulomb is the unit of electric charge, and it equals the charge of approximately  $6 \times 10^{18}$  electrons.) For reasons that will become clear later, the opportunities to talk about nanovolts ( $1\text{ nV} = 10^{-9}\text{ V}$ ) and megavolts ( $1\text{ MV} = 10^6\text{ V}$ ) are rare.

**Current** (symbol  $I$ ). Current is the rate of flow of electric charge past a point. The unit of measure is the ampere, or amp, with currents usually expressed in amperes (A), milliamperes ( $1\text{ mA} = 10^{-3}\text{ A}$ ), microamperes ( $1\text{ }\mu\text{A} = 10^{-6}\text{ A}$ ), nanoamperes ( $1\text{ nA} = 10^{-9}\text{ A}$ ), or occasionally picoamperes ( $1\text{ pA} = 10^{-12}\text{ A}$ ). A current of 1 amp equals a flow of 1 coulomb of charge per second. By convention, current in a circuit is considered to flow from a more positive point to a more negative point, even though the actual electron flow is in the opposite direction.

*Important:* from these definitions you can see that currents flow *through* things, and voltages are applied (or appear) *across* things. So you've got to say it right: always refer to the voltage *between* two points or *across* two points in a circuit. Always refer to current *through* a device or connection in a circuit.

To say something like “the voltage through a resistor ...” is nonsense. However, we do frequently speak of the voltage *at a point* in a circuit. This is always understood to mean the voltage between that point and “ground,” a common point in the circuit that everyone seems to know about. Soon you will, too.

We *generate* voltages by doing work on charges in devices such as batteries (conversion of electrochemical energy), generators (conversion of mechanical energy by magnetic forces), solar cells (photovoltaic conversion of the energy of photons), etc. We *get* currents by placing voltages across things.

At this point you may well wonder how to “see” voltages and currents. The single most useful electronic instrument is the oscilloscope, which allows you to look at voltages (or occasionally currents) in a circuit as a function of time.<sup>3</sup> We will deal with oscilloscopes, and also voltmeters, when we discuss signals shortly; for a preview see Appendix O, and the multimeter box later in this chapter.

<sup>3</sup> It has been said that engineers in other disciplines are envious of electrical engineers, because we have such a splendid visualization tool.

In real circuits we connect things together with wires (metallic conductors), each of which has the same voltage on it everywhere (with respect to ground, say).<sup>4</sup> We mention this now so that you will realize that an actual circuit doesn't have to look like its schematic diagram, because wires can be rearranged.

Here are some simple rules about voltage and current:

1. The sum of the currents into a point in a circuit equals the sum of the currents out (conservation of charge). This is sometimes called Kirchhoff's current law (KCL). Engineers like to refer to such a point as a *node*. It follows that, for a series circuit (a bunch of two-terminal things all connected end-to-end), the current is the same everywhere.

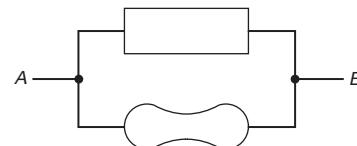


Figure 1.1. Parallel connection.

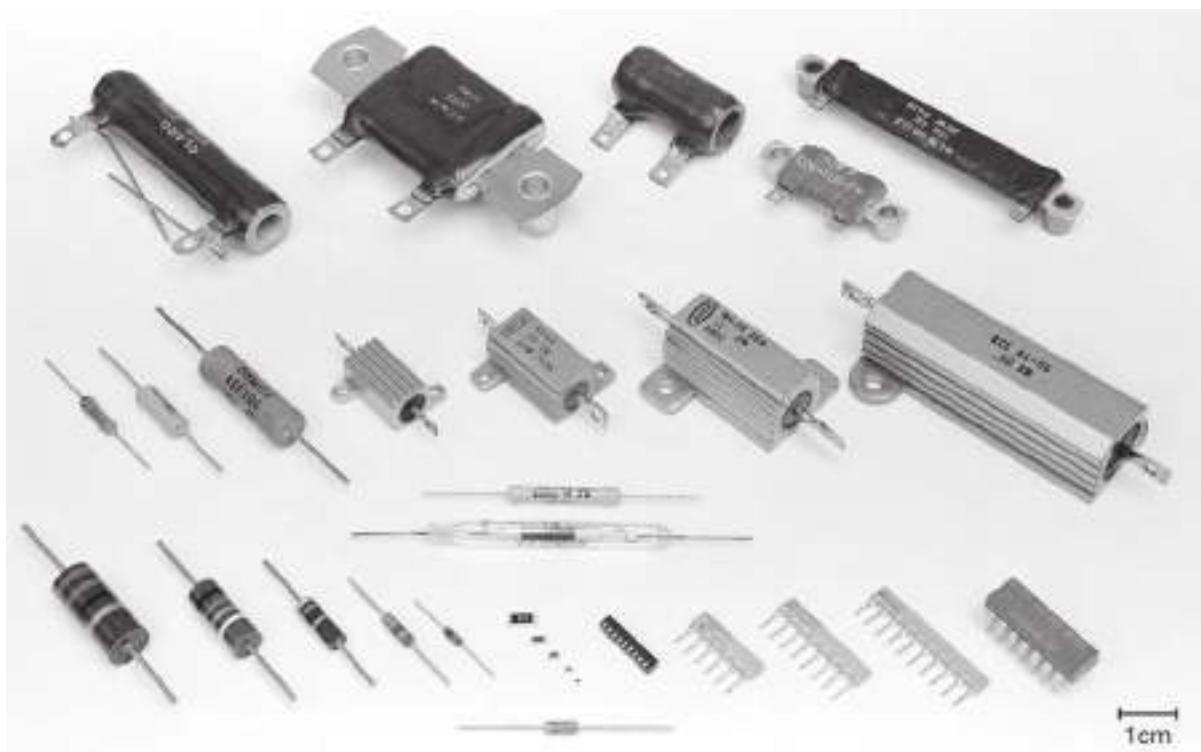
2. Things hooked in parallel (Figure 1.1) have the same voltage across them. Restated, the sum of the “voltage drops” from A to B via one path through a circuit equals the sum by any other route, and is simply the voltage between A and B. Another way to say it is that the sum of the voltage drops around any closed circuit is zero. This is Kirchhoff's voltage law (KVL).
3. The power (energy per unit time) consumed by a circuit device is

$$P = VI \quad (1.1)$$

This is simply  $(\text{energy}/\text{charge}) \times (\text{charge}/\text{time})$ . For  $V$  in volts and  $I$  in amps,  $P$  comes out in watts. A watt is a joule per second ( $1\text{ W} = 1\text{ J/s}$ ). So, for example, the current flowing through a 60W lightbulb running on 120 V is 0.5 A.

Power goes into heat (usually), or sometimes mechanical work (motors), radiated energy (lamps, transmitters), or stored energy (batteries, capacitors, inductors). Managing the heat load in a complicated system (e.g., a large computer, in which many kilowatts of electrical energy are converted to heat, with the energetically insignificant by-product of a few pages of computational results) can be a crucial part of the system design.

<sup>4</sup> In the domain of high frequencies or low impedances, that isn't strictly true, and we will have more to say about this later, and in Chapter IX. For now, it's a good approximation.



**Figure 1.2.** A selection of common resistor types. Top row, left to right (wirewound ceramic power resistors): 20W vitreous enamel with leads, 20W with mounting studs, 30W vitreous enamel, 5W and 20W with mounting studs. Middle row (wirewound power resistors): 1W, 3W, and 5W axial ceramic; 5W, 10W, 25W, and 50W conduction-cooled (“Dale-type”). Bottom row: 2W, 1W,  $\frac{1}{2}$ W,  $\frac{1}{4}$ W, and  $\frac{1}{8}$ W carbon composition; surface-mount thick-film (2010, 1206, 0805, 0603, and 0402 sizes); surface-mount resistor array; 6-, 8-, and 10-pin single in-line package arrays; dual in-line package array. The resistor at bottom is the ubiquitous RN55D  $\frac{1}{4}$ W, 1% metal-film type; and the pair of resistors above are Victoreen high-resistance types (glass, 2 G $\Omega$ ; ceramic, 5 G $\Omega$ ).

Soon, when we deal with periodically varying voltages and currents, we will have to generalize the simple equation  $P = VI$  to deal with *average* power, but it's correct as a statement of *instantaneous* power just as it stands.

Incidentally, don't call current “amperage”; that's strictly bush league.<sup>5</sup> The same caution will apply to the term “ohmage”<sup>6</sup> when we get to resistance in the next section.

### 1.2.2 Relationship between voltage and current: resistors

This is a long and interesting story. It is the heart of electronics. Crudely speaking, the name of the game is to make and use gadgets that have interesting and useful  $I$ -versus- $V$  characteristics. Resistors ( $I$  simply proportional to  $V$ ),

capacitors ( $I$  proportional to rate of change of  $V$ ), diodes ( $I$  flows in only one direction), thermistors (temperature-dependent resistor), photoresistors (light-dependent resistor), strain gauges (strain-dependent resistor), etc., are examples. Perhaps more interesting still are *three-terminal* devices, such as transistors, in which the current that can flow between a pair of terminals is controlled by the voltage applied to a third terminal. We will gradually get into some of these exotic devices; for now, we will start with the most mundane (and most widely used) circuit element, the resistor (Figure 1.3).



**Figure 1.3.** Resistor.

#### A. Resistance and resistors

It is an interesting fact that the current through a metallic conductor (or other partially conducting material) is proportional to the voltage across it. (In the case of wire

<sup>5</sup> Unless you're a power engineer working with giant 13 kV transformers and the like – those guys are allowed to say amperage.

<sup>6</sup> ...also, Dude, “ohmage” is not the preferred nomenclature: *resistance*, please.

## PREFIXES

<i>Multiple</i>	<i>Prefix</i>	<i>Symbol</i>	<i>Derivation</i>
$10^{24}$	yotta	Y	end-1 of Latin alphabet, hint of Greek <i>iota</i>
$10^{21}$	zetta	Z	end of Latin alphabet, hint of Greek <i>zeta</i>
$10^{18}$	exa	E	Greek <i>hexa</i> (six: power of 1000)
$10^{15}$	peta	P	Greek <i>penta</i> (five: power of 1000)
$10^{12}$	tera	T	Greek <i>teras</i> (monster)
$10^9$	giga	G	Greek <i>gigas</i> (giant)
$10^6$	mega	M	Greek <i>megas</i> (great)
$10^3$	kilo	k	Greek <i>khilioi</i> (thousand)
$10^{-3}$	milli	m	Latin <i>milli</i> (thousand)
$10^{-6}$	micro	$\mu$	Greek <i>mikros</i> (small)
$10^{-9}$	nano	n	Greek <i>nanos</i> (dwarf)
$10^{-12}$	pico	p	from Italian/Spanish <i>piccolo/pico</i> (small)
$10^{-15}$	femto	f	Danish/Norwegian <i>femten</i> (fifteen)
$10^{-18}$	atto	a	Danish/Norwegian <i>atten</i> (eighteen)
$10^{-21}$	zepto	z	end of Latin alphabet, mirrors <i>zetta</i>
$10^{-24}$	yocto	y	end-1 of Latin alphabet, mirrors <i>yotta</i>

These prefixes are universally used to scale units in science and engineering. Their etymological derivations are a matter of some controversy and should not be considered historically reliable. When abbreviating a unit with a prefix, the symbol for the unit follows the prefix without space. Be careful about uppercase and lowercase letters (especially m and M) in both prefix and unit: 1 mW

is a milliwatt, or one-thousandth of a watt; 1 MHz is a megahertz or 1 million hertz. In general, units are spelled with lowercase letters, even when they are derived from proper names. The unit name is not capitalized when it is spelled out and used with a prefix, only when abbreviated. Thus: hertz and kilohertz, but Hz and kHz; watt, milliwatt, and megawatt, but W, mW, and MW.

conductors used in circuits, we usually choose a thick-enough gauge of wire so that these “voltage drops” will be negligible.) This is by no means a universal law for all objects. For instance, the current through a neon bulb is a highly nonlinear function of the applied voltage (it is zero up to a critical voltage, at which point it rises dramatically). The same goes for a variety of interesting special devices – diodes, transistors, lightbulbs, etc. (If you are interested in understanding why metallic conductors behave this way, read §§4.4–4.5 in Purcell and Morin’s splendid text *Electricity and Magnetism*).

A resistor is made out of some conducting stuff (carbon, or a thin metal or carbon film, or wire of poor conductivity), with a wire or contacts at each end. It is characterized by its resistance:

$$R = V/I; \quad (1.2)$$

$R$  is in ohms for  $V$  in volts and  $I$  in amps. This is known as Ohm’s law. Typical resistors of the most frequently used type (metal-oxide film, metal film, or carbon film) come in

values from 1 ohm ( $1\Omega$ ) to about 10 megohms ( $10\text{M}\Omega$ ). Resistors are also characterized by how much power they can safely dissipate (the most commonly used ones are rated at 1/4 or 1/8 W), their physical size,<sup>7</sup> and by other parameters such as tolerance (accuracy), temperature coefficient, noise, voltage coefficient (the extent to which  $R$  depends on applied  $V$ ), stability with time, inductance, etc. See the box on resistors, Chapter 1x, and Appendix C for further details. Figure 1.2 shows a collection of resistors, with most of the available morphologies represented.

Roughly speaking, resistors are used to convert a

<sup>7</sup> The sizes of *chip resistors* and other components intended for surface mounting are specified by a four-digit size code, in which each pair of digits specifies a dimension in units of  $0.010''$  (0.25 mm). For example, an 0805 size resistor is  $2\text{mm} \times 1.25\text{ mm}$ , or  $80\text{ mils} \times 50\text{ mils}$  (1 mil is  $0.001''$ ); the height must be separately specified. To add confusion to this simple scheme, the four-digit size code may instead be *metric* (sometimes without saying so!), in units of 0.1 mm: thus an “0805” (English) is also a “2012” (metric).

## RESISTORS

Resistors are truly ubiquitous. There are almost as many types as there are applications. Resistors are used in amplifiers as loads for active devices, in bias networks, and as feedback elements. In combination with capacitors they establish time constants and act as filters. They are used to set operating currents and signal levels. Resistors are used in power circuits to reduce voltages by dissipating power, to measure currents, and to discharge capacitors after power is removed. They are used in precision circuits to establish currents, to provide accurate voltage ratios, and to set precise gain values. In logic circuits they act as bus and line terminators and as “pullup” and “pull-down” resistors. In high-voltage circuits they are used to measure voltages and to equalize leakage currents among diodes or capacitors connected in series. In radiofrequency (RF) circuits they set the bandwidth of resonant circuits, and they are even used as coil forms for inductors.

Resistors are available with resistances from  $0.0002\ \Omega$  through  $10^{12}\ \Omega$ , standard power ratings from  $1/8$  watt through 250 watts, and accuracies from 0.005% through 20%. Resistors can be made from metal films, metal-oxide films, or carbon films; from carbon-composition or

ceramic-composition moldings; from metal foil or metal wire wound on a form; or from semiconductor elements similar to field-effect transistors (FETs). The most commonly used resistor type is formed from a carbon, metal, or oxide film, and comes in two widely used “packages”: the cylindrical *axial-lead* type (typified by the generic RN55D 1%  $1/4$  W metal-film resistor),<sup>8</sup> and the much smaller *surface-mount* “chip resistor.” These common types come in 5%, 2%, and 1% tolerances, in a standard set of values ranging from  $1\ \Omega$  to  $10\ M\Omega$ . The 1% types have 96 values per decade, whereas the 2% and 5% types have 48 and 24 values per decade (see Appendix C). Figure 1.2 illustrates most of the common resistor packages.

Resistors are so easy to use and well behaved that they’re often taken for granted. They’re not perfect, though, and you should be aware of some of their limitations so that you won’t be surprised someday. The principal defects are variations in resistance with temperature, voltage, time, and humidity. Other defects relate to inductance (which may be serious at high frequencies), the development of thermal hot spots in power applications, or electrical noise generation in low-noise amplifiers. We treat these in the advanced Chapter 1x.

voltage to a current, and vice versa. This may sound awfully trite, but you will soon see what we mean.

### B. Resistors in series and parallel

From the definition of  $R$ , some simple results follow:

1. The resistance of two resistors in series (Figure 1.4) is

$$R = R_1 + R_2. \quad (1.3)$$

By putting resistors in series, you always get a *larger* resistor.

2. The resistance of two resistors in parallel (Figure 1.5) is

$$R = \frac{R_1 R_2}{R_1 + R_2} \quad \text{or} \quad R = \frac{1}{\frac{1}{R_1} + \frac{1}{R_2}}. \quad (1.4)$$

By putting resistors in parallel, you always get a *smaller* resistor. Resistance is measured in ohms ( $\Omega$ ), but in practice we frequently omit the  $\Omega$  symbol when referring to resistors that are more than  $1000\ \Omega$  ( $1\ k\Omega$ ). Thus, a  $4.7\ k\Omega$  resistor is often referred to as a  $4.7\text{k}$  resistor, and a  $1\ M\Omega$



Figure 1.4. Resistors in series.

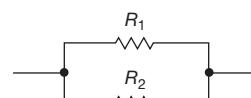


Figure 1.5. Resistors in parallel.

resistor as a  $1\text{M}$  resistor (or  $1\text{ meg}$ ).<sup>9</sup> If these preliminaries bore you, please have patience – we’ll soon get to numerous amusing applications.

**Exercise 1.1.** You have a  $5\text{k}$  resistor and a  $10\text{k}$  resistor. What is their combined resistance (a) in series and (b) in parallel?

**Exercise 1.2.** If you place a  $1\ \Omega$  resistor across a  $12$  volt car battery, how much power will it dissipate?

**Exercise 1.3.** Prove the formulas for series and parallel resistors.

<sup>8</sup> Conservatively rated at  $1/8$  watt in its RN55 military grade (“MIL-spec”), but rated at  $1/4$  watt in its CMF-55 industrial grade.

<sup>9</sup> A popular “international” alternative notation replaces the decimal point with the unit multiplier, thus  $4k7$  or  $1M0$ . A  $2.2\ \Omega$  resistor becomes  $2R2$ . There is an analogous scheme for capacitors and inductors.

**Exercise 1.4.** Show that several resistors in parallel have resistance

$$R = \frac{1}{\frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3} + \dots} \quad (1.5)$$

Beginners tend to get carried away with complicated algebra in designing or trying to understand electronics. Now is the time to begin learning intuition and shortcuts. Here are a couple of good tricks:

**Shortcut #1** A large resistor in series (parallel) with a small resistor has the resistance of the larger (smaller) one, roughly. So you can “trim” the value of a resistor up or down by connecting a second resistor in series or parallel: to trim *up*, choose an available resistor value below the target value, then add a (much smaller) series resistor to make up the difference; to trim *down*, choose an available resistor value above the target value, then connect a (much larger) resistor in parallel. For the latter you can approximate with proportions – to lower the value of a resistor by 1%, say, put a resistor 100 times as large in parallel.<sup>10</sup>

**Shortcut #2** Suppose you want the resistance of 5k in parallel with 10k. If you think of the 5k as two 10k's in parallel, then the whole circuit is like three 10k's in parallel. Because the resistance of  $n$  equal resistors in parallel is  $1/n$ th the resistance of the individual resistors, the answer in this case is 10k/3, or 3.33k. This trick is handy because it allows you to analyze circuits quickly in your head, without distractions. We want to encourage mental designing, or at least “back-of-the-envelope” designing, for idea brainstorming.

Some more home-grown philosophy: there is a tendency among beginners to want to compute resistor values and other circuit component values to many significant places, particularly with calculators and computers that readily oblige. There are two reasons you should try to avoid falling into this habit: (a) the components themselves are of finite precision (resistors typically have tolerances of  $\pm 5\%$  or  $\pm 1\%$ ; for capacitors it's typically  $\pm 10\%$  or  $\pm 5\%$ ; and the parameters that characterize transistors, say, frequently are known only to a factor of 2); (b) one mark of a good circuit design is insensitivity of the finished circuit to precise values of the components (there are exceptions, of course). You'll also learn circuit intuition more quickly if you get into the habit of doing approximate calculations in your head, rather than watching meaningless numbers pop up on a calculator display. We believe strongly that reliance on formulas and equations early in your electronic circuit

education is a fine way to prevent you from understanding what's really going on.

In trying to develop intuition about resistance, some people find it helpful to think about *conductance*,  $G = 1/R$ . The current through a device of conductance  $G$  bridging a voltage  $V$  is then given by  $I = GV$  (Ohm's law). A small resistance is a large conductance, with correspondingly large current under the influence of an applied voltage. Viewed in this light, the formula for parallel resistors is obvious: when several resistors or conducting paths are connected across the same voltage, the total current is the sum of the individual currents. Therefore the net conductance is simply the sum of the individual conductances,  $G = G_1 + G_2 + G_3 + \dots$ , which is the same as the formula for parallel resistors derived earlier.

Engineers are fond of defining reciprocal units, and they have designated as the unit of conductance the siemens ( $S = 1/\Omega$ ), also known as the mho (that's ohm spelled backward, given the symbol  $\mathcal{S}$ ). Although the concept of conductance is helpful in developing intuition, it is not used widely;<sup>11</sup> most people prefer to talk about resistance instead.

### C. Power in resistors

The power dissipated by a resistor (or any other device) is  $P = IV$ . Using Ohm's law, you can get the equivalent forms  $P = I^2R$  and  $P = V^2/R$ .

**Exercise 1.5.** Show that it is not possible to exceed the power rating of a 1/4 watt resistor of resistance greater than 1k, no matter how you connect it, in a circuit operating from a 15 volt battery.

**Exercise 1.6.** Optional exercise: New York City requires about  $10^{10}$  watts of electrical power, at 115 volts<sup>12</sup> (this is plausible: 10 million people averaging 1 kilowatt each). A heavy power cable might be an inch in diameter. Let's calculate what will happen if we try to supply the power through a cable 1 foot in diameter made of pure copper. Its resistance is  $0.05 \mu\Omega$  ( $5 \times 10^{-8}$  ohms) per foot. Calculate (a) the power lost per foot from “ $I^2R$  losses,” (b) the length of cable over which you will lose all  $10^{10}$  watts, and (c) how hot the cable will get, if you know the physics involved ( $\sigma = 6 \times 10^{-12} \text{ W/K}^4 \text{ cm}^2$ ). If you have done your computations correctly, the result should seem preposterous. What is the solution to this puzzle?

<sup>11</sup> Although the elegant *Millman's theorem* has its admirers: it says that the output voltage from a set of resistors (call them  $R_i$ ) that are driven from a set of corresponding input voltages ( $V_i$ ) and connected together at the output is  $V_{\text{out}} = (\sum V_i G_i) / \sum G_i$ , where the  $G_i$  are the conductances ( $G_i = 1/R_i$ ).

<sup>12</sup> Although the “official” line voltage is 120 V  $\pm 5\%$ , you'll sometimes see 110 V, 115 V, or 117 V. This loose language is OK (and we use it in this book), because (a) the median voltage at the wall plug is 3 to 5 volts lower, when powering stuff; and (b) the *minimum* wall-plug voltage is 110 V. See ANSI standard C84.1.

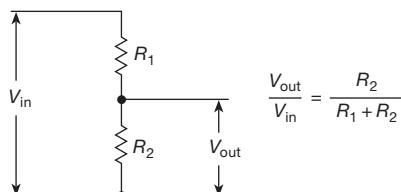
<sup>10</sup> With an error, in this case, of just 0.01%.

## D. Input and output

Nearly all electronic circuits accept some sort of applied *input* (usually a voltage) and produce some sort of corresponding *output* (which again is often a voltage). For example, an audio amplifier might produce a (varying) output voltage that is 100 times as large as a (similarly varying) input voltage. When describing such an amplifier, we imagine measuring the output voltage for a given applied input voltage. Engineers speak of the *transfer function*  $\mathbf{H}$ , the ratio of (measured) output divided by (applied) input; for the audio amplifier above,  $\mathbf{H}$  is simply a constant ( $\mathbf{H} = 100$ ). We'll get to amplifiers soon enough, in the next chapter. However, with only resistors we can already look at a very important circuit fragment, the *voltage divider* (which you might call a "de-amplifier").

### 1.2.3 Voltage dividers

We now come to the subject of the voltage divider, one of the most widespread electronic circuit fragments. Show us any real-life circuit and we'll show you half a dozen voltage dividers. To put it very simply, a voltage divider is a circuit that, given a certain voltage input, produces a predictable fraction of the input voltage as the output voltage. The simplest voltage divider is shown in Figure 1.6.

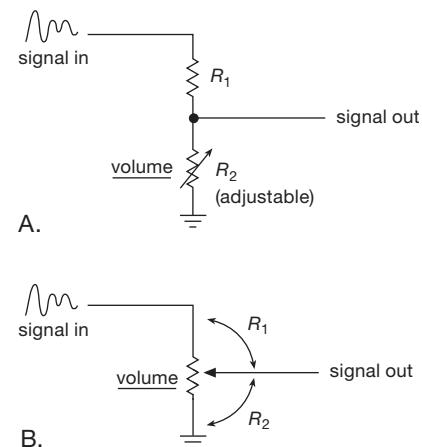


**Figure 1.6.** Voltage divider. An applied voltage  $V_{\text{in}}$  results in a (smaller) output voltage  $V_{\text{out}}$ .

An important word of explanation: when engineers draw a circuit like this, it's generally assumed that the  $V_{\text{in}}$  on the left is a voltage that you are applying to the circuit, and that the  $V_{\text{out}}$  on the right is the resulting output voltage (produced by the circuit) that you are measuring (or at least are interested in). You are supposed to know all this (a) because of the convention that signals generally flow from left to right, (b) from the suggestive names ("in," "out") of the signals, and (c) from familiarity with circuits like this. This may be confusing at first, but with time it becomes easy.

What is  $V_{\text{out}}$ ? Well, the current (same everywhere, assuming no "load" on the output; i.e., nothing connected across the output) is

$$I = \frac{V_{\text{in}}}{R_1 + R_2}.$$



**Figure 1.7.** An adjustable voltage divider can be made from a fixed and variable resistor, or from a potentiometer. In some contemporary circuits you'll find instead a long series chain of equal-value resistors, with an arrangement of electronic switches that lets you choose any one of the junctions as the output; this sounds much more complicated – but it has the advantage that you can adjust the voltage ratio *electrically* (rather than mechanically).

(We've used the definition of resistance and the series law.) Then, for  $R_2$ ,

$$V_{\text{out}} = IR_2 = \frac{R_2}{R_1 + R_2} V_{\text{in}}. \quad (1.6)$$

Note that the output voltage is always less than (or equal to) the input voltage; that's why it's called a divider. You could get amplification (more output than input) if one of the resistances were negative. This isn't as crazy as it sounds; it is possible to make devices with negative "incremental" resistances (e.g., the component known as a *tunnel diode*) or even true negative resistances (e.g., the negative-impedance converter that we will talk about later in the book, §6.2.4B). However, these applications are rather specialized and need not concern you now.

Voltage dividers are often used in circuits to generate a particular voltage from a larger fixed (or varying) voltage. For instance, if  $V_{\text{in}}$  is a varying voltage and  $R_2$  is an adjustable resistor (Figure 1.7A), you have a "volume control"; more simply, the combination  $R_1R_2$  can be made from a single variable resistor, or *potentiometer* (Figure 1.7B). This and similar applications are common, and potentiometers come in a variety of styles, some of which are shown in Figure 1.8.

The humble voltage divider is even more useful, though, as a way of *thinking* about a circuit: the input voltage and upper resistance might represent the output of an amplifier, say, and the lower resistance might represent the input of



**Figure 1.8.** Most of the common potentiometer styles are shown here. Top row, left to right (panel mount): power wirewound, “type AB” 2W carbon composition, 10-turn wirewound/plastic hybrid, ganged dual pot. Middle row (panel mount): optical encoder (continuous rotation, 128 cycles per turn), single-turn cermet, single-turn carbon, screw-adjust single-turn locking. Front row (board-mount trimmers): multturn side-adjust (two styles), quad single-turn, 3/8" (9.5 mm) square single-turn, 1/4" (6.4 mm) square single-turn, 1/4" (6.4 mm) round single-turn, 4 mm square single-turn surface mount, 4 mm square multturn surface mount, 3/8" (9.5 mm) square multturn, quad nonvolatile 256-step integrated pot (E<sup>2</sup>POT) in 24-pin small-outline IC.

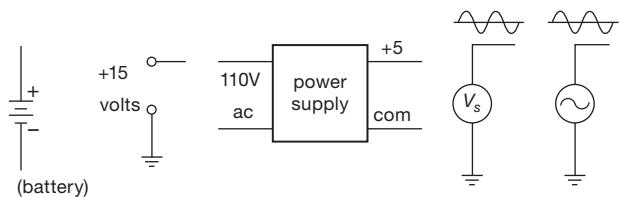
the following stage. In this case the voltage-divider equation tells you how much signal gets to the input of that last stage. This will all become clearer after you know about a remarkable fact (Thévenin’s theorem) that will be discussed later. First, though, a short aside on voltage sources and current sources.

#### 1.2.4 Voltage sources and current sources

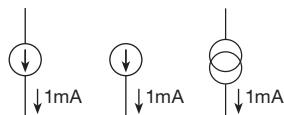
A perfect *voltage source* is a two-terminal “black box” that maintains a fixed voltage drop across its terminals, regardless of load resistance. This means, for instance, that it must supply a current  $I = V/R$  when a resistance  $R$  is attached to its terminals. A real voltage source can supply only a finite maximum current, and in addition it generally behaves like a perfect voltage source with a small resistance in series. Obviously, the smaller this series resistance, the better. For example, a standard 9 volt alkaline battery behaves approximately like a perfect 9 volt voltage source in series with a  $3\ \Omega$  resistor, and it can provide a maximum

current (when shorted) of 3 amps (which, however, will kill the battery in a few minutes). A voltage source “likes” an open-circuit load and “hates” a short-circuit load, for obvious reasons. (The meaning of “open circuit” and “short circuit” sometimes confuse the beginner: an open circuit has nothing connected to it, whereas a short circuit is a piece of wire bridging the output.) The symbols used to indicate a voltage source are shown in Figure 1.9.

A perfect *current source* is a two-terminal black box that maintains a constant current through the external circuit, regardless of load resistance or applied voltage. To do this it must be capable of supplying any necessary voltage across its terminals. Real current sources (a much-neglected subject in most textbooks) have a limit to the voltage they can provide (called the *output-voltage compliance*, or just *compliance*), and in addition they do not provide absolutely constant output current. A current source “likes” a short-circuit load and “hates” an open-circuit load. The symbols used to indicate a current source are shown in Figure 1.10.

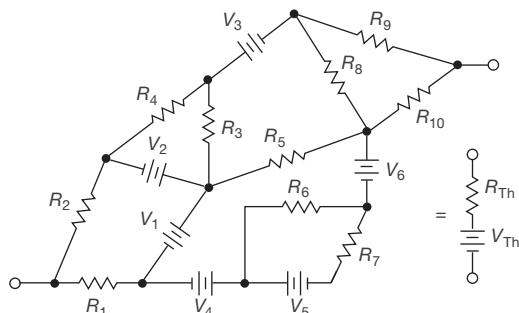


**Figure 1.9.** Voltage sources can be either steady (dc) or varying (ac).



**Figure 1.10.** Current-source symbols.

A battery is a real-life approximation to a voltage source (there is no analog for a current source). A standard D-size flashlight cell, for instance, has a terminal voltage of 1.5 V, an equivalent series resistance of about  $0.25 \Omega$ , and a total energy capacity of about 10,000 watt–seconds (its characteristics gradually deteriorate with use; at the end of its life, the voltage may be about 1.0 V, with an internal series resistance of several ohms). It is easy to construct voltage sources with far better characteristics, as you will learn when we come to the subject of feedback; this is a major topic of Chapter 9. Except in the important class of devices intended for portability, the use of batteries in electronic devices is rare.



**Figure 1.11.** The Thévenin equivalent circuit.

## 1.2.5 Thévenin equivalent circuit

Thévenin's theorem states<sup>12</sup> that any two-terminal network of resistors and voltage sources is equivalent to a single

resistor  $R$  in series with a single voltage source  $V$ . This is remarkable. Any mess of batteries and resistors can be mimicked with one battery and one resistor (Figure 1.11). (Incidentally, there's another theorem, Norton's theorem, that says you can do the same thing with a current source in parallel with a resistor.)

How do you figure out the Thévenin equivalent  $R_{\text{Th}}$  and  $V_{\text{Th}}$  for a given circuit? Easy!  $V_{\text{Th}}$  is the open-circuit voltage of the Thévenin equivalent circuit; so if the two circuits behave identically, it must also be the open-circuit voltage of the given circuit (which you get by calculation, if you know what the circuit is, or by measurement, if you don't). Then you find  $R_{\text{Th}}$  by noting that the short-circuit current of the equivalent circuit is  $V_{\text{Th}}/R_{\text{Th}}$ . In other words,

$$V_{\text{Th}} = V \text{ (open circuit)},$$

$$R_{\text{Th}} = \frac{V \text{ (open circuit)}}{I \text{ (short circuit)}}. \quad (1.7)$$

Let's apply this method to the voltage divider, which must have a Thévenin equivalent:

1. The open-circuit voltage is

$$V = V_{\text{in}} \frac{R_2}{R_1 + R_2}.$$

2. The short-circuit current is

$$V_{\text{in}}/R_1.$$

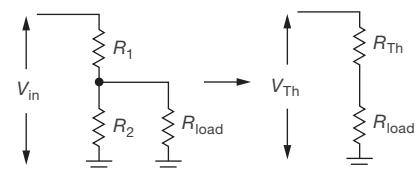
So the Thévenin equivalent circuit is a voltage source,

$$V_{\text{Th}} = V_{\text{in}} \frac{R_2}{R_1 + R_2}, \quad (1.8)$$

in series with a resistor,

$$R_{\text{Th}} = \frac{R_1 R_2}{R_1 + R_2}. \quad (1.9)$$

(It is not a coincidence that this happens to be the parallel resistance of  $R_1$  and  $R_2$ . The reason will become clear later.)



**Figure 1.12.** Thévenin equivalent of a voltage divider.

From this example it is easy to see that a voltage divider is not a very good battery, in the sense that its output voltage drops severely when a load is attached. As an example, consider Exercise 1.10. You now know everything you need to know to calculate exactly how much the output will

<sup>12</sup> We provide a proof, for those who are interested, in Appendix D.

## MULTIMETERS

There are numerous instruments that let you measure voltages and currents in a circuit. The oscilloscope is the most versatile; it lets you “see” voltages versus time at one or more points in a circuit. Logic probes and logic analyzers are special-purpose instruments for troubleshooting digital circuits. The simple multimeter provides a good way to measure voltage, current, and resistance, often with good precision; however, it responds slowly, and thus it cannot replace the oscilloscope where changing voltages are of interest. Multimeters are of two varieties: those that indicate measurements on a conventional scale with a moving pointer, and those that use a digital display.

The traditional (and now largely obsolete) VOM (volt-ohm-milliammeter) multimeter uses a meter movement that measures current (typically  $50\ \mu\text{A}$  full scale). (See a less-design-oriented electronics book for pretty pictures of the innards of meter movements; for our purposes, it suffices to say that it uses coils and magnets.) To measure *voltage*, the VOM puts a resistor in series with the basic movement. For instance, one kind of VOM will generate a 1 V (full-scale) range by putting a 20k resistor in series with the standard  $50\ \mu\text{A}$  movement; higher voltage ranges use correspondingly larger resistors. Such a VOM is specified as  $20,000\ \Omega/\text{V}$ , meaning that it looks like a resistor whose value is 20k multiplied by the full-scale voltage of the particular range selected. Full scale on any voltage range is  $1/20,000$  amps, or  $50\ \mu\text{A}$ . It should be clear that one of these voltmeters disturbs a circuit less on a higher range, since it looks like a higher resistance (think of the voltmeter as the lower leg of a voltage divider, with the Thévenin resistance of the circuit you are measuring as the upper resistor). Ideally, a voltmeter should have infinite input resistance.

Most contemporary multimeters use electronic amplification and have an input resistance of  $10\ \text{M}\Omega$  to  $1000\ \text{M}\Omega$  when measuring voltage; they display their results digitally, and are known collectively as digital multimeters (DMMs). A word of caution: sometimes the input resistance of these meters is very high on the most sensitive ranges, dropping to a lower resistance for the higher ranges. For instance, you might typically have an input resistance of  $10^9\ \Omega$  on the 0.2 V and 2 V ranges, and  $10^7\ \Omega$  on all higher ranges. Read the specifications carefully! However, for most circuit measurements these high input resistances will produce negligible loading effects. In any case, it is easy to calculate how serious the effect is by using the voltage-divider equation. Typically, multimeters provide voltage ranges from a volt (or less) to a kilovolt (or more), full scale.

A multimeter usually includes *current-measuring* capability, with a

set of switchable ranges. Ideally, a current-measuring meter should have zero resistance<sup>13</sup> in order not to disturb the circuit under test, since it must be put in series with the circuit. In practice, you tolerate a few tenths of a volt drop (sometimes called “voltage burden”) with both VOMs and digital multimeters. For either kind of meter, selecting a current range puts a small resistor across the meter’s input terminals, typically of resistance value to create a voltage drop of 0.1 V to 0.25 V for the chosen full-scale current; the voltage drop is then converted to a corresponding current indication.<sup>14</sup> Typically, multimeters provide current ranges from  $50\ \mu\text{A}$  (or less) to an amp (or more), full scale.

Multimeters also have one or more batteries in them to power the resistance measurement. By supplying a small current and measuring the voltage drop, they measure resistance, with several ranges to cover values from  $1\ \Omega$  (or less) to  $10\ \text{M}\Omega$  (or more).

*Important:* don’t try to measure “the current of a voltage source,” by sticking the meter across the wall plug; the same applies for ohms. This is a leading cause of blown-out meters.

**Exercise 1.7.** What will a  $20,000\ \Omega/\text{V}$  meter read, on its 1 V scale, when attached to a 1 V source with an internal resistance of  $10\text{k}$ ? What will it read when attached to a  $10\text{k}$ – $10\text{k}$  voltage divider driven by a “stiff” (zero source resistance) 1 V source?

**Exercise 1.8.** A  $50\ \mu\text{A}$  meter movement has an internal resistance of  $5\text{k}$ . What shunt resistance is needed to convert it to a 0–1 A meter? What series resistance will convert it to a 0–10 V meter?

**Exercise 1.9.** The very high internal resistance of *digital* multimeters, in their voltage-measuring ranges, can be used to measure extremely low *currents* (even though the DMM may not offer a low current range explicitly). Suppose, for example, you want to measure the small current that flows through a  $1000\ \text{M}\Omega$  “leakage” resistance (that term is used to describe a small current that ideally should be absent entirely, for example through the insulation of an underground cable). You have available a standard DMM, whose 2 V dc range has  $10\ \text{M}\Omega$  internal resistance, and you have available a dc source of +10 V. How can you use what you’ve got to measure accurately the leakage resistance?

drop for a given load resistance: use the Thévenin equivalent circuit, attach a load, and calculate the new output, noting that the new circuit is nothing but a voltage divider (Figure 1.12).

**Exercise 1.10.** For the circuit shown in Figure 1.12, with

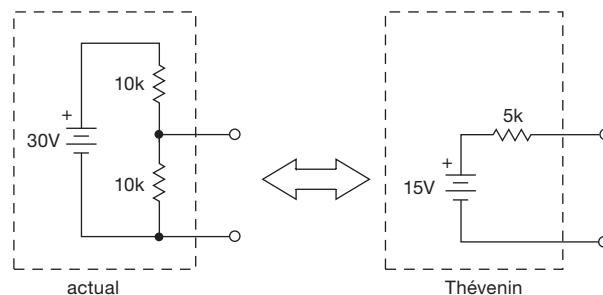
$V_{\text{in}}=30\ \text{V}$  and  $R_1 = R_2 = 10\text{k}$ , find (a) the output voltage with no load attached (the open-circuit voltage); (b) the output voltage with a  $10\text{k}$  load (treat as a voltage divider, with  $R_2$  and  $R_{\text{load}}$  combined into a single resistor); (c) the Thévenin equivalent circuit; (d) the same as in part (b), but using the Thévenin equivalent circuit [again, you wind up with a voltage divider; the answer should agree with the result in part (b)]; (e) the power dissipated in each of the resistors.

### A. Equivalent source resistance and circuit loading

As we have just seen, a voltage divider powered from some fixed voltage is equivalent to some smaller voltage source

<sup>13</sup> This is the opposite of an ideal voltage-measuring meter, which should present infinite resistance across its input terminals.

<sup>14</sup> A special class of current meters known as *electrometers* operate with very small voltage burdens (as little as  $0.1\ \text{mV}$ ) by using the technique of feedback, something we’ll learn about in Chapters 2 and 4.



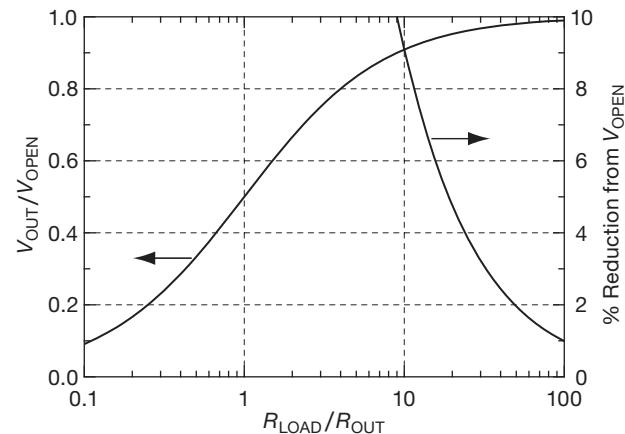
**Figure 1.13.** Voltage divider example.

in series with a resistor. For example, the output terminals of a 10k–10k voltage divider driven by a perfect 30 volt battery are precisely equivalent to a perfect 15 volt battery in series with a 5k resistor (Figure 1.13). Attaching a load resistor causes the voltage divider's output to drop, owing to the finite *source resistance* (Thévenin equivalent resistance of the voltage divider output, viewed as a source of voltage). This is often undesirable. One solution to the problem of making a stiff voltage source (“stiff” is used in this context to describe something that doesn't bend under load) might be to use much smaller resistors in a voltage divider. Occasionally this brute-force approach is useful. However, it is usually best to construct a voltage source, or power supply, as it's commonly called, using active components like transistors or operational amplifiers, which we will treat in Chapters 2–4. In this way you can easily make a voltage source with internal (Thévenin equivalent) resistance as small as milliohms (thousandths of an ohm), without the large currents and dissipation of power characteristic of a low-resistance voltage divider delivering the same performance. In addition, with an active power supply it is easy to make the output voltage adjustable. These topics are treated extensively in Chapter 9.

The concept of equivalent internal resistance applies to all sorts of sources, not just batteries and voltage dividers. Signal sources (e.g., oscillators, amplifiers, and sensing devices) all have an equivalent internal resistance. Attaching a load whose resistance is less than or even comparable to the internal resistance will reduce the output considerably. This undesirable reduction of the open-circuit voltage (or signal) by the load is called “circuit loading.” Therefore you should strive to make  $R_{\text{load}} \gg R_{\text{internal}}$ , because a high-resistance load has little attenuating effect on the source (Figure 1.14).<sup>15</sup> We will see numerous circuit

examples in the chapters ahead. This high-resistance condition ideally characterizes measuring instruments such as voltmeters and oscilloscopes.

A word on language: you frequently hear things like “the resistance looking into the voltage divider” or “the output sees a load of so-and-so many ohms,” as if circuits had eyes. It's OK (in fact, it's a rather good way of keeping straight which resistance you're talking about) to say what part of the circuit is doing the “looking.”<sup>16</sup>



**Figure 1.14.** To minimize the attenuation of a signal source below its open-circuit voltage, keep the load resistance large compared with the output resistance.

## B. Power transfer

Here is an interesting problem: what load resistance will result in maximum power being transferred to the load for a given source resistance? (The terms *source resistance*, *internal resistance*, and *Thévenin equivalent resistance* all mean the same thing.) It is easy to see that either  $R_{\text{load}}=0$  or  $R_{\text{load}}=\infty$  results in zero power transferred, because  $R_{\text{load}}=0$  means that  $V_{\text{load}}=0$  and  $I_{\text{load}}=V_{\text{source}}/R_{\text{source}}$ , so that  $P_{\text{load}}=V_{\text{load}}I_{\text{load}}=0$ . But  $R_{\text{load}}=\infty$  means that  $V_{\text{load}}=V_{\text{source}}$  and  $I_{\text{load}}=0$ , so that again  $P_{\text{load}}=0$ . There has to be a maximum in between.

**Exercise 1.11.** Show that  $R_{\text{load}} = R_{\text{source}}$  maximizes the power in the load for a given source resistance. Note: skip this exercise if you don't know calculus, and take it on faith that the answer is true.

dio frequencies and transmission lines, you must “match impedances” (i.e., set  $R_{\text{load}}=R_{\text{internal}}$ ) in order to prevent reflection and loss of power. See Appendix H on transmission lines.

<sup>15</sup> There are two important exceptions to this general principle: (1) a current source has a high (ideally infinite) internal resistance and should drive a load of relatively low load resistance; (2) when dealing with ra-

dio frequencies and transmission lines, you must “match impedances” (i.e., set  $R_{\text{load}}=R_{\text{internal}}$ ) in order to prevent reflection and loss of power. See Appendix H on transmission lines.

<sup>16</sup> The urge to anthropomorphize runs deep in the engineering and scientific community, despite warnings like “don't anthropomorphize computers . . . they don't like it.”

Lest this example leave the wrong impression, we would like to emphasize again that circuits are ordinarily designed so that the load resistance is much greater than the source resistance of the signal that drives the load.

### 1.2.6 Small-signal resistance

We often deal with electronic devices for which  $I$  is not proportional to  $V$ ; in such cases there's not much point in talking about resistance, since the ratio  $V/I$  will depend on  $V$ , rather than being a nice constant, independent of  $V$ . For these devices it is sometimes useful to know instead the *slope* of the  $V$ - $I$  curve, in other words, the ratio of a small change in applied voltage to the resulting change in current through the device,  $\Delta V/\Delta I$  (or  $dV/dI$ ). This quantity has the units of resistance (ohms) and substitutes for resistance in many calculations. It is called the small-signal resistance, incremental resistance, or dynamic resistance.

#### A. Zener diodes

As an example, consider the *zener diode*, which has the  $I$ - $V$  curve shown in Figure 1.15. Zeners are used to create a constant voltage inside a circuit somewhere, simply done by providing them with a (roughly constant) current derived from a higher voltage within the circuit.<sup>17</sup> For example, the zener diode in Figure 1.15 will convert an applied current in the range shown to a corresponding (but fractionally narrower) range of voltages. It is important to know how the resulting zener voltage will change with applied current; this is a measure of its “regulation” against changes in the driving current provided to it. Included in the specifications of a zener will be its dynamic resistance, given at a certain current. For example, a zener might have a dynamic resistance of  $10\Omega$  at  $10\text{ mA}$ , at its specified zener voltage of  $5\text{ V}$ . Using the definition of dynamic resistance, we find that a  $10\%$  change in applied current will therefore result in a change in voltage of

$$\Delta V = R_{\text{dyn}}\Delta I = 10 \times 0.1 \times 0.01 = 10\text{ mV}$$

or

$$\Delta V/V = 0.002 = 0.2\%,$$

thus demonstrating good voltage-regulating ability. In this sort of application you frequently get the zener current

<sup>17</sup> Zeners belong to the more general class of *diodes* and *rectifiers*, important devices that we'll see later in the chapter (§1.6), and indeed throughout the book. The ideal diode (or rectifier) acts as a perfect conductor for current flow in one direction, and a perfect insulator for current flow in the reverse direction; it is a “one-way valve” for current.

through a resistor from a higher voltage available somewhere in the circuit, as in Figure 1.16.

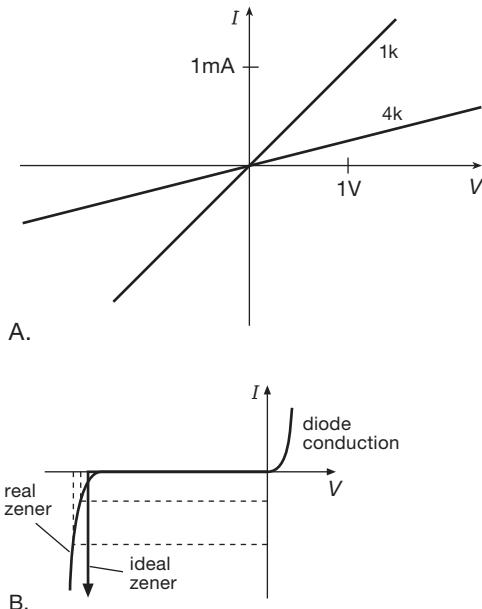


Figure 1.15.  $I$ - $V$  curves: A. Resistor (linear). B. Zener diode (non-linear).

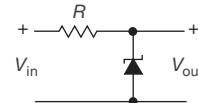


Figure 1.16. Zener regulator.

Then,

$$I = \frac{V_{\text{in}} - V_{\text{out}}}{R}$$

and

$$\Delta I = \frac{\Delta V_{\text{in}} - \Delta V_{\text{out}}}{R},$$

so

$$\Delta V_{\text{out}} = R_{\text{dyn}}\Delta I = \frac{R_{\text{dyn}}}{R}(\Delta V_{\text{in}} - \Delta V_{\text{out}})$$

and finally

$$\Delta V_{\text{out}} = \frac{R_{\text{dyn}}}{R + R_{\text{dyn}}} \Delta V_{\text{in}}.$$

Aha – the voltage-divider equation, again! Thus, for changes in voltage, the circuit behaves like a voltage divider, with the zener replaced by a resistor equal to its dynamic resistance at the operating current. This is the

utility of incremental resistance. For instance, suppose in the preceding circuit we have an input voltage ranging between 15 and 20 V, and we use a 1N4733 (5.1 V, 1W zener diode) in order to generate a stable 5.1 V power supply. We choose  $R = 300\Omega$ , for a maximum zener current of 50 mA:  $(20\text{ V} - 5.1\text{ V})/300\Omega$ . We can now estimate the output-voltage regulation (variation in output voltage), knowing that this particular zener has a specified maximum dynamic resistance of  $7.0\Omega$  at 50 mA. The zener current varies from 50 mA to 33 mA over the input-voltage range; this 17 mA change in current then produces a voltage change at the output of  $\Delta V = R_{\text{dyn}}\Delta I$ , or 0.12 V.

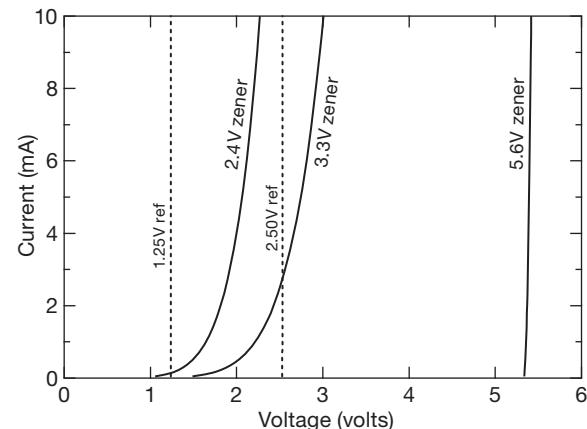
It’s a useful fact, when dealing with zener diodes, that the dynamic resistance of a zener diode varies roughly in inverse proportion to current. It’s worth knowing, also, that there are ICs designed to substitute for zener diodes; these “two-terminal voltage references” have superior performance – much lower dynamic resistance (less than  $1\Omega$ , even at currents as small as 0.1 mA; that’s a thousand times better than the zener we just used), and excellent temperature stability (better than  $0.01\text{/C}$ ). We will see more of zeners and voltage references in §§2.2.4 and 9.10.

In real life, a zener will provide better regulation if driven by a current source, which has, by definition,  $R_{\text{incr}}=\infty$  (the same current, regardless of voltage). But current sources are more complex, and therefore in practice we often resort to the humble resistor. When thinking about zeners, it’s worth remembering that low-voltage units (e.g., 3.3 V) behave rather poorly, in terms of constancy of voltage versus current (Figure 1.17); if you think you need a low voltage zener, use a two-terminal reference instead (§9.10).

### 1.2.7 An example: “It’s too hot!”

Some people like to turn the thermostat way up, annoying other people who like their houses cool. Here’s a little gadget (Figure 1.18) that lets folks of the latter persuasion know when to complain – it lights up a red light-emitting diode (LED) indicator when the room is warmer than  $30^\circ\text{C}$  ( $86^\circ\text{F}$ ). It also shows how to use the humble voltage divider (and even humbler Ohm’s law), and how to deal with an LED, which behaves like a zener diode (and is sometimes used as such).

The triangular symbol is a *comparator*, a handy device (discussed in §12.3) that switches its output according to the relative voltages at its two input terminals. The temperature sensing device is  $R_4$ , which decreases in resistance by about  $4\%/\text{C}$ , and which is  $10\text{k}\Omega$  at  $25^\circ\text{C}$ . So we’ve made



**Figure 1.17.** Low-voltage zeners are pretty dismal, as seen in these measured  $I$  vs.  $V$  curves (for three members of the 1N5221–67 series), particularly in contrast to the excellent measured performance of a pair of “IC voltage references” (LM385Z-1.2 and LM385Z-2.5, see §9.10 and Table 9.7). However, zener diodes in the neighborhood of 6 V (such as the 5.6 V 1N5232B or 6.2 V 1N5234B) exhibit admirably steep curves, and are useful parts.

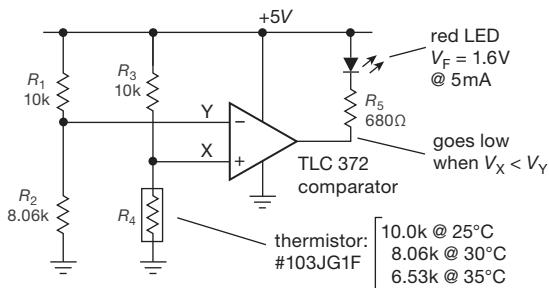
it the lower leg of a voltage divider ( $R_3R_4$ ), whose output is compared with the temperature-insensitive divider  $R_1R_2$ . When it’s hotter than  $30^\circ\text{C}$ , point “X” is at a lower voltage than point “Y”, so the comparator pulls its output to ground.

At the output there’s an LED, which behaves electrically like a 1.6 V zener diode; and when current is flowing, it lights up. Its lower terminal is then at  $5\text{ V} - 1.6\text{ V}$ , or  $+3.4\text{ V}$ . So we’ve added a series resistor, sized to allow 5 mA when the comparator output is at ground:  $R_5=3.4\text{ V}/5\text{ mA}$ , or  $680\Omega$ .

If you wanted to, you could make the setpoint adjustable by replacing  $R_2$  with a 5k pot in series with a 5k fixed resistor. We’ll see later that it’s also a good idea to add some *hysteresis*, to encourage the comparator to be decisive. Note that this circuit is insensitive to the exact power-supply voltage because it compares *ratios*. Ratiometric techniques are good; we’ll see them again later.

## 1.3 Signals

A later section in this chapter will deal with capacitors, devices whose properties depend on the way the voltages and currents in a circuit are *changing*. Our analysis of dc circuits so far (Ohm’s law, Thévenin equivalent circuits, etc.) still holds, even if the voltages and currents are changing in time. But for a proper understanding of alternating-current (ac) circuits, it is useful to have in mind certain common



**Figure 1.18.** The LED lights up when it's hotter than 30°C. The comparator (which we'll see later, in Chapters 4 and 12) pulls its output to ground when the voltage at "X" is less than the voltage at "Y."  $R_4$  is a thermistor, which is a resistor with a deliberate negative temperature coefficient; that is, its resistance decreases with increasing temperature – about 4%/°C.

types of *signals*, voltages that change in time in a particular way.

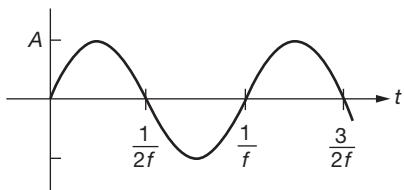
### 1.3.1 Sinusoidal signals

Sinusoidal signals are the most popular signals around; they're what you get out of the wall plug. If someone says something like "take a 10  $\mu$ V signal at 1 MHz," they mean a sinewave. Mathematically, what you have is a voltage described by

$$V = A \sin 2\pi ft \quad (1.10)$$

where  $A$  is called the amplitude and  $f$  is the frequency in hertz (cycles per second). A sinewave looks like the wave shown in Figure 1.19. Sometimes it is important to know the value of the signal at some arbitrary time  $t = 0$ , in which case you may see a *phase*  $\phi$  in the expression:

$$V = A \sin(2\pi ft + \phi).$$



**Figure 1.19.** Sinewave of amplitude  $A$  and frequency  $f$ .

The other variation on this simple theme is the use of *angular frequency*, which looks like this:

$$V = A \sin \omega t.$$

Here  $\omega$  is the angular frequency, measured in radians per

second. Just remember the important relation  $\omega = 2\pi f$  and you won't go wrong.

The great merit of sinewaves (and the cause of their perennial popularity) is the fact that they are the solutions to certain linear differential equations that happen to describe many phenomena in nature as well as the properties of linear circuits. A linear circuit has the property that its output, when driven by the sum of two input signals, equals the sum of its individual outputs when driven by each input signal in turn; i.e., if  $\mathcal{O}(A)$  represents the output when driven by signal  $A$ , then a circuit is linear if  $\mathcal{O}(A + B) = \mathcal{O}(A) + \mathcal{O}(B)$ . A linear circuit driven by a sinewave always responds with a sinewave, although in general the phase and amplitude are changed. No other periodic signal can make this statement. It is standard practice, in fact, to describe the behavior of a circuit by its *frequency response*, by which we mean the way the circuit alters the amplitude of an applied sinewave as a function of frequency. A stereo amplifier, for instance, should be characterized by a "flat" frequency response over the range 20 Hz to 20 kHz, at least.

The sinewave frequencies we usually deal with range from a few hertz to a few tens of megahertz. Lower frequencies, down to 0.0001 Hz or lower, can be generated with carefully built circuits, if needed. Higher frequencies, up to say 2000 MHz (2 GHz) and above, can be generated, but they require special transmission-line techniques. Above that, you're dealing with microwaves, for which conventional wired circuits with lumped-circuit elements become impractical, and exotic waveguides or "striplines" are used instead.

### 1.3.2 Signal amplitudes and decibels

In addition to its amplitude, there are several other ways to characterize the magnitude of a sinewave or any other signal. You sometimes see it specified by *peak-to-peak amplitude* (pp amplitude), which is just what you would guess, namely, twice the amplitude. The other method is to give the *root-mean-square amplitude* (rms amplitude), which is  $V_{\text{rms}} = (1/\sqrt{2})A = 0.707A$  (this is for sinewaves only; the ratio of pp to rms will be different for other waveforms). Odd as it may seem, this is the usual method, because rms voltage is what's used to compute power. The nominal voltage across the terminals of a wall socket (in the United States) is 120 volts rms, 60 Hz. The *amplitude* is 170 volts (339 volts pp).<sup>18</sup>

<sup>18</sup> Occasionally you'll encounter devices (e.g., mechanical moving-pointer meters) that respond to the *average* magnitude of an ac signal.

### A. Decibels

How do you compare the relative amplitudes of two signals? You could say, for instance, that signal  $X$  is twice as large as signal  $Y$ . That's fine, and useful for many purposes. But because we often deal with ratios as large as a million, it is better to use a logarithmic measure, and for this we present the decibel (it's one-tenth as large as something called a bel, which no one ever uses). By definition, the ratio of two signals, in decibels (dB), is

$$\text{dB} = 10 \log_{10} \frac{P_2}{P_1}, \quad (1.11)$$

where  $P_1$  and  $P_2$  represent the *power* in the two signals. We are often dealing with signal *amplitudes*, however, in which case we can express the ratio of two signals having the same waveform as

$$\text{dB} = 20 \log_{10} \frac{A_2}{A_1}, \quad (1.12)$$

where  $A_1$  and  $A_2$  are the two signal amplitudes. So, for instance, one signal of twice the amplitude of another is +6 dB relative to it, since  $\log_{10} 2 = 0.3010$ . A signal 10 times as large is +20 dB; a signal one-tenth as large is -20 dB.

Although decibels are ordinarily used to specify the ratio of two signals, they are sometimes used as an absolute measure of amplitude. What is happening is that you are assuming some reference signal level and expressing any other level in decibels relative to it. There are several standard levels (which are unstated, but understood) that are used in this way; the most common references are (a) 0 dBV (1 V rms); (b) 0 dBm (the voltage corresponding to 1 mW into some assumed load impedance, which for radiofrequencies is usually  $50\Omega$ , but for audio is often  $600\Omega$ ; the corresponding 0 dBm amplitudes, when loaded by those impedances, are then 0.22 V rms and 0.78 V rms); and (c) the small noise voltage generated by a resistor at room temperature (this surprising fact is discussed in §8.1.1). In addition to these, there are reference amplitudes used for measurements in other fields of engineering and science. For instance, in acoustics, 0 dB SPL (sound pressure level) is a wave whose rms pressure is  $20\mu\text{Pa}$  (that's  $2 \times 10^{-10}$  atm); in audio communications, levels can be stated in dBnC (relative noise reference weighted in frequency by "curve C"). When stating amplitudes this way, it is best to be specific about the 0 dB reference amplitude;

---

For a sinewave the relationship is  $V_{\text{avg}}=V_{\text{rms}}/1.11$ . However, such meters are usually calibrated so that they indicate the rms sinewave amplitude. For signals other than sinewaves their indication is in error; be sure to use a "true rms" meter if you want the right answer.

say something like "an amplitude of 27 decibels relative to 1 V rms," or abbreviate "27 dB re 1 V rms," or define a term like "dBV."<sup>19</sup>

**Exercise 1.12.** Determine the voltage and power ratios for a pair of signals with the following decibel ratios: (a) 3 dB, (b) 6 dB, (c) 10 dB, (d) 20 dB.

**Exercise 1.13.** We might call this amusing exercise "Desert Island dBs": in the table below we've started entering some values for power ratios corresponding to the first dozen integral dBs, using the results for parts (a) and (c) of the last exercise. Your job is to complete the table, without recourse to a calculator. A possibly helpful hint: starting at 10 dB, go down the table in steps of 3 dB, then up in a step of 10 dB, then down again. Finally, get rid of yucky numbers like 3.125 (and its near relatives) by noticing that it's charmingly close to  $\pi$ .

dB	ratio ( $P/P_0$ )
0	1
1	
2	
3	2
4	
5	
6	4
7	
8	
9	8
10	10
11	

### 1.3.3 Other signals

#### A. Ramp

The ramp is a signal that looks like the one shown in Figure 1.20A. It is simply a voltage rising (or falling) at a constant rate. That can't go on forever, of course, even in science fiction movies. It is sometimes approximated by a finite ramp (Figure 1.20B) or by a periodic ramp (known as a *sawtooth*, Figure 1.20C).

#### B. Triangle

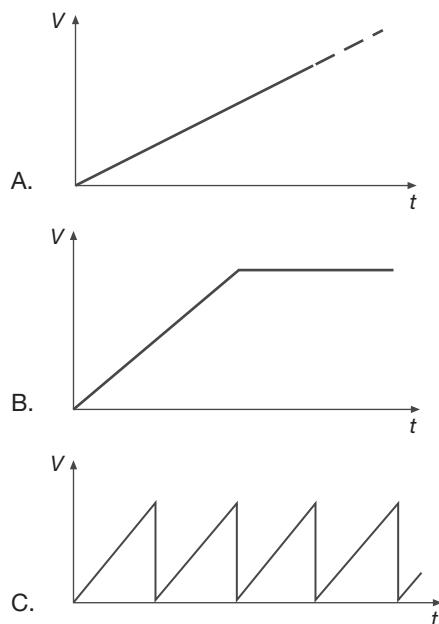
The triangle wave is a close cousin of the ramp; it is simply a symmetrical ramp (Figure 1.21).

#### C. Noise

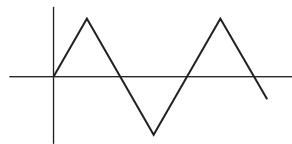
Signals of interest are often mixed with *noise*; this is a catch-all phrase that usually applies to random noise of thermal origin. Noise voltages can be specified by their

---

<sup>19</sup> One of the authors, when asked by his nontechnical spouse how much we spent on that big plasma screen, replied "36 dB\$."



**Figure 1.20.** A: Voltage-ramp waveform. B: Ramp with limit. C: Sawtooth wave.



**Figure 1.21.** Triangle wave.

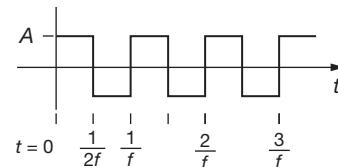


**Figure 1.22.** Noise.

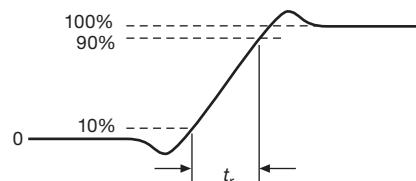
frequency spectrum (power per hertz) or by their amplitude distribution. One of the most common kind of noise is *band-limited white Gaussian noise*, which means a signal with equal power per hertz in some band of frequencies and that exhibits a Gaussian (bell-shaped) distribution of amplitudes when many instantaneous measurements of its amplitude are made. This kind of noise is generated by a resistor (Johnson noise or Nyquist noise), and it plagues sensitive measurements of all kinds. On an oscilloscope it appears as shown in Figure 1.22. We will discuss noise and low-noise techniques in considerable detail in Chapter 8.

#### D. Square wave

A square wave is a signal that varies in time as shown in Figure 1.23. Like the sinewave, it is characterized by amplitude and frequency (and perhaps phase). A linear circuit driven by a square wave rarely responds with a square wave. For a square wave, the peak amplitude and the rms amplitude are the same.

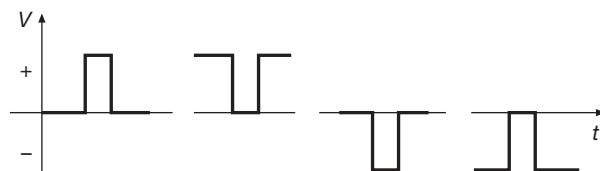


**Figure 1.23.** Square wave.



**Figure 1.24.** Rise time of a step waveform.

The edges of a square wave are not perfectly square; in typical electronic circuits the *rise time*  $t_r$  ranges from a few nanoseconds to a few microseconds. Figure 1.24 shows the sort of thing usually seen. The rise time is conventionally defined as the time required for the signal to go from 10% to 90% of its total transition.



**Figure 1.25.** Positive- and negative-going pulses of both polarities.

#### E. Pulses

A pulse is a signal that looks like the objects shown in Figure 1.25. It is defined by amplitude and pulse width. You can generate a train of periodic (equally spaced) pulses, in which case you can talk about the frequency, or pulse repetition rate, and the “duty cycle,” the ratio of pulse width to repetition period (duty cycle ranges from zero to 100%). Pulses can have positive or negative polarity; in addition, they can be “positive-going” or “negative-going.” For

instance, the second pulse in Figure 1.25 is a negative-going pulse of positive polarity.

### F. Steps and spikes

Steps and spikes are signals that are talked about a lot but are not so often used. They provide a nice way of describing what happens in a circuit. If you could draw them, they would look something like the example in Figure 1.26. The step function is part of a square wave; the spike is simply a jump of vanishingly short duration.



Figure 1.26. Steps and spikes.

### 1.3.4 Logic levels

Pulses and square waves are used extensively in digital electronics, in which predefined voltage levels represent one of two possible states present at any point in the circuit. These states are called simply HIGH and LOW, and correspond to the 1 (true) and 0 (false) states of Boolean logic (the algebra that describes such two-state systems).

Precise voltages are not necessary in digital electronics. You need only to distinguish which of the two possible states is present. Each digital logic family therefore specifies legal HIGH and LOW states. For example, the "74LVC" digital logic family runs from a single +3.3 V supply, with output levels that are typically 0 V (LOW) and 3.3 V (HIGH), and an input decision threshold of 1.5 V. However, actual outputs can be as much as 0.4 V away from ground or from +3.3 V without malfunction. We'll have much more to say about logic levels in Chapters 10 through 12.

### 1.3.5 Signal sources

Often the source of a signal is some part of the circuit you are working on. But for test purposes a flexible signal source is invaluable. They come in three flavors: signal generators, pulse generators, and function generators.

#### A. Signal generators

Signal generators are sinewave oscillators, usually equipped to give a wide range of frequency coverage,

with provision for precise control of amplitude (using a resistive divider network called an *attenuator*). Some units let you *modulate* (i.e., vary in time) the output amplitude ("AM" for "amplitude modulated") or frequency ("FM" for "frequency modulated"). A variation on this theme is the *sweep generator*, a signal generator that can sweep its output frequency repeatedly over some range. These are handy for testing circuits whose properties vary with frequency in a particular way, e.g., "tuned circuits" or filters. Nowadays these devices, as well as most test instruments, are available in configurations that allow you to program the frequency, amplitude, etc., from a computer or other digital instrument.

For many signal generators the signal source is a *frequency synthesizer*, a device that generates sinewaves whose frequencies can be set precisely. The frequency is set digitally, often to eight significant figures or more, and is internally synthesized from a precise standard (a stand-alone quartz-crystal oscillator or rubidium frequency standard, or a GPS-derived oscillator) by digital methods we will discuss later (§13.13.6). Typical of synthesizers is the programmable SG384 from Stanford Research Systems, with a frequency range of 1  $\mu$ Hz to 4 GHz, an amplitude range of -110 dBm to +16.5 dBm (0.7  $\mu$ V to 1.5 V, rms), and various modulation modes such as AM, FM, and  $\Phi$ M; it costs about \$4,600. You can get synthesized sweep generators, and you can get synthesizers that produce other waveforms (see *Function Generators*, below). If your requirement is for no-nonsense accurate frequency generation, you can't beat a synthesizer.

#### B. Pulse generators

Pulse generators make only pulses, but what pulses! Pulse width, repetition rate, amplitude, polarity, rise time, etc., may all be adjustable. The fastest ones go up to gigahertz pulse rates. In addition, many units allow you to generate pulse pairs, with settable spacing and repetition rate, or even programmable patterns (they are sometimes called pattern generators). Most contemporary pulse generators are provided with logic-level outputs for easy connection to digital circuitry. As with signal generators, these come in the programmable variety.

#### C. Function generators

In many ways function generators are the most flexible signal sources of all. You can make sine, triangle, and square waves over an enormous frequency range (0.01 Hz to 30 MHz is typical), with control of amplitude and dc offset (a constant-dc voltage added to the signal). Many of them have provision for frequency sweeping, often in

several modes (linear or logarithmic frequency variation versus time). They are available with pulse outputs (although not with the flexibility you get with a pulse generator), and some of them have provision for modulation.

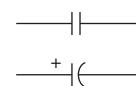
Traditional function generators used analog circuitry, but contemporary models generally are synthesized digital function generators, exhibiting all the flexibility of a function generator along with the stability and accuracy of a frequency synthesizer. In addition, they let you program an “arbitrary” waveform, specifying the amplitude at a set of equally spaced points. An example is the Tektronix AFG3102, with a lower frequency limit of 1 microhertz, which can make sine and square waves to 100 MHz, pulses and “noise” to 50 MHz, and arbitrary waveforms (up to 128k points) to 50 MHz. It has modulation (five kinds), sweep (linear and log), and burst modes (1 to  $10^6$  cycles), and everything is programmable, including frequency, pulse width and rise times, modulation, and amplitude (20 mV to 10 Vpp); it even includes some bizarre built-in waveforms such as  $\sin(x)/x$ , exponential rise and fall, Gaussian, and Lorentzian. It has two independent outputs and costs about \$5k. For general use, if you can have only one signal source, the function generator is for you.

## 1.4 Capacitors and ac circuits

Once we enter the world of *changing* voltages and currents, or “signals,” we encounter two very interesting circuit elements that are useless in purely dc circuits: capacitors and inductors. As you will see, these humble devices, combined with resistors, complete the triad of passive linear circuit elements that form the basis of nearly all circuitry.<sup>20</sup> Capacitors, in particular, are essential in nearly every circuit application. They are used for waveform generation, filtering, and blocking and bypass applications. They are used in integrators and differentiators. In combination with inductors, they make possible sharp filters for separating desired signals from background. You will see some of these applications as we continue with this chapter, and there will be numerous interesting examples in later chapters.

Let’s proceed, then, to look at capacitors in detail. Por-

<sup>20</sup> Readers of the scientific journal *Nature* (London) were greeted, in 2008, with an article titled “The missing memristor found” (D. B. Strukov et al., **453**, 80, 2008), purporting to have found a heretofore missing “fourth fundamental [passive circuit] element.” We are skeptical. However the controversy is ultimately resolved, it should be noted that the memristor is a *nonlinear* element; there are only three *linear* passive 2-terminal circuit elements.



**Figure 1.27.** Capacitors. The curved electrode indicates the negative terminal of a polarized capacitor, or the “outer foil” of a wrapped-film capacitor.

tions of the treatment that follows are necessarily mathematical in nature; the reader with little mathematical preparation may find the math review in Appendix A helpful. In any case, an understanding of the details is less important in the long run than an understanding of the results.

### 1.4.1 Capacitors

A capacitor (Figure 1.27) (the old-fashioned name was *condenser*) is a device that has two wires sticking out of it and has the property

$$Q = CV. \quad (1.13)$$

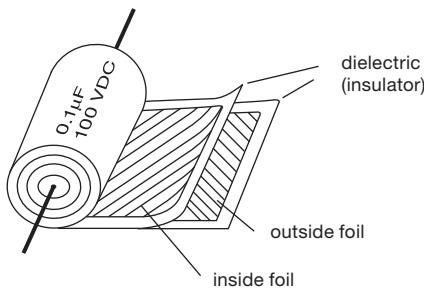
Its basic form is a pair of closely-spaced metal plates, separated by some insulating material, as in the rolled-up “axial-film capacitor” of Figure 1.28. A capacitor of  $C$  farads with  $V$  volts across its terminals has  $Q$  coulombs of stored charge on one plate and  $-Q$  on the other. The capacitance is proportional to the area and inversely proportional to the spacing. For the simple parallel-plate capacitor, with separation  $d$  and plate area  $A$  (and with the spacing  $d$  much less than the dimensions of the plates), the capacitance  $C$  is given by

$$C = 8.85 \times 10^{-14} \epsilon A/d \text{ F}, \quad (1.14)$$

where  $\epsilon$  is the dielectric constant of the insulator, and the dimensions are measured in centimeters. It takes a lot of area, and tiny spacing, to make the sort of capacitances commonly used in circuits.<sup>21</sup> For example, a pair of  $1 \text{ cm}^2$  plates separated by 1 mm is a capacitor of slightly less than  $10^{-12} \text{ F}$  (a picofarad); you’d need 100,000 of them just to create the  $0.1 \mu\text{F}$  capacitor of Figure 1.28 (which is nothing special; we routinely use capacitors with many microfarads of capacitance). Ordinarily you don’t need to calculate capacitances, because you buy a capacitor as an electronic component.

To a first approximation, capacitors are devices that might be considered simply frequency-dependent resistors.

<sup>21</sup> And it doesn’t hurt to have a high dielectric constant, as well: air has  $\epsilon=1$ , but plastic films have  $\epsilon=2.1$  (polypropylene) or 3.1 (polyester). And certain ceramics are popular among capacitor makers:  $\epsilon=45$  (“C0G” type) or 3000 (“X7R” type).



**Figure 1.28.** You get a lot of area by rolling up a pair of metallized plastic films. And it's great fun unrolling one of these axial-lead Mylar capacitors (ditto for the old-style golf balls with their lengthy wound-up rubber band).

They allow you to make frequency-dependent voltage dividers, for instance. For some applications (bypass, coupling) this is almost all you need to know, but for other applications (filtering, energy storage, resonant circuits) a deeper understanding is needed. For example, ideal capacitors cannot dissipate power, even though current can flow through them, because the voltage and current are  $90^\circ$  out of phase.

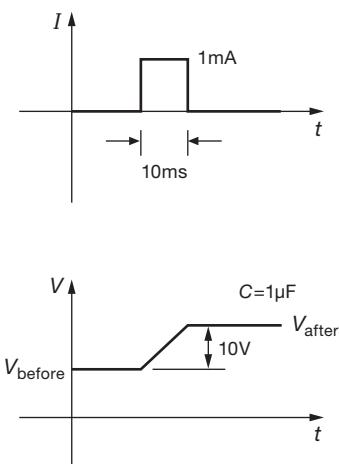
Before launching into the details of capacitors in the following dozen pages (including some necessary mathematics that describes their behavior in time and in frequency), we wish to emphasize those first two applications – bypass and coupling – because they are the most common uses of capacitors, and they are easy to understand at the simplest level. We'll see these in detail later (§§1.7.1C and 1.7.1A), but no need to wait – it's easy, and intuitive. Because a capacitor looks like an open circuit at dc, it lets you couple a varying signal while blocking its average dc level. This is a *blocking* capacitor (also called a *coupling* capacitor), as in Figure 1.93. Likewise, because a capacitor looks like a short circuit at high frequencies, it suppresses ("bypasses") signals where you don't want them, for example on the dc voltages that power your circuits, as in Figure 8.80A (where capacitors are suppressing signals on the +5 V and -5 V dc supply voltages, and also on the base terminal of transistor  $Q_2$ ).<sup>22</sup> Demographically, these two applications account for the vast majority of capacitors that are wired into the world's circuits.

Taking the derivative of the defining equation 1.13, you get

<sup>22</sup> Ironically, these essential bypass capacitors are so taken for granted that they are usually omitted from schematic diagrams (a practice we follow in this book). Don't make the mistake of omitting them also from your actual circuits!

$$I = C \frac{dV}{dt}. \quad (1.15)$$

So a capacitor is more complicated than a resistor: the current is not simply proportional to the voltage, but rather to the rate of change of voltage. If you change the voltage across a farad by 1 volt per second, you are supplying an amp. Conversely, if you supply an amp, its voltage changes by 1 volt per second. A farad is an enormous capacitance, and you usually deal in microfarads ( $\mu\text{F}$ ), nanofarads (nF), or picofarads (pF).<sup>23</sup> For instance, if you supply a current of 1 mA to  $1\mu\text{F}$ , the voltage will rise at 1000 volts per second. A 10 ms pulse of this current will increase the voltage across the capacitor by 10 volts (Figure 1.29).



**Figure 1.29.** The voltage across a capacitor changes when a current flows through it.

When you charge up a capacitor, you're supplying energy. The capacitor doesn't get hot; instead, it stores the energy in its internal electric fields. It's an easy exercise to discover for yourself that the amount of stored energy in a charged capacitor is just

$$U_C = \frac{1}{2} CV^2, \quad (1.16)$$

where  $U_C$  is in joules for  $C$  in farads and  $V$  in volts. This is an important result; we'll see it often.

**Exercise 1.14.** Take the energy challenge: imagine charging up a capacitor of capacitance  $C$ , from 0 V to some final voltage  $V_f$ . If you do it right, the result won't depend on how you get there,

<sup>23</sup> To make matters confusing to the uninitiated, the units are often omitted on capacitor values specified in schematic diagrams. You have to figure it out from the context.



**Figure 1.30.** Capacitors masquerade as anything they like! Here is a representative collection. In the lower left are small-value variable capacitors (one air, three ceramic), with large-value polarized aluminum electrolytics above them (the three on the left have *radial* leads, the three on the right have *axial* leads, and the specimen with screw terminals at top is often called a *computer electrolytic*). Next in line across the top is a low-inductance film capacitor (note the wide strap terminals), then an oil-filled paper capacitor, and last, a set of disc ceramic capacitors running down the right. The four rectangular objects below are film capacitors (polyester, polycarbonate, or polypropylene). The D-subminiature connector seems misplaced – but it is a *filtered* connector, with a 1000 pF capacitor from each pin to the shell. To its left is a group of seven polarized tantalum electrolytics (five with axial leads, one radial, and one surface-mount). The three capacitors above them are axial-film capacitors. The ten capacitors at bottom center are all ceramic types (four with radial leads, two axial, and four surface-mount *chip capacitors*); above them are high-voltage capacitors – an axial-glass capacitor, and a ceramic *transmitting capacitor* with screw terminals. Finally, below them and to the left are four mica capacitors and a pair of diode-like objects known as *varactors*, which are voltage-variable capacitors made from a diode junction.

so you don't need to assume constant current charging (though you're welcome to do so). At any instant the rate of flow of energy into the capacitor is  $VI$  (joules/s); so you need to integrate  $dU = VI dt$  from start to finish. Take it from there.

Capacitors come in an amazing variety of shapes and sizes (Figure 1.30 shows examples of most of them); with time, you will come to recognize their more common incarnations. For the smallest capacitances you may see examples of the basic parallel-plate (or cylindrical piston) construction. For greater capacitance, you need more area and

closer spacing; the usual approach is to plate some conductor onto a thin insulating material (the dielectric), for instance, aluminized plastic film rolled up into a small cylindrical configuration. Other popular types are thin ceramic wafers (ceramic chip capacitors), metal foils with oxide insulators (electrolytic capacitors), and metallized mica. Each of these types has unique properties; for a brief rundown, see the section on capacitors in Chapter 1x. In general, ceramic and polyester types are used for most non-critical circuit applications; capacitors with polycarbonate, polystyrene, polypropylene, Teflon, or glass dielectric are

used in demanding applications; tantalum capacitors are used where greater capacitance is needed; and aluminum electrolytics are used for power-supply filtering.

### A. Capacitors in parallel and series

The capacitance of several capacitors in parallel is the sum of their individual capacitances. This is easy to see: put voltage  $V$  across the parallel combination; then

$$\begin{aligned} C_{\text{total}}V &= Q_{\text{total}} = Q_1 + Q_2 + Q_3 + \dots \\ &= C_1V + C_2V + C_3V + \dots \\ &= (C_1 + C_2 + C_3 + \dots)V \end{aligned}$$

or

$$C_{\text{total}} = C_1 + C_2 + C_3 + \dots \quad (1.17)$$

For capacitors in series, the formula is like that for resistors in parallel:

$$C_{\text{total}} = \frac{1}{\frac{1}{C_1} + \frac{1}{C_2} + \frac{1}{C_3} + \dots} \quad (1.18)$$

or (two capacitors only)

$$C_{\text{total}} = \frac{C_1 C_2}{C_1 + C_2}.$$

**Exercise 1.15.** Derive the formula for the capacitance of two capacitors in series. Hint: because there is no external connection to the point where the two capacitors are connected together, they must have equal stored charges.

The current that flows in a capacitor during charging ( $I = CdV/dt$ ) has some unusual features. Unlike resistive current, it's not proportional to voltage, but rather to the rate of change (the “time derivative”) of voltage. Furthermore, unlike the situation in a resistor, the power ( $V \times I$ ) associated with capacitive current is not turned into heat, but is stored as energy in the capacitor’s internal electric field. You get all that energy back when you discharge the capacitor. We’ll see another way to look at these curious properties when we talk about *reactance*, beginning in §1.7.

### 1.4.2 RC circuits: V and I versus time

When dealing with ac circuits (or, in general, any circuits that have changing voltages and currents), there are two possible approaches. You can talk about  $V$  and  $I$  versus time, or you can talk about amplitude versus signal frequency. Both approaches have their merits, and you find yourself switching back and forth according to which description is most convenient in each situation. We begin our

study of ac circuits in the *time domain*. Starting with §1.7, we will tackle the *frequency domain*.

What are some of the features of circuits with capacitors? To answer this question, let’s begin with the simple *RC* circuit (Figure 1.31). Application of the capacitor rules gives

$$C \frac{dV}{dt} = I = -\frac{V}{R}. \quad (1.19)$$

This is a differential equation, and its solution is

$$V = Ae^{-t/RC}. \quad (1.20)$$

So a charged capacitor placed across a resistor will discharge as in Figure 1.32. Intuition serves well here: the current that flows is (from the resistor equation) proportional to the remaining voltage; but the slope of the discharge is (from the capacitor equation) proportional to that current. So the discharge curve has to be a function whose derivative is proportional to its value, i.e., an exponential.

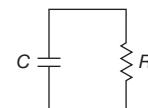


Figure 1.31. The simplest *RC* circuit.

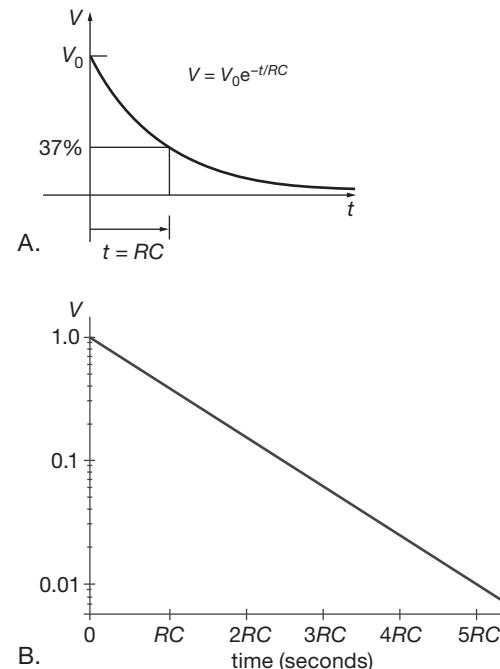


Figure 1.32. *RC* discharge waveform, plotted with (A) linear and (B) logarithmic voltage axes.

### A. Time constant

The product  $RC$  is called the *time constant* of the circuit. For  $R$  in ohms and  $C$  in farads, the product  $RC$  is in seconds. A microfarad across 1.0k has a time constant of 1 ms; if the capacitor is initially charged to 1.0 V, the initial current is 1.0 mA.

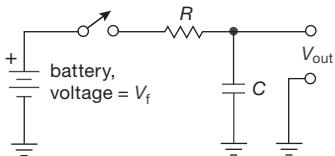


Figure 1.33. *RC* charging circuit.

Figure 1.33 shows a slightly different circuit. At time  $t = 0$ , someone connects the battery. The equation for the circuit is then

$$I = C \frac{dV}{dt} = \frac{V_f - V_{\text{out}}}{R},$$

with the solution

$$V_{\text{out}} = V_f + A e^{-t/RC}.$$

(Please don't worry if you can't follow the mathematics. What we are doing is getting some important results, which you should remember. Later we will use the results often, with no further need for the mathematics used to derive them. For readers whose knowledge of math is somewhat, uh, *rusty*, the brief review in Appendix A may prove helpful.) The constant  $A$  is determined by initial conditions (Figure 1.34):  $V = 0$  at  $t = 0$ ; therefore,  $A = -V_f$ , and

$$V_{\text{out}} = V_f (1 - e^{-t/RC}). \quad (1.21)$$

Once again there's good intuition: as the capacitor charges up, the slope (which is proportional to current, because it's a capacitor) is proportional to the *remaining* voltage (because that's what appears across the resistor, producing the current); so we have a waveform whose slope decreases proportionally to the vertical distance it has still to go – an exponential.

You can turn around the last equation to figure out the time required to reach a voltage  $V$  on the way to the final voltage  $V_f$ . Try it! (Refer to Appendix A if you need help with logarithms.) You should get

$$t = RC \log_e \left( \frac{V_f}{V_f - V} \right) \quad (1.22)$$

### B. Decay to equilibrium

Eventually (when  $t \gg RC$ ),  $V$  reaches  $V_f$ . (Presenting the "5RC rule of thumb": a capacitor charges or decays to

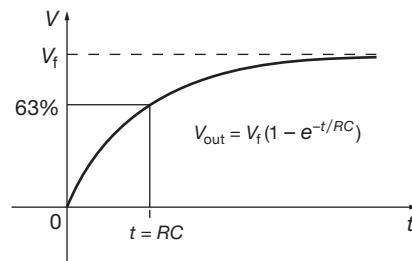


Figure 1.34. *RC* charging waveform.

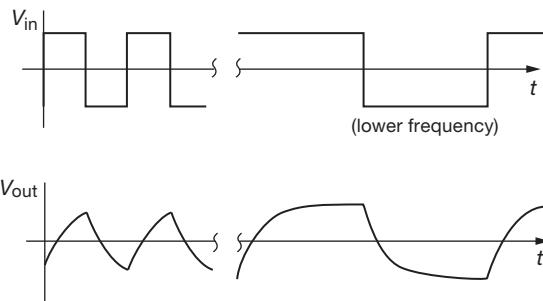


Figure 1.35. Output (lower waveforms) across a capacitor, when driven by square waves through a resistor.

within 1% of its final value in five time constants.) If we then change the battery voltage to some other value (say, 0 V),  $V$  will decay toward that new value with an exponential  $e^{-t/RC}$ . For example, replacing the battery's step input from 0 to  $+V_f$  with a square-wave input  $V_{\text{in}}(t)$  would produce the output shown in Figure 1.35.

**Exercise 1.16.** Show that the rise time (the time required for going from 10% to 90% of its final value) of this signal is  $2.2RC$ .

You might ask the obvious next question: what about  $V(t)$  for arbitrary  $V_{\text{in}}(t)$ ? The solution involves an inhomogeneous differential equation and can be solved by standard methods (which are, however, beyond the scope of this book). You would find

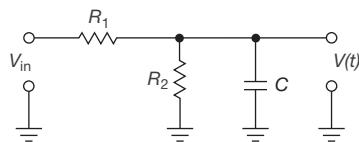
$$V(t) = \frac{1}{RC} \int_{-\infty}^t V_{\text{in}}(\tau) e^{-(t-\tau)/RC} d\tau.$$

That is, the *RC* circuit averages past history at the input with a weighting factor of

$$e^{-\Delta t/RC}.$$

In practice, you seldom ask this question. Instead, you deal in the *frequency domain*, in which you ask how much of each frequency component present in the input gets through. We will get to this important topic soon (§1.7). Before we do, though, there are a few other interesting

circuits we can analyze simply with this time-domain approach.

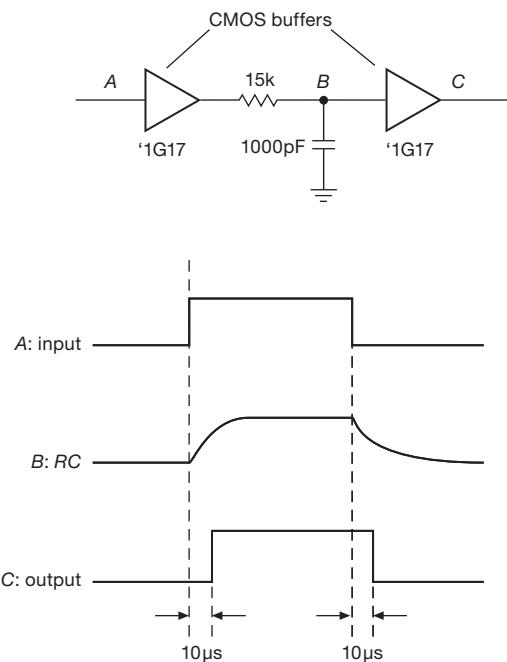


**Figure 1.36.** Looks complicated, but it's not! (Thévenin to the rescue.)

### C. Simplification by Thévenin equivalents

We could go ahead and analyze more complicated circuits by similar methods, writing down the differential equations and trying to find solutions. For most purposes it simply isn't worth it. This is as complicated an *RC* circuit as we will need. Many other circuits can be reduced to it; take, for example, the circuit in Figure 1.36. By just using the Thévenin equivalent of the voltage divider formed by  $R_1$  and  $R_2$ , you can find the output  $V(t)$  produced by a step input for  $V_{in}$ .

**Exercise 1.17.** In the circuit shown in Figure 1.36,  $R_1 = R_2 = 10\text{k}$ , and  $C = 0.1 \mu\text{F}$ . Find  $V(t)$  and sketch it.



**Figure 1.37.** Producing a delayed digital waveform with the help of an *RC* and a pair of LVC-family logic buffers (tiny parts with a huge part number: SN74LVC1G17DCKR!).

### D. A circuit example: time-delay circuit

Let's take a short detour to try out these theoretical ideas on a couple of real circuits. Textbooks usually avoid such pragmatism, especially in early chapters, but we think it's fun to apply electronics to practical applications. We'll need to introduce a few "black-box" components to get the job done, but you'll learn about them in detail later, so don't worry.

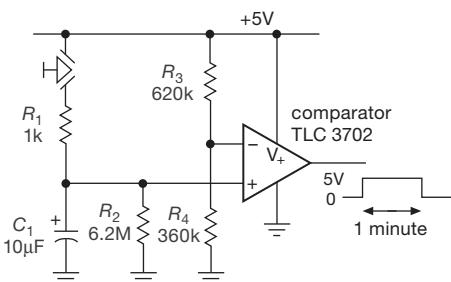
We have already mentioned logic levels, the voltages that digital circuits live on. Figure 1.37 shows an application of capacitors to produce a delayed pulse. The triangular symbols are "CMOS<sup>24</sup> buffers." They give a HIGH output if the input is HIGH (more than one-half the dc power-supply voltage used to power them), and vice versa. The first buffer provides a replica of the input signal, but with low source resistance, to prevent input loading by the *RC* (recall our earlier discussion of circuit loading in §1.2.5A). The *RC* output has the characteristic decays and causes the output buffer to switch 10  $\mu\text{s}$  after the input transitions (an *RC* reaches 50% output after a time  $t = 0.7RC$ ). In an actual application you would have to consider the effect of the buffer input threshold deviating from one-half the supply voltage, which would alter the delay and change the output pulse width. Such a circuit is sometimes used to delay a pulse so that something else can happen first. In designing circuits you try not too often to rely on tricks like this, but they're occasionally handy.

### E. Another circuit example: "One Minute of Power"

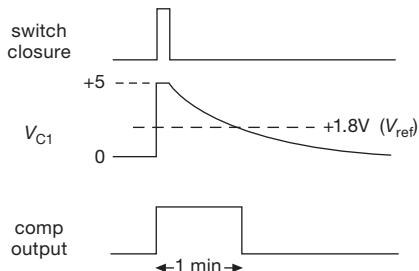
Figure 1.38 shows another example of what can be done with simple *RC* timing circuits. The triangular symbol is a *comparator*, something we'll treat in detail later, in Chapters 4 and 10; all you need to know, for now, is that (a) it is an IC (containing a bunch of resistors and transistors), (b) it is powered from some positive dc voltage that you connect to the pin labeled " $V_+$ ", and (c) it drives its output (the wire sticking out to the right) either to  $V_+$  or to ground, depending on whether the input labeled "+" is more or less positive than the input labeled "-", respectively. (These are called the *non-inverting* and the *inverting* inputs, respectively.) It doesn't draw any current from its inputs, but it happily drives loads that require up to 20 mA or so. And a comparator is decisive: its output is either "HIGH" (at  $V_+$ ) or "LOW" (ground).

Here's how the circuit works: the voltage divider  $R_3R_4$  holds the (-) input at 37% of the supply voltage, in this case about +1.8 V; let's call that the "reference voltage."

<sup>24</sup> Complementary metal-oxide semiconductor, the dominant form of digital logic, as we'll see from Chapter 10 onward.



**Figure 1.38.** RC timing circuit: one push → one minute!.

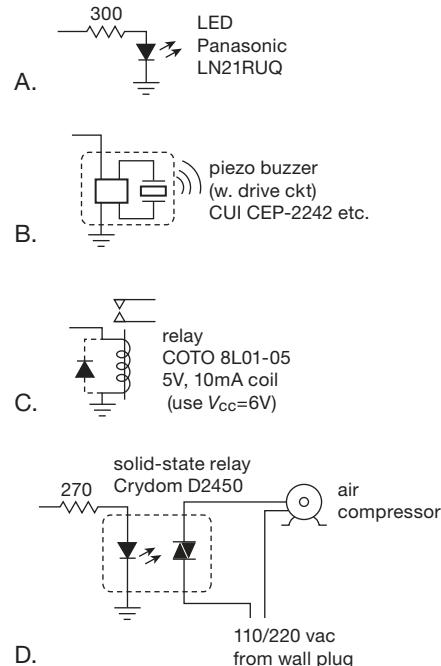


**Figure 1.39.** Producing a delayed digital waveform for the circuit of Figure 1.38. The voltage  $V_{C1}$  has a rise time of  $R_1 C_1 \approx 10 \text{ ms}$ .

So if the circuit has been sitting there for a while,  $C_1$  is fully discharged, and the comparator's output is at ground. When you push the START button momentarily,  $C_1$  charges quickly (10 ms time constant) to +5 V, which makes the comparator's output switch to +5 V; see Figure 1.39. After the button is released, the capacitor discharges exponentially toward ground, with a time constant of  $\tau = R_2 C_1$ , which we've set to be 1 minute. At that time its voltage crosses the reference voltage, so the comparator's output switches rapidly back to ground. (Note that we've conveniently chosen the reference voltage to be a fraction  $1/e$  of  $V_+$ , so it takes exactly one time constant  $\tau$  for that to happen. For  $R_2$  we used the closest standard value to  $6 \text{ M}\Omega$ ; see Appendix C.) The bottom line is that the output spends 1 minute at +5 V, after the button is pushed.

We'll add a few details shortly, but first let's use the output to do some interesting things, which are shown in Figures 1.40A–D. You can make a self-stopping flashlight key fob by connecting its output to an LED; you need to put a resistor in series, to set the current (we'll say much more about this later). If you prefer to make some noise, you could connect a *piezo beeper* to beep continuously (or intermittently) for a minute (this might be an end-of-cycle signal for a clothes dryer). Another possibility is to attach a small electromechanical *relay*, which is just an electrically operated mechanical switch, to provide a pair of contacts

that can activate pretty much any load you care to switch on and off. The use of a relay has the important property that the load – the circuit being switched by the relay – is electrically isolated from the +5 V and ground of the timing circuit itself.



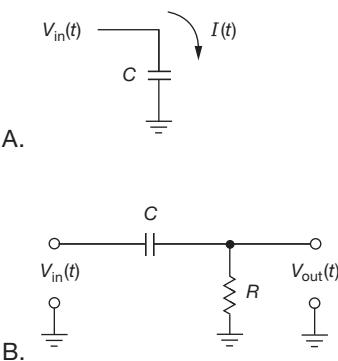
**Figure 1.40.** Driving interesting stuff from the output of the timer circuit in Figure 1.38.

Finally, for turning serious industrial machinery on and off, you would probably use a hefty *solid-state relay* (SSR, §12.7), which has within it an infrared LED coupled to an ac switching device known as a *triac*. When activated, the triac acts as an excellent mechanical switch, capable of switching many amperes, and (like the electromechanical relay) is fully isolated electrically from its input circuit. The example shows this thing hooked to an air compressor, so your friends will get a minute's worth of air to inflate their tires at your home "gas station" (literally!) after they drop a quarter into your coin-initiated timer. You could do an analogous thing with a coin-operated hot shower (but, hey, we get only *one minute*?!).

Some details: (a) in the circuit of Figure 1.38 you could omit  $R_1$  and the circuit would still work, but there would be a large transient current when the discharged capacitor was initially connected across the +5 V supply (recall  $I = C dV/dt$ : here you would be trying to produce 5 V of " $dV$ " in roughly 0 s of " $dt$ "). By adding a series resistor  $R_1$  you limit the peak current to a modest 5 mA while charging

the capacitor fast enough ( $> 99\%$  in  $5 RC$  time constants, or 0.05 s). (b) The comparator output would likely bounce around a bit (see Figure 4.31), just as the (+) input crosses the reference voltage in its leisurely exponential promenade toward ground, owing to unavoidable bits of electrical noise. To fix this problem you usually see the circuit arranged so that some of the output is coupled back to the input in a way that reinforces the switching (this is officially called *hysteresis*, or *positive feedback*; we'll see it in Chapters 4 and 10). (c) In electronic circuits it's always a good idea to *bypass* the dc supply by connecting one or more capacitors between the dc "rail" and ground. The capacitance is noncritical – values of  $0.1 \mu\text{F}$  to  $10 \mu\text{F}$  are commonly used; see §1.7.16A.

Our simple examples above all involved turning some load on and off. But there are other uses for an electronic *logic* signal, like the output of the comparator, that is in one of two possible binary states, called HIGH and LOW (in this case +5 V and ground), 1 and 0, or TRUE and FALSE. For example, such a signal can enable or disable the operation of some other circuit. Imagine that the opening of a car door triggers our 1 min HIGH output, which then enables a keypad to accept a security code so you can start the car. After a minute, if you haven't managed to type the magic code, it shuts off, thus ensuring a certain minimum of operator sobriety.



**Figure 1.41.** Differentiators. A. Perfect (except it has no output terminal). B. Approximate (but at least it has an output!).

### 1.4.3 Differentiators

You can make a simple circuit that differentiates an input signal; that is,  $V_{out} \propto dV_{in}/dt$ . Let's take it in two steps.

1. First look at the (impractical) circuit in Figure 1.41A: The input voltage  $V_{in}(t)$  produces a current through the capacitor of  $I_{cap} = C dV_{in}/dt$ . That's just what we want – if we

could somehow use the current through  $C$  as our "output"! But we can't.<sup>25</sup>

2. So we add a small resistor from the low side of the capacitor to ground, to act as a "current-sensing" resistor (Figure 1.41B). The good news is that we now have an output proportional to the current through the capacitor. The bad news is that the circuit is no longer a perfect mathematical differentiator. That's because the voltage across  $C$  (whose derivative produces the current we are sensing with  $R$ ) is no longer equal to  $V_{in}$ ; it now equals the difference between  $V_{in}$  and  $V_{out}$ . Here's how it goes: the voltage across  $C$  is  $V_{in} - V_{out}$ , so

$$I = C \frac{d}{dt} (V_{in} - V_{out}) = \frac{V_{out}}{R}.$$

If we choose  $R$  and  $C$  small enough so that  $dV_{out}/dt \ll dV_{in}/dt$ , then

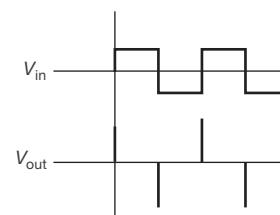
$$C \frac{dV_{in}}{dt} \approx \frac{V_{out}}{R}$$

or

$$V_{out}(t) \approx RC \frac{d}{dt} V_{in}(t).$$

That is, we get an output proportional to the rate of change of the input waveform.

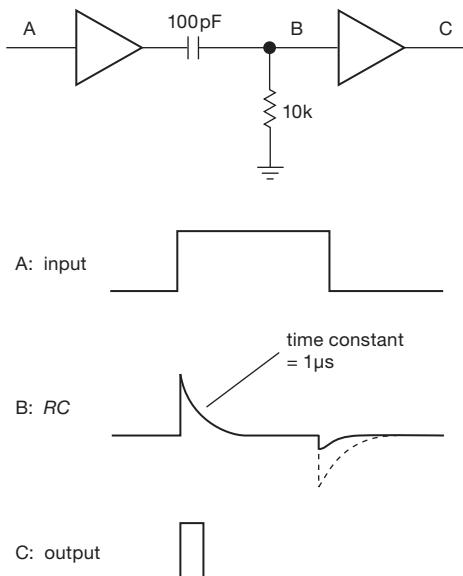
To keep  $dV_{out}/dt \ll dV_{in}/dt$ , we make the product  $RC$  small, taking care not to "load" the input by making  $R$  too small (at the transition the change in voltage across the capacitor is zero, so  $R$  is the load seen by the input). We will have a better criterion for this when we look at things in the frequency domain (§1.7.10). If you drive this circuit with a square wave, the output will be as shown in Figure 1.42.



**Figure 1.42.** Output waveform (bottom) from differentiator driven by a square wave.

Differentiators are handy for detecting *leading edges* and *trailing edges* in pulse signals, and in digital circuitry you sometimes see things like those depicted in Figure 1.43. The  $RC$  differentiator generates spikes at the transitions of the input signal, and the output buffer converts

<sup>25</sup> Devotees of the cinema will be reminded of Dr. Strangelove's outburst: "The whole point of a doomsday machine is lost ... if you keep it *secret!*"



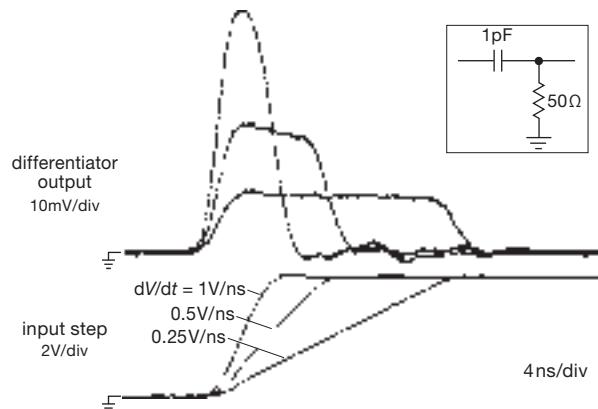
**Figure 1.43.** Leading-edge detector.

the spikes to short square-topped pulses. In practice, the negative spike will be small because of a diode (a handy device discussed in §1.6) built into the buffer.

To inject some real-world realism here, we hooked up and made some measurements on a differentiator that we configured for high-speed signals. For this we used  $C=1\text{ pF}$  and  $R=50\Omega$  (the latter is the world-wide standard for high-speed circuits, see Appendix H), we drove it with a 5 V step of settable slew rate (i.e.,  $dV/dt$ ). Figure 1.44 shows both input and output waveforms, for three choices of  $dV_{in}/dt$ . At these speeds (note the horizontal scale: 4 nanoseconds per division!) circuits often depart from ideal performance, as can be seen in the fastest risetime. The two slower steps show reasonable behavior; that is, a flat-top output waveform during the input's upward ramp; check for yourself that the output amplitude is correctly predicted by the formula.

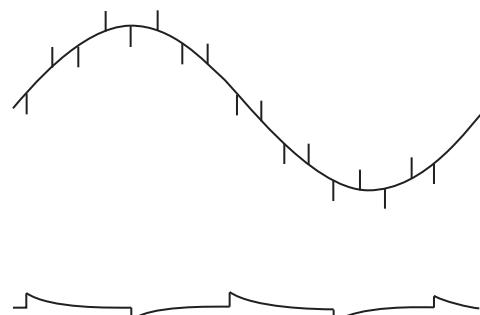
#### A. Unintentional capacitive coupling

Differentiators sometimes crop up unexpectedly, in situations where they're not welcome. You may see signals like those shown in Figure 1.45. The first case is caused by a square wave somewhere in the circuit coupling capacitively to the signal line you're looking at; that might indicate a missing resistor termination on your signal line. If not, you must either reduce the source resistance of the signal line or find a way to reduce capacitive coupling from the offending square wave. The second case is typical of what you might see when you look at a square wave, but have a



**Figure 1.44.** Three fast step waveforms, differentiated by the  $RC$  network shown. For the fastest waveform ( $10^9$  volts per second!), imperfections in the components and measuring instruments cause deviation from the ideal.

broken connection somewhere, usually at the scope probe. The very small capacitance of the broken connection combines with the scope input resistance to form a differentiator. Knowing that you've got a differentiated “something” can help you find the trouble and eliminate it.



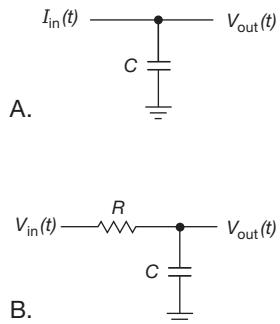
**Figure 1.45.** Two examples of unintentional capacitive coupling.

#### 1.4.4 Integrators

If  $RC$  circuits can take derivatives, why not integrals? As before, let's take it in two steps.

1. Imagine that we have an input signal that is a time-varying *current* versus time,  $I_{in}(t)$  (Figure 1.46A).<sup>26</sup> That input current is precisely the current through the capacitor, so  $I_{in}(t) = C dV(t)/dt$ , and therefore  $V(t) = \int I_{in}(t) dt$ .

<sup>26</sup> We're used to thinking of signals as time-varying *voltages*; but we'll see how we can convert such signals to proportional time-varying *currents*, by using “voltage-to-current converters” (with the fancier name “transconductance amplifiers”).



**Figure 1.46.** Integrator. A. Perfect (but requires a *current* input signal). B. Approximate (see text).

That's just what we wanted! Thus a simple capacitor, with one side grounded, is an integrator, *if* we have an input signal in the form of a current  $I_{in}(t)$ . Most of the time we don't, though.

2. So we connect a resistor in series with the more usual input *voltage* signal  $V_{in}(t)$ , to convert it to a current (Figure 1.46B). The good news is that it works, sort of. The bad news is that the circuit is no longer a perfect integrator. That's because the current through  $C$  (whose integral produces the output voltage) is no longer proportional to  $V_{in}$ ; it is now proportional to the difference between  $V_{in}$  and  $V$ . Here's how it goes: the voltage across  $R$  is  $V_{in} - V$ , so

$$I = C \frac{dV}{dt} = \frac{V_{in} - V}{R}.$$

If we manage to keep  $V \ll V_{in}$ , by keeping the product  $RC$  large,<sup>27</sup> then

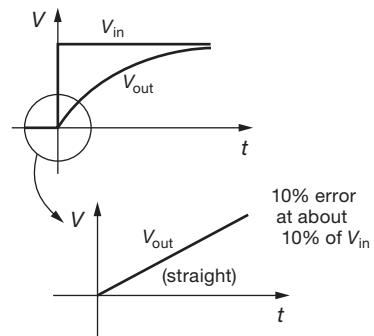
$$C \frac{dV}{dt} \approx \frac{V_{in}}{R}$$

or

$$V(t) = \frac{1}{RC} \int^t V_{in}(t) dt + \text{constant.}$$

That is, we get an output proportional to the integral over time of the input waveform. You can see how the approximation works for a square-wave input:  $V(t)$  is then the exponential charging curve we saw earlier (Figure 1.47). The first part of the exponential is a ramp, the integral of a constant; as we increase the time constant  $RC$ , we pick off a smaller part of the exponential, i.e., a better approximation to a perfect ramp.

Note that the condition  $V \ll V_{in}$  is the same as saying that  $I$  is proportional to  $V_{in}$ , which was our first integrator



**Figure 1.47.** Integrator approximation is good when  $V_{out} \ll V_{in}$ .

circuit. A large voltage across a large resistance approximates a current source and, in fact, is frequently used as one.

Later, when we get to operational amplifiers and feedback, we will be able to build integrators without the restriction  $V_{out} \ll V_{in}$ . They will work over large frequency and voltage ranges with negligible error.

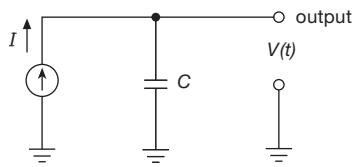
The integrator is used extensively in analog computation. It is a useful subcircuit that finds application in control systems, feedback, analog-digital conversion, and waveform generation.

### A. Ramp generators

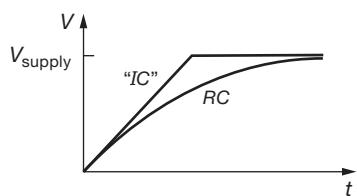
At this point it is easy to understand how a ramp generator works. This nice circuit is extremely useful, for example in timing circuits, waveform and function generators, analog oscilloscope sweep circuits, and analog-digital conversion circuitry. The circuit uses a constant current to charge a capacitor (Figure 1.48). From the capacitor equation  $I = C(dV/dt)$ , you get  $V(t) = (I/C)t$ . The output waveform is as shown in Figure 1.49. The ramp stops when the current source “runs out of voltage,” i.e., reaches the limit of its compliance. On the same figure is shown the curve for a simple  $RC$ , with the resistor tied to a voltage source equal to the compliance of the current source, and with  $R$  chosen so that the current at zero output voltage is the same as that of the current source. (Real current sources generally have output compliances limited by the power-supply voltages used in making them, so the comparison is realistic.) In the next chapter, which deals with transistors, we will design some current sources, with some refinements to follow in the chapters on operational amplifiers (op-amps) and FETs. Exciting things to look forward to!

<sup>27</sup> Just as with the differentiator, we'll have another way of framing this criterion in §1.7.10.

**Exercise 1.18.** A current of 1 mA charges a 1  $\mu\text{F}$  capacitor. How long does it take the ramp to reach 10 volts?



**Figure 1.48.** A constant-current source charging a capacitor generates a ramp voltage waveform.



**Figure 1.49.** Constant-current charging (with finite compliance) versus  $RC$  charging.

#### 1.4.5 Not quite perfect...

Real capacitors (the kind you can see, and touch, and pay money for) generally behave according to theory; but they have some additional “features” that can cause problems in some demanding applications. For example, all capacitors exhibit *some series resistance* (which may be a function of frequency), and *some series inductance* (see the next section), along with some frequency-dependent parallel resistance. Then there’s a “memory” effect (known as *dielectric absorption*), which is rarely discussed in polite society: if you charge a capacitor up to some voltage  $V_0$  and hold it there for a while, and then discharge it to 0 V, then when you remove the short across its terminals it will tend to drift back a bit toward  $V_0$ .

Don’t worry about this stuff, for now. We’ll treat in detail these effects (and other oddities of real-world components) in the advanced topics Chapter 1x.

## 1.5 Inductors and transformers

### 1.5.1 Inductors

If you understand capacitors, you shouldn’t have great trouble with inductors (Figure 1.50). They’re closely related to capacitors: the rate of current change in an inductor is proportional to the voltage applied across it (for a capacitor it’s the other way around – the rate of *voltage* change is proportional to the *current* through it). The defining equation for an inductor is

$$V = L \frac{dI}{dt}, \quad (1.23)$$

where  $L$  is called the *inductance* and is measured in henrys (or mH,  $\mu$ H, nH, etc.). Putting a constant voltage across an inductor causes the current to rise as a ramp (compare with a capacitor, in which a constant *current* causes the *voltage* to rise as a ramp); 1 V across 1 H produces a current that increases at 1 amp per second.



**Figure 1.50.** Inductors. The parallel-bar symbol represents a core of magnetic material.

Just as with capacitors, the energy invested in ramping up the current in an inductor is stored internally, here in the form of magnetic fields. And the analogous formula is

$$U_L = \frac{1}{2} L I^2, \quad (1.24)$$

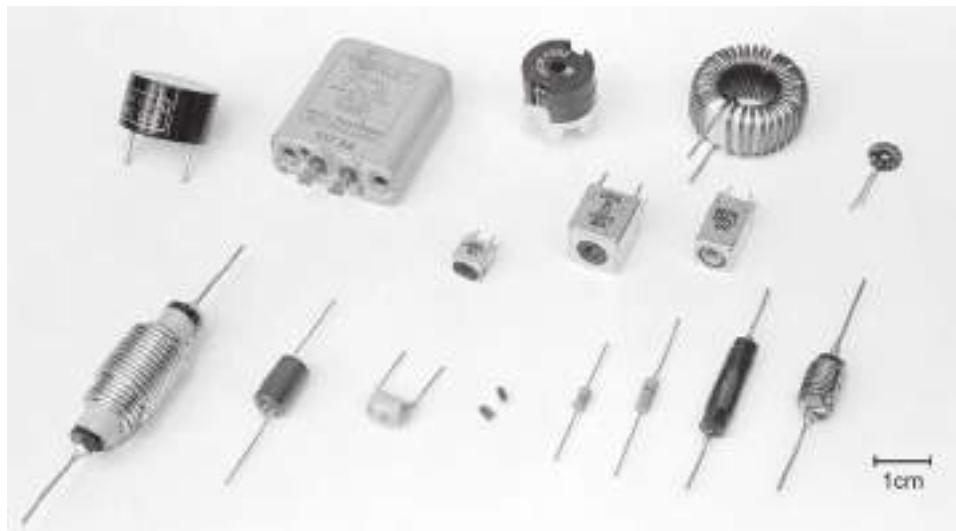
where  $U_L$  is in joules (watt seconds) for  $L$  in henrys and  $I$  in amperes. As with capacitors, this is an important result, one which lies at the core of switching power conversion (exemplified by those little black “wall-warts” that provide power to all manner of consumer electronic gadgets). We’ll see lots more of this in Chapter 9.

The symbol for an inductor looks like a coil of wire; that’s because, in its simplest form, that’s all it is. Its somewhat peculiar behavior comes about because inductors are magnetic devices, in which two things are going on: current flowing through the coil creates a magnetic field aligned along the coil’s axis; and then changes in that field produce a voltage (sometimes called “back EMF”) in a way that tries to cancel out those changes (an effect known as Lenz’s law). The inductance  $L$  of a coil is simply the ratio of magnetic flux passing through the coil divided by the current through the coil that produces that flux (multiplied by an overall constant). Inductance depends on the coil geometry (e.g., diameter and length) and the properties of any magnetic material (the “core”) that may be used to confine the magnetic field. That’s all you need to understand why the inductance of a coil of given geometry is proportional to the square of the number of turns.

**Exercise 1.19.** Explain why  $L \propto n^2$  for an inductor consisting of a coil of  $n$  turns of wire, maintaining fixed diameter and length as  $n$  is varied.

We’ll get into some more detail in the Chapter 1x. But it’s worth displaying a semi-empirical formula for the approximate inductance  $L$  of a coil of diameter  $d$  and length  $l$ , in which the  $n^2$  dependence is on display:

$$L \approx K \frac{d^2 n^2}{18d + 40l} \quad \mu\text{H},$$



**Figure 1.51.** Inductors. Top row, left to right: encapsulated toroid, hermetically-sealed toroid, board-mount pot core, bare toroid (two sizes). Middle row: slug-tuned ferrite-core inductors (three sizes). Bottom row: high-current ferrite-core choke, ferrite-bead choke, dipped radial-lead ferrite-core inductor, surface-mount ferrite chokes, molded axial-lead ferrite-core chokes (two styles), lacquered ferrite-core inductors (two styles).

where  $K = 1.0$  or  $0.395$  for dimensions in inches or centimeters, respectively. This is known as Wheeler's formula and is accurate to  $1\%$  as long as  $l > 0.4d$ .

As with capacitive current, inductive current is not simply proportional to voltage (as in a resistor). Furthermore, unlike the situation in a resistor, the power associated with inductive current ( $V \times I$ ) is not turned into heat, but is stored as energy in the inductor's magnetic field (recall that for a capacitor the power associated with capacitive current is likewise not dissipated as heat, but is stored as energy in the capacitor's electric field). You get all that energy back when you interrupt the inductor's current (with a capacitor you get all the energy back when you discharge the voltage to zero).

The basic inductor is a coil, which may be just a loop with one or more turns of wire; or it may be a coil with some length, known as a solenoid. Variations include coils wound on various core materials, the most popular being iron (or iron alloys, laminations, or powder) and ferrite (a gray, nonconductive, brittle magnetic material). These are all ploys to multiply the inductance of a given coil by the "permeability" of the core material. The core may be in the shape of a rod, a toroid (doughnut), or even more bizarre shapes, such as a "pot core" (which has to be seen to be understood; the best description we can come up with is a doughnut mold split horizontally in half, if doughnuts were

made in molds). See Figure 1.51 for some typical geometries.

Inductors find heavy use in radiofrequency (RF) circuits, serving as RF "chokes" and as parts of tuned circuits (§1.7.14). A pair of closely coupled inductors forms the interesting object known as a transformer. We will talk briefly about them shortly.

An inductor is, in a real sense, the opposite of a capacitor.<sup>28</sup> You will see how that works out later in the chapter when we deal with the important subject of *impedance*.

#### A. A look ahead: some magic with inductors

Just to give a taste of some of the tricks that you can do with inductors, take a look at Figure 1.52. Although we'll understand these circuits a lot better when we go at them in Chapter 9, it's possible to see what's going on with what we know already. In Figure 1.52A the left-hand side of inductor  $L$  is alternately switched between a dc input voltage  $V_{in}$  and ground, at some rapid rate, spending equal times

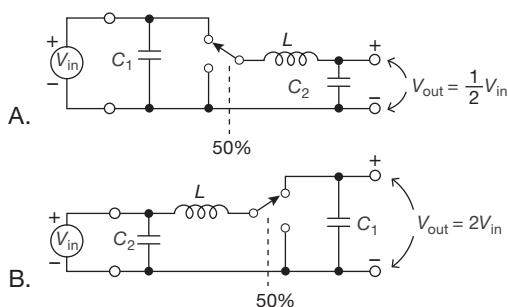
<sup>28</sup> In practice, however, capacitors are much more widely used in electronic circuits. That is because practical inductors depart significantly from ideal performance – by having winding resistance, core losses, and self-capacitance – whereas practical capacitors are nearly perfect (more on this in Chapter 1x). Inductors are indispensable, however, in *switching power converters*, as well as in tuned  $LC$  circuits for RF applications.

connected to each (a “50% duty cycle”). But the defining equation  $V=LdI/dt$  requires that the *average* voltage across an inductor must be zero, otherwise the magnitude of its average current is rising without limit. (This is sometimes called the *volt-second balance rule*.) From this it follows that the average output voltage is half the input voltage (make sure you understand why). In this circuit  $C_2$  acts as a storage capacitor for steadyng the output voltage (more on this later, and in Chapter 9).

Producing an output that is half the voltage of an input is not very exciting; after all, a simple voltage divider can do that. But, unlike a voltage divider, this circuit does not waste any energy; apart from non-idealities of the components, it is 100% efficient. And in fact this circuit is widely used in power conversion; it's called a “synchronous buck converter.”

But look now at Figure 1.52B, which is just a turned-around version of Figure 1.52A. This time, volt-second balance requires that the output voltage be *twice* the input voltage. You can't do *that* with a voltage divider! Once again, the output capacitor ( $C_1$  this time) serves to hold the output voltage steady by storing charge. This configuration is called a “synchronous boost converter.”

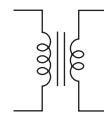
These and other switching converters are discussed extensively in Chapter 9, where Table 9.5 lists some fifty representative types.



**Figure 1.52.** Inductors let you do neat tricks, such as *increasing* a dc input voltage.

## 1.5.2 Transformers

A transformer is a device consisting of two closely coupled coils (called primary and secondary). An ac voltage applied to the primary appears across the secondary, with a voltage multiplication proportional to the turns ratio of the transformer, and with a current multiplication inversely proportional to the turns ratio. Power is conserved. Figure 1.53 shows the circuit symbol for a laminated-core transformer (the kind used for 60 Hz ac power conversion).



**Figure 1.53.** Transformer.

Transformers are quite efficient (output power is very nearly equal to input power); thus, a step-up transformer gives higher voltage at lower current. Jumping ahead for a moment, a transformer of turns ratio  $n$  increases the impedance by  $n^2$ . There is very little primary current if the secondary is unloaded.

*Power transformers* (meant for use from the 115 V powerline) serve two important functions in electronic instruments: they change the ac line voltage to a useful (usually lower) value that can be used by the circuit, and they “isolate” the electronic device from actual connection to the powerline, because the windings of a transformer are electrically insulated from each other. They come in an enormous variety of secondary voltages and currents: outputs as low as 1 volt or so up to several thousand volts, current ratings from a few millamps to hundreds of amps. Typical transformers for use in electronic instruments might have secondary voltages from 10 to 50 volts, with current ratings of 0.1 to 5 amps or so. A related class of transformers is used in electronic power conversion, in which plenty of power is flowing, but typically as pulse or square waveforms, and at much higher frequencies (50 kHz to 1 MHz is typical).

Transformers for signals at audio frequencies and radio frequencies are also available. At radio frequencies you sometimes use tuned transformers if only a narrow range of frequencies is present. There is also an interesting class of transmission-line transformers. In general, transformers for use at high frequencies must use special core materials or construction to minimize core losses, whereas low-frequency transformers (e.g., ac powerline transformers) are burdened instead by large and heavy cores. The two kinds of transformers are in general not interchangeable.

## A. Problems, problems...

This simple “first-look” description ignores interesting – and important – issues. For example, there are inductances associated with the transformer, as suggested by its circuit symbol: an effective parallel inductance (called the *magnetizing inductance*) and an effective series inductance (called the *leakage inductance*). Magnetizing inductance causes a primary current even with no secondary load; more significantly, it means that you cannot make a “dc

transformer.” And leakage inductance causes a voltage drop that depends on load current, as well as bedeviling circuits that have fast pulses or edges. Other departures from ideal performance include winding resistance, core losses, capacitance, and magnetic coupling to the outside world. Unlike capacitors (which behave nearly ideally in most circuit applications), the deficiencies of inductors have significant effects in real-world circuit applications. We’ll deal with these in Chapter 1x and Chapter 9.

## 1.6 Diodes and diode circuits

We are not done with capacitors and inductors! We have dealt with them in the *time domain* (*RC* circuits, exponential charge and discharge, differentiators and integrators, and so on), but we have not yet tackled their behavior in the *frequency domain*.

We will get to that soon enough. But this is a good time to take a break from “*RLC*” and put our knowledge to use with some clever and useful circuits. We begin by introducing a new device, the *diode*. It’s our first example of a nonlinear device, and you can do nifty things with it.

### 1.6.1 Diodes

The circuit elements we’ve discussed so far (resistors, capacitors, and inductors) are all *linear*, meaning that a doubling of the applied signal (a voltage, say) produces a doubling of the response (a current, say). This is true even for the reactive devices (capacitors and inductors). These components are also *passive*, as opposed to *active* devices, the latter exemplified by transistors, which are semiconductor devices that control the flow of power. And they are all two-terminal devices, which is self-explanatory.

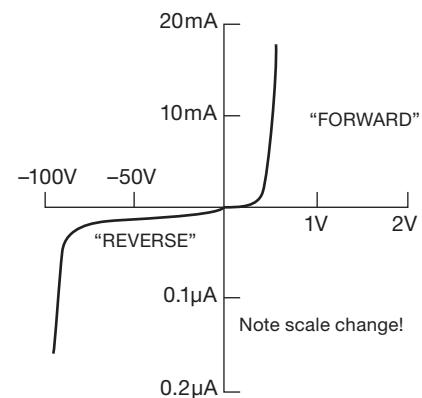


**Figure 1.54.** Diode.

The diode (Figure 1.54) is an important and useful two-terminal passive *nonlinear* device. It has the *V–I* curve shown in Figure 1.55. (In keeping with the general philosophy of this book, we will not attempt to describe the solid-state physics that makes such devices possible.)

The diode’s arrow (the anode terminal) points in the direction of forward current flow. For example, if the diode is in a circuit in which a current of 10 mA is flowing from anode to cathode, then (from the graph) the anode is approximately 0.6 V more positive than the cathode; this is called the “forward voltage drop.” The reverse current,

which is measured in the nanoamp range for a general-purpose diode (note the hugely different scales in the graph for forward and reverse current), is almost never of any consequence until you reach the reverse breakdown voltage (also called the peak inverse voltage, PIV), typically 75 volts for a general-purpose diode like the 1N4148. (Normally you never subject a diode to voltages large enough to cause reverse breakdown; the exception is the zener diode we mentioned earlier.) Frequently, also, the forward voltage drop of about 0.5 to 0.8 V is of little concern, and the diode can be treated as a good approximation to an ideal one-way conductor. There are other important characteristics that distinguish the thousands of diode types available, e.g., maximum forward current, capacitance, leakage current, and reverse recovery time; Table 1.1 includes a few popular diodes, to give a sense of the capabilities of these little devices.



**Figure 1.55.** Diode *V–I* curve.

Before jumping into some circuits with diodes, we should point out two things: (a) a diode doesn’t have a resistance (it doesn’t obey Ohm’s law). (b) If you put some diodes in a circuit, it won’t have a Thévenin equivalent.

### 1.6.2 Rectification

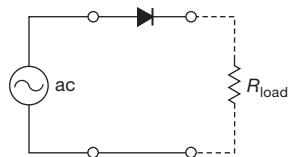
A rectifier changes ac to dc; this is one of the simplest and most important applications of diodes (which are sometimes called rectifiers). The simplest circuit is shown in Figure 1.56. The “ac” symbol represents a source of ac voltage; in electronic circuits it is usually provided by a transformer, powered from the ac powerline. For a sine-wave input that is much larger than the forward drop (about 0.6 V for silicon diodes, the usual type), the output will look like that in Figure 1.57. If you think of the diode as a one-way conductor, you won’t have any trouble

**Table 1.1 Representative Diodes**

Part #	$V_R$ (max) (V)	$I_R$ (typ, 25°C) (A @ V)	$V_F$ @ $I_F$ (mV)	$I_F$ (mA)	Capacitance (pF @ $V_R$ )	SMT <sup>a</sup> p/n	Comments
<i>Silicon</i>							
PAD5	45	0.25pA	20V	800	1	0.5pF 5V	SSTPAD5 metal + glass can
1N4148	75	10nA	20V	750	10	0.9pF 0V	1N4148W jellybean sig diode
1N4007	1000	50nA	800V	800	250	12pF 10V	DL4007 1N4004 lower V
1N5406	600	<10μA	600V	1.0V	10A	18pF 10V	none heat through leads
<i>Schottky</i> <sup>b</sup>							
1N6263	60	7nA	20V	400	1	0.6pF 10V	1N6263W see also 1N5711
1N5819	40	10μA	32V	400	1000	150pF 1V	1N5819HW jellybean
1N5822	40	40μA	32V	480	3000	450pF 1V	none power Schottky
MBRP40045	45	500μA	40V	540	400A	3500pF 10V	you jest! Moby dual Schottky

Notes: (a) SMT, surface-mount technology. (b) Schottky diodes have lower forward voltage and zero reverse-recovery time, but more capacitance.

understanding how the circuit works. This circuit is called a *half-wave rectifier*, because only half of the input waveform is used.

**Figure 1.56.** Half-wave rectifier.

which the diode drop becomes significant, you have to remember that.<sup>29</sup>

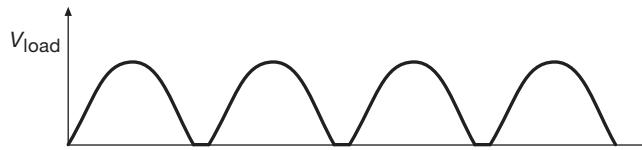
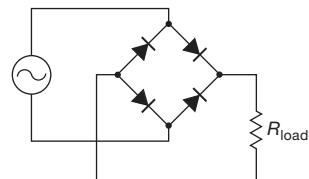
**Figure 1.59.** Full-wave output voltage (unfiltered).**Figure 1.57.** Half-wave output voltage (unfiltered).**Figure 1.58.** Full-wave bridge rectifier.

Figure 1.58 shows another rectifier circuit, a “full-wave bridge.” Figure 1.59 shows the voltage across the load; note that the entire input waveform is used. The gaps at zero voltage occur because of the diodes’ forward voltage drop. In this circuit, two diodes are always in series with the input; when you design low-voltage power supplies, for

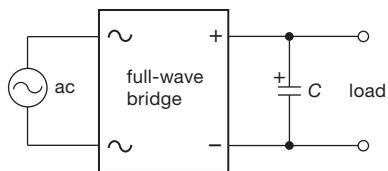
### 1.6.3 Power-supply filtering

The preceding rectified waveforms aren’t good for much as they stand. They’re “dc” only in the sense that they don’t change polarity. But they still have a lot of “ripple” (periodic variations in voltage about the steady value) that has to be smoothed out in order to generate genuine dc. This we do by attaching a relatively large value capacitor (Figure 1.60); it charges up to the peak output voltage during the diode conduction, and its stored charge ( $Q = CV$ ) provides the output current in between charging cycles. Note that the diodes prevent the capacitor from discharging back through the ac source. In this application you should think of the capacitor as an energy storage device, with stored energy  $U = \frac{1}{2}CV^2$  (recall §1.4.1; for  $C$  in farads and  $V$  in volts,  $U$  comes out in joules, or equivalently, watt seconds).

The capacitor value is chosen so that

$$R_{\text{load}}C \gg 1/f,$$

<sup>29</sup> The diode drop can be eliminated with *active switching* (or *synchronous switching*, a technique in which the diodes are replaced by transistor switches, actuated in synchronism with the input ac waveform (see §9.5.3B).



**Figure 1.60.** Full-wave bridge with output storage (“filter”) capacitor.

(where  $f$  is the ripple frequency, here 120 Hz) in order to ensure small ripple by making the time constant for discharge much longer than the time between recharging. We make this vague statement clearer now.

### A. Calculation of ripple voltage

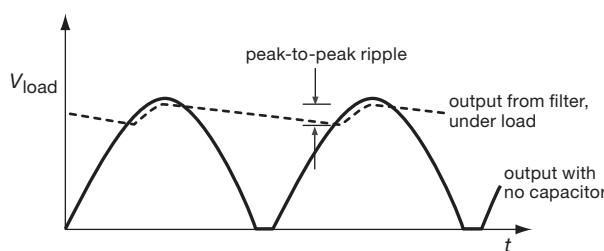
It is easy to calculate the approximate ripple voltage, particularly if it is small compared with the dc (see Figure 1.61). The load causes the capacitor to discharge somewhat between cycles (or half-cycles, for full-wave rectification). If you assume that the load current stays constant (it will, for small ripple), you have

$$\Delta V = \frac{I}{C} \Delta t \quad \left( \text{from } I = C \frac{dV}{dt} \right). \quad (1.25)$$

Just use  $1/f$  (or  $1/2f$  for full-wave rectification) for  $\Delta t$  (this estimate is a bit on the safe side, because the capacitor begins charging again in less than a half-cycle). You get<sup>30</sup>

$$\Delta V = \frac{I_{\text{load}}}{fC} \quad (\text{half-wave}),$$

$$\Delta V = \frac{I_{\text{load}}}{2fC} \quad (\text{full-wave}).$$



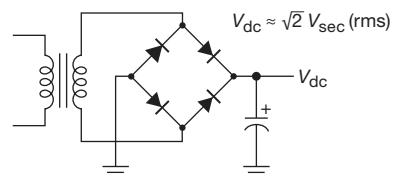
**Figure 1.61.** Power-supply ripple calculation.

If you wanted to do the calculation without any approximation, you would use the exact exponential discharge formula. You would be misguided in insisting on that kind

of accuracy, though, for two reasons. (a) The discharge is an exponential only if the load is a resistance; many loads are not. In fact, the most common load, a *voltage regulator*, looks like a constant-current load. (b) Power supplies are built with capacitors with typical tolerances of 20% or more. Realizing the manufacturing spread, you design conservatively, allowing for the worst-case combination of component values.

In this case, viewing the initial part of the discharge as a ramp is in fact quite accurate, especially if the ripple is small, and in any case it errs in the direction of conservative design – it overestimates the ripple.

**Exercise 1.20.** Design a full-wave bridge rectifier circuit to deliver 10 V dc with less than 0.1 V (pp) ripple into a load drawing up to 10 mA. Choose the appropriate ac input voltage, assuming 0.6 V diode drops. Be sure to use the correct ripple frequency in your calculation.



**Figure 1.62.** Bridge rectifier circuit. The polarity marking and curved electrode indicate a polarized capacitor, which must not be allowed to charge with the opposite polarity.

### 1.6.4 Rectifier configurations for power supplies

#### A. Full-wave bridge

A dc power supply with the bridge circuit we just discussed looks as shown in Figure 1.62. In practice, you generally buy the bridge as a prepackaged module. The smallest ones come with maximum current ratings of 1 A average, with a selection of rated minimum breakdown voltages going from 100 V to 600 V, or even 1000 V. Giant bridge rectifiers are available with current ratings of 25 A or more.

#### B. Center-tapped full-wave rectifier

The circuit in Figure 1.63 is called a center-tapped full-wave rectifier. The output voltage is half what you get if you use a bridge rectifier. It is not the most efficient circuit in terms of transformer design, because each half of the secondary is used only half the time. To develop some intuition on this subtle point, consider two different configurations that produce the same rectified dc output voltage: (a) the circuit of Figure 1.63, and (b) the same transformer, this time with its secondary cut at the center tap and

<sup>30</sup> While teaching electronics we've noticed that students love to memorize these equations! An informal poll of the authors showed that two out of two engineers don't memorize them. Please don't waste brain cells that way – instead, learn how to derive them.

rewired with the two halves in parallel, the resultant combined secondary winding connected to a full-wave bridge. Now, to deliver the same output power, each half winding in (a), during its conduction cycle, must supply the same current as the parallel pair in (b). But the power dissipated in the winding resistances goes like  $I^2R$ , so the power lost to heating in the transformer secondary windings reduced by a factor of 2 for the bridge configuration (b).

Here's another way to see the problem: imagine we use the same transformer as in (a), but for our comparison circuit we replace the pair of diodes with a bridge, as in Figure 1.62, and we leave the center tap unconnected. Now, to deliver the same output *power*, the current through the winding during that time is twice what it would be for a true full-wave circuit. To expand on this subtle point: heating in the windings, calculated from Ohm's law, is  $I^2R$ , so you have four times the heating for half the time, or twice the average heating of an equivalent full-wave bridge circuit. You would have to choose a transformer with a current rating 1.4 (square root of 2) times as large compared with the (better) bridge circuit; besides costing more, the resulting supply would be bulkier and heavier.

**Exercise 1.21.** This illustration of  $I^2R$  heating may help you understand the disadvantage of the center-tapped rectifier circuit. What fuse rating (minimum) is required for passing the current waveform shown in Figure 1.64, which has 1 amp average current? Hint: a fuse "blows out" by melting a metallic link ( $I^2R$  heating), for steady currents larger than its rating. Assume for this problem that the thermal time constant of the fusible link is much longer than the time scale of the square wave, i.e., that the fuse responds to the value of  $I^2$  averaged over many cycles.

### C. Split supply

A popular variation of the center-tapped full-wave circuit is shown in Figure 1.65. It gives you split supplies (equal plus and minus voltages), which many circuits need. It is an efficient circuit, because both halves of the input waveform are used in each winding section.

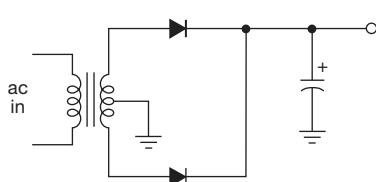


Figure 1.63. Full-wave rectifier using center-tapped transformer.

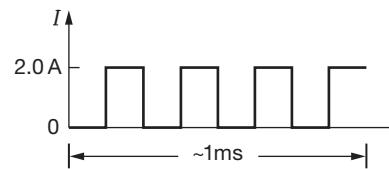


Figure 1.64. Illustrating greater  $I^2R$  heating with discontinuous current flow.

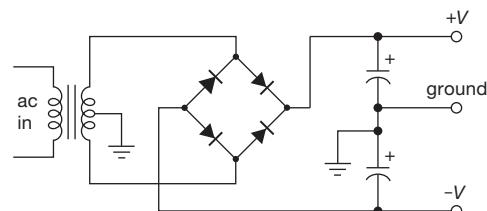


Figure 1.65. Dual-polarity (split) supply.

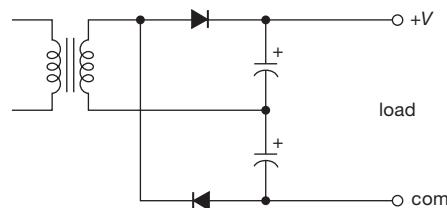


Figure 1.66. Voltage doubler.

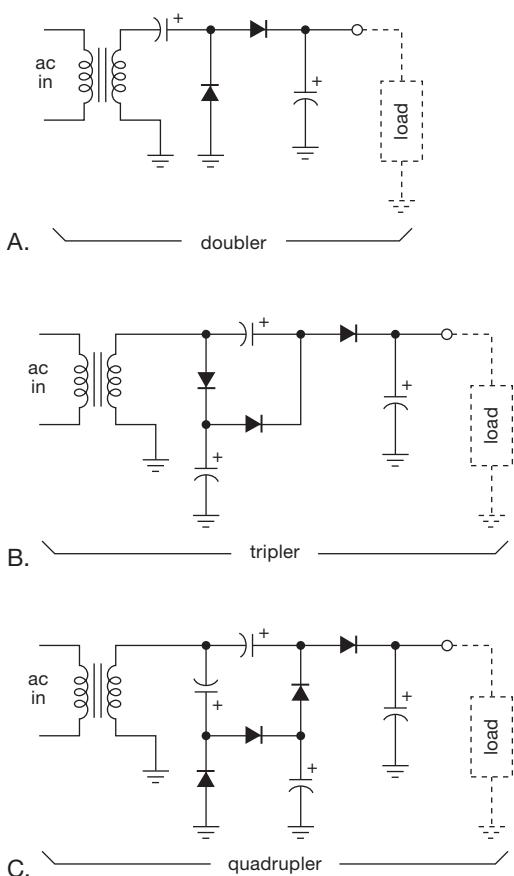
### D. Voltage multipliers

The circuit shown in Figure 1.66 is called a voltage doubler. Think of it as two half-wave rectifier circuits in series. It is officially a full-wave rectifier circuit because both halves of the input waveform are used – the ripple frequency is twice the ac frequency (120 Hz for the 60 Hz line voltage in the United States).

Variations of this circuit exist for voltage triplers, quadruplers, etc. Figure 1.67 shows doubler, tripler, and quadrupler circuits that let you ground one side of the transformer. You can extend this scheme as far as you want, producing what's called a Cockcroft–Walton generator; these are used in arcane applications (such as particle accelerators) and in everyday applications (such as image intensifiers, air ionizers, laser copiers, and even bug zappers) that require a high dc voltage but hardly any current.

### 1.6.5 Regulators

By choosing capacitors that are sufficiently large, you can reduce the ripple voltage to any desired level. This brute-force approach has three disadvantages.

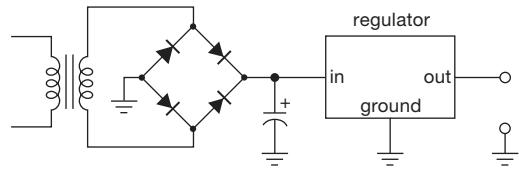


**Figure 1.67.** Voltage multipliers; these configurations don't require a floating voltage source.

- The required capacitors may be prohibitively bulky and expensive.
- The very short interval of current flow during each cycle<sup>31</sup> (only very near the top of the sinusoidal waveform) produces more  $I^2R$  heating.
- Even with the ripple reduced to negligible levels, you still have variations of output voltage that are due to other causes, e.g., the dc output voltage will be roughly proportional to the ac input voltage, giving rise to fluctuations caused by input line voltage variations. In addition, changes in load current will still cause the output voltage to change because of the finite internal resistances of the transformer, diode, etc. In other words, the Thévenin equivalent circuit of the dc power supply has  $R > 0$ .

A better approach to power-supply design is to use enough capacitance to reduce ripple to low levels (perhaps

<sup>31</sup> Called the *conduction angle*.



**Figure 1.68.** Regulated dc power supply.

10% of the dc voltage), then use an active *feedback circuit* to eliminate the remaining ripple. Such a feedback circuit “looks at” the output, making changes in a controllable series resistor (a transistor) as necessary to keep the output voltage constant (Figure 1.68). This is known as a “linear regulated dc power supply.”<sup>32</sup>

These voltage regulators are used almost universally as power supplies for electronic circuits. Nowadays complete voltage regulators are available as inexpensive ICs (priced under \$1). A power supply built with a voltage regulator can be made easily adjustable and self-protecting (against short circuits, overheating, etc.), with excellent properties as a voltage source (e.g., internal resistance measured in millionohms). We will deal with regulated dc power supplies in Chapter 9.

## 1.6.6 Circuit applications of diodes

### A. Signal rectifier

There are other occasions when you use a diode to make a waveform of one polarity only. If the input waveform isn't a sinewave, you usually don't think of it as a rectification in the sense of a power supply. For instance, you might want a train of pulses corresponding to the rising edge of a square wave. The easiest way is to rectify the differentiated wave (Figure 1.69). Always keep in mind the 0.6 V(approximately) forward drop of the diode. This circuit, for instance, gives no output for square waves smaller than 0.6 V pp. If this is a problem, there are various tricks to circumvent this limitation. One possibility is to use *hot carrier diodes* (Schottky diodes), with a forward drop of about 0.25 V.

A possible *circuit solution* to this problem of finite diode drop is shown in Figure 1.70. Here  $D_1$  compensates  $D_2$ 's forward drop by providing 0.6 V of bias to hold  $D_2$  at the threshold of conduction. Using a diode ( $D_1$ ) to provide the bias (rather than, say, a voltage divider) has several

<sup>32</sup> A popular variant is the regulated switching power converter. Although its operation is quite different in detail, it uses the same feedback principle to maintain a constant output voltage. See Chapter 9 for much more on both techniques.

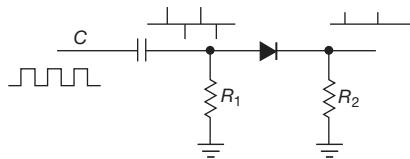


Figure 1.69. Signal rectifier applied to differentiator output.

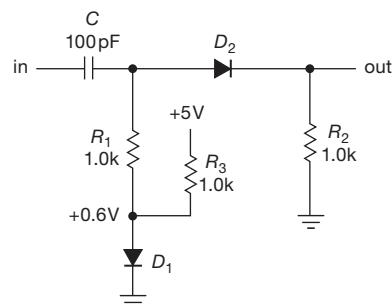


Figure 1.70. Compensating the forward voltage drop of a diode signal rectifier.

advantages: (a) there is nothing to adjust, (b) the compensation will be nearly perfect, and (c) changes of the forward drop (e.g., with changing temperature) will be compensated properly. Later we will see other instances of matched-pair compensation of forward drops in diodes, transistors, and FETs. It is a simple and powerful trick.

### B. Diode gates

Another application of diodes, which we will recognize later under the general heading of *logic*, is to pass the higher of two voltages without affecting the lower. A good example is *battery backup*, a method of keeping something running (e.g., the “real-time clock” chip in a computer, which keeps a running count of date and time) that must continue running even when the device is switched off. Figure 1.71 shows a circuit that does the job. The battery does nothing until the +5 V power is switched off; then it takes over without interruption.

### C. Diode clamps

Sometimes it is desirable to limit the range of a signal (i.e., prevent it from exceeding certain voltage limits) somewhere in a circuit. The circuit shown in Figure 1.72 will accomplish this. The diode prevents the output from exceeding about +5.6 V, with no effect on voltages less than that (including negative voltages); the only limitation is that the input must not go so negative that the reverse breakdown voltage of the diode is exceeded (e.g., -75 V for a 1N4148). The series resistor limits the diode current during

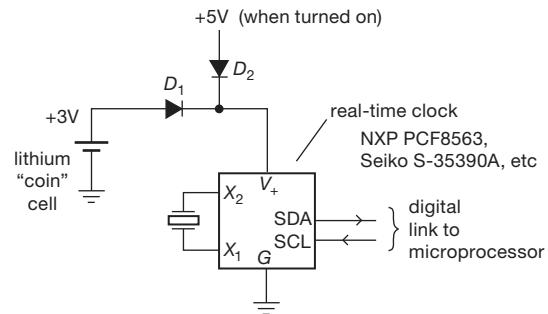


Figure 1.71. Diode OR gate: battery backup. The real-time clock chips are specified to operate properly with supply voltages from +1.8 V to +5.5 V. They draw a paltry  $0.25 \mu\text{A}$ , which calculates to a 1-million-hour life (a hundred years) from a standard CR2032 coin cell!

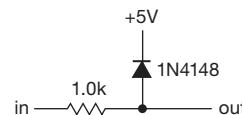


Figure 1.72. Diode voltage clamp.

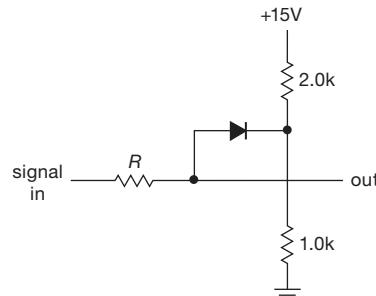


Figure 1.73. Voltage divider providing clamping voltage.

clamping action; however, a side effect is that it adds  $1\text{k}\Omega$  of series resistance (in the Thévenin sense) to the signal, so its value is a compromise between maintaining a desirable low source (Thévenin) resistance and a desirable low clamping current. Diode clamps are standard equipment on all inputs in contemporary CMOS digital logic. Without them, the delicate input circuits are easily destroyed by static electricity discharges during handling.

**Exercise 1.22.** Design a symmetrical clamp, i.e., one that confines a signal to the range -5.6 to +5.6 V.

A voltage divider can provide the reference voltage for a clamp (Figure 1.73). In this case you must ensure that the resistance looking into the voltage divider ( $R_{vd}$ ) is small compared with  $R$  because what you have looks as shown

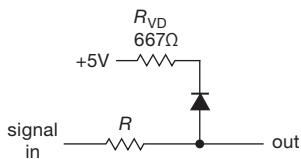


Figure 1.74. Clamping to voltage divider: equivalent circuit.

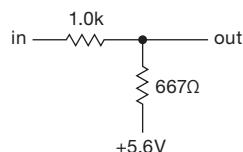


Figure 1.75. Poor clamping: voltage divider not stiff enough.

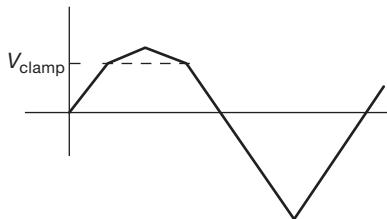


Figure 1.76. Clamping waveform for circuit of Figure 1.73.

in Figure 1.74 when the voltage divider is replaced with its Thévenin equivalent circuit. When the diode conducts (input voltage exceeds clamp voltage), the output is really just the output of a voltage divider, with the Thévenin equivalent resistance of the voltage reference as the lower resistor (Figure 1.75). So, for the values shown, the output of the clamp for a triangle-wave input would look as shown in Figure 1.76. The problem is that the voltage divider doesn't provide a stiff reference, in the language of electronics. A stiff voltage source is one that doesn't bend easily, i.e., it has low internal (Thévenin) resistance.

In practice, the problem of finite impedance of the voltage-divider reference can be easily solved by use of a transistor or an op-amp. This is usually a better solution than using very small resistor values, because it doesn't consume large currents, yet it provides a voltage reference with a Thévenin resistance of a few ohms or less. Furthermore, there are other ways to construct a clamp, using an op-amp as part of the clamp circuit. You will see these methods in Chapter 4.

Alternatively, a simple way to stiffen the clamp circuit of Figure 1.73, for time-varying signals only, is to add a so-called *bypass capacitor* across the lower ( $1\text{k}\Omega$ ) resistor. To understand this fully we need to know about capacitors in

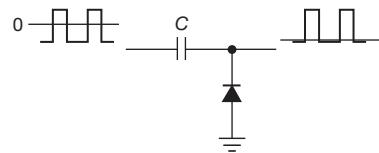


Figure 1.77. dc restoration.

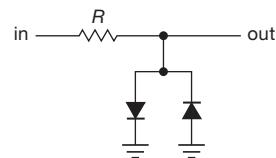


Figure 1.78. Diode limiter.

the frequency domain, a subject we'll take up shortly. For now we'll simply say that you can put a capacitor across the  $1\text{k}$  resistor, and its stored charge acts to maintain that point at constant voltage. For example, a  $15\text{\mu F}$  capacitor to ground would make the divider look as if it had a Thévenin resistance of less than  $10\Omega$  for frequencies above  $1\text{kHz}$ . (You could similarly add a bypass capacitor across  $D_1$  in Figure 1.70.) As we'll learn, the effectiveness of this trick decreases at low frequencies, and it does nothing at dc.

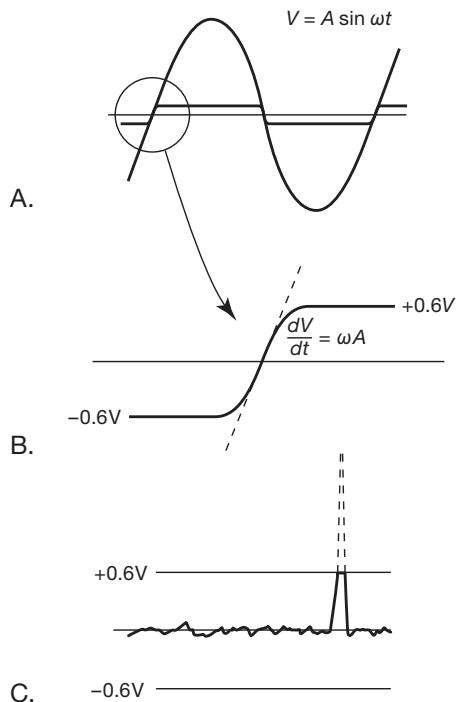
One interesting clamp application is "dc restoration" of a signal that has been ac coupled (capacitively coupled). Figure 1.77 shows the idea. This is particularly important for circuits whose inputs look like diodes (e.g., a transistor with grounded emitter, as we'll see in the next chapter); otherwise an ac-coupled signal will just fade away, as the coupling capacitor charges up to the signal's peak voltage.

#### D. Limiter

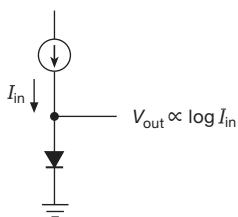
One last clamp circuit is shown in Figure 1.78. This circuit limits the output "swing" (again, a common electronics term) to one diode drop in either polarity, roughly  $\pm 0.6\text{V}$ . That might seem awfully small, but if the next stage is an amplifier with large voltage amplification, its input will always be near  $0\text{V}$ ; otherwise the output is in "saturation" (e.g., if the next stage has a gain of 1000 and operates from  $\pm 15\text{V}$  supplies, its input must stay in the range  $\pm 15\text{mV}$  in order for its output not to saturate). Figure 1.79 shows what a limiter does to oversize sinewaves and spikes. This clamp circuit is often used as input protection for a high-gain amplifier.

#### E. Diodes as nonlinear elements

To a good approximation the forward current through a diode is proportional to an exponential function of the



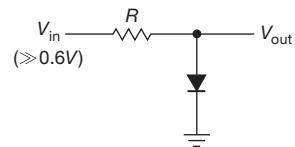
**Figure 1.79.** A. Limiting large-amplitude sinewaves; B. details; and C. spikes.



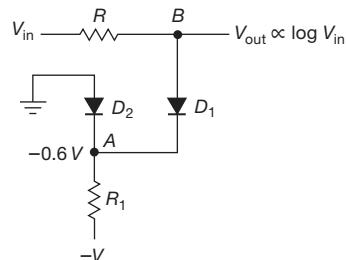
**Figure 1.80.** Exploiting the diode's nonlinear  $V-I$  curve: logarithmic converter.

voltage across it at a given temperature (for a discussion of the exact law, see §2.3.1). So you can use a diode to generate an output voltage proportional to the logarithm of a current (Figure 1.80). Because  $V$  hovers in the region of 0.6 V, with only small voltage changes that reflect input current variations, you can generate the input current with a resistor if the input voltage is much larger than a diode drop (Figure 1.81).

In practice, you may want an output voltage that isn't offset by the 0.6 V diode drop. In addition, it would be nice to have a circuit that is insensitive to changes in temperature (a silicon diode's voltage drop decreases approximately  $2 \text{ mV}/^\circ\text{C}$ ). The method of diode drop compensation is helpful here (Figure 1.82).  $R_1$  makes  $D_2$  conduct, holding



**Figure 1.81.** Approximate log converter.



**Figure 1.82.** Diode drop compensation in the logarithmic converter.

point  $A$  at about  $-0.6 \text{ V}$ . Point  $B$  is then near ground (making  $I_{in}$  accurately proportional to  $V_{in}$ , incidentally). As long as the two (identical) diodes are at the same temperature, there is good cancellation of the forward drops, except, of course, for the difference owing to input current through  $D_1$ , which produces the desired output. In this circuit,  $R_1$  should be chosen so that the current through  $D_2$  is significantly larger than the maximum input current in order to keep  $D_2$  in conduction.

In Chapter 2x we will examine better ways of constructing logarithmic converter circuits, along with careful methods of temperature compensation. With such methods it is possible to construct logarithmic converters accurate to a few percent over six decades or more of input current. A better understanding of diode and transistor characteristics, along with an understanding of op-amps, is necessary first. This section is meant to serve only as an introduction for things to come.

### 1.6.7 Inductive loads and diode protection

What happens if you open a switch that is providing current to an inductor? Because inductors have the property

$$V = L dI/dt,$$

it is not possible to turn off the current suddenly, because that would imply an infinite voltage across the inductor's terminals. What happens instead is that the voltage across the inductor rises abruptly and keeps rising until it forces current to flow. Electronic devices controlling inductive

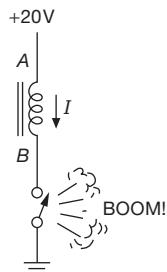


Figure 1.83. Inductive “kick.”

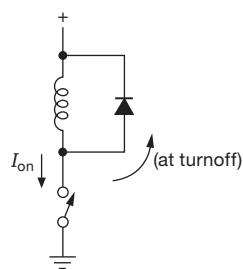
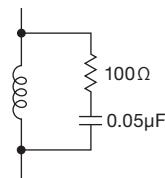


Figure 1.84. Blocking inductive kick.

loads can be easily damaged, especially the component that “breaks down” in order to satisfy the inductor’s craving for continuity of current. Consider the circuit in Figure 1.83. The switch is initially closed, and current is flowing through the inductor (which might be a relay, as described later). When the switch is opened, the inductor tries to keep current flowing from *A* to *B*, as it had been. In other words, it tries to make current flow out of *B*, which it does by forcing *B* to a high positive voltage (relative to *A*). In a case like this, in which there’s no connection to terminal *B*, it may go 1000 V positive before the switch contact “blows over.” This shortens the life of the switch and also generates impulsive interference that may affect other circuits nearby. If the switch happens to be a transistor, it would be an understatement to say that its life is shortened; its life is *ended*.

The best solution usually is to put a diode across the inductor, as in Figure 1.84. When the switch is on, the diode is back-biased (from the dc drop across the inductor’s winding resistance). At turn-off the diode goes into conduction, putting the switch terminal a diode drop above the positive supply voltage. The diode must be able to handle the initial diode current, which equals the steady current that had been flowing through the inductor; something like a 1N4004 is fine for nearly all cases.

The only disadvantage of this simple protection circuit is that it lengthens the decay of current through the inductor, because the rate of change of inductor current

Figure 1.85. *RC* “snubber” for suppressing inductive kick.

is proportional to the voltage across it. For applications in which the current must decay quickly (high-speed actuators or relays, camera shutters, magnet coils, etc.), it may be better to put a resistor across the inductor, choosing its value so that  $V_{\text{supply}} + IR$  is less than the maximum allowed voltage across the switch. For the fastest decay with a given maximum voltage, a zener with series diode (or other voltage-clamping device) can be used instead, giving a linear-like ramp-down of current rather than an exponential decay (see discussion in Chapter 1*x*).

For inductors driven from ac (transformers, ac relays), the diode protection just described will not work, because the diode will conduct on alternate half-cycles when the switch is closed. In that case a good solution is an *RC* “snubber” network (Figure 1.85). The values shown are typical for small inductive loads driven from the ac powerline. Such a snubber should be included in all instruments that run from the ac powerline, because the power transformer is inductive.<sup>33</sup>

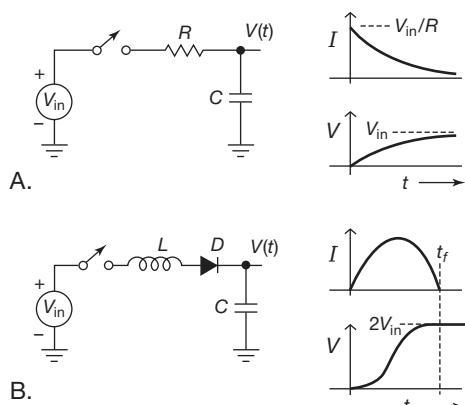
An alternative to the *RC* snubber is the use of a bidirectional zener-like voltage-clamping element. Among these the most common are the bidirectional “TVS” (transient voltage suppressor) zener and the metal-oxide varistor (“MOV”); the latter is an inexpensive device that looks something like a disc ceramic capacitor and behaves electrically like a bidirectional zener diode. Both classes are designed for transient voltage protection, are variously available at voltage ratings from 10 to 1000 volts, and can handle transient currents up to thousands of amperes (see Chapter 9*x*). Including a transient suppressor (with appropriate fusing) across the ac powerline terminals makes good sense in a piece of electronic equipment, not only to prevent inductive spike interference to other nearby instruments but also to prevent occasional large powerline spikes from damaging the instrument itself.

## 1.6.8 Interlude: inductors as friends

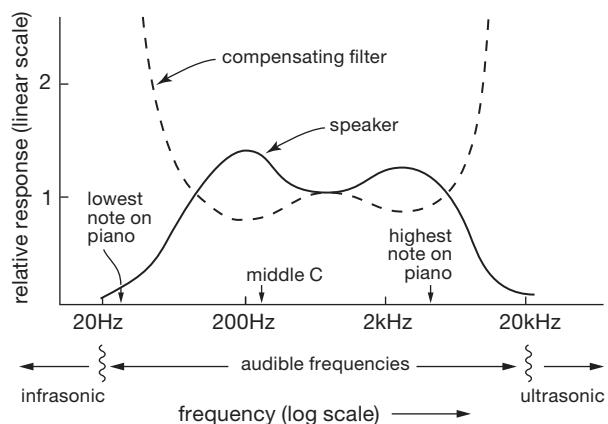
Lest we leave the impression that inductance and inductors are things only to be feared, let’s look at the circuit in

<sup>33</sup> As explained in §9.5.1, you should choose a capacitor rated for “across-the-line” service.

Figure 1.86. The goal is to charge up the capacitor from a source of dc voltage  $V_{in}$ . In the top circuit (Figure 1.86A) we've done it the conventional way, with a series resistor to limit the peak current demanded from the voltage source. OK, it does work – but it has a drawback that can be serious, namely that half the energy is lost as heat in the resistor. By contrast, in the circuit with the inductor (Figure 1.86B) no energy is lost (assuming ideal components); and, as a bonus, the capacitor gets charged to twice the input voltage. The output-voltage waveform is a sinusoidal half-cycle at the resonant frequency  $f = 1/2\pi\sqrt{LC}$ , a topic we'll see soon (§1.7.14).<sup>34,35</sup>



**Figure 1.86.** Resonant charging is lossless (with ideal components) compared with the 50% efficiency of resistive charging. Charging is complete after  $t_f$ , equal to a half-cycle of the resonant frequency. The series diode terminates the cycle, which would otherwise continue to oscillate between 0 and  $2V_{in}$ .



**Figure 1.87.** Example of frequency analysis: “boom box” loudspeaker equalization. The lowest and highest piano notes, called A0 and C8, are at 27.5 Hz and 4.2 kHz; they are four octaves below A440 and four octaves above middle C, respectively.

## 1.7 Impedance and reactance

*Warning:* this section is somewhat mathematical; you may wish to skip over the mathematics, but be sure to pay attention to the results and graphs.

Circuits with capacitors and inductors are more complicated than the resistive circuits we talked about earlier, in that their behavior depends on frequency: a “voltage divider” containing a capacitor or inductor will have a frequency-dependent division ratio. In addition, circuits containing these components (known collectively as *reactive* components) “corrupt” input waveforms such as square waves, as we saw earlier.

However, both capacitors and inductors are *linear* devices, meaning that the amplitude of the output waveform, whatever its shape, increases exactly in proportion to the input waveform’s amplitude. This linearity has many consequences, the most important of which is probably the following: *the output of a linear circuit, driven with a sinewave at some frequency f, is itself a sinewave at the same frequency (with, at most, changed amplitude and phase).*

Because of this remarkable property of circuits containing resistors, capacitors, and inductors (and, later, linear amplifiers), it is particularly convenient to analyze any such circuit by asking how the output voltage (amplitude and phase) depends on the input voltage *for sinewave input at a single frequency*, even though this may not be the intended use. A graph of the resulting *frequency response*, in which the ratio of output to input is plotted for each sinewave frequency, is useful for thinking about many kinds of

<sup>34</sup> A mechanical analogy may be helpful here. Imagine dropping packages onto a conveyor belt that is moving at speed  $v$ ; the packages are accelerated to that speed by friction, with 50% efficiency, finally reaching the belt speed  $v$ , at which speed they ride into the sunset. That’s resistive charging. Now we try something completely different, namely, we rig up a conveyor belt with little catchers attached by springs to the belt; and alongside it we have a second belt, running at twice the speed ( $2v$ ). Now when we drop a package onto the first conveyor it compresses a spring, then rebounds at  $2v$ ; and it makes a soft landing onto the second conveyor. No energy is lost (ideal springs), and the package rides off into the sunset at  $2v$ . That’s reactive charging.

<sup>35</sup> Resonant charging is used for the high-voltage supply in flashlamps and stroboscopes, with the advantages of (a) full charge between flashes (spaced no closer than  $t_f$ ), and (b) no current immediately after discharge (see waveforms), thus permitting the flashlamp to “quench” after each flash.

waveforms. As an example, a certain “boom-box” loudspeaker might have the frequency response shown in Figure 1.87, in which the “output” in this case is of course sound pressure, not voltage. It is desirable for a speaker to have a “flat” response, meaning that the graph of sound pressure versus frequency is constant over the band of audible frequencies. In this case the speaker’s deficiencies can be corrected by the introduction of a passive filter with the inverse response (as shown) within the amplifiers of the radio.

As we will see, it is possible to generalize Ohm’s law, replacing the word “resistance” with “impedance,” in order to describe any circuit containing these linear passive devices (resistors, capacitors, and inductors). You could think of the subject of impedance (generalized resistance) as Ohm’s law for circuits that include capacitors and inductors.

Some terminology: impedance (**Z**) is the “generalized resistance”; inductors and capacitors, for which the voltage and current are always  $90^\circ$  out of phase, are *reactive*; they have *reactance* (*X*). Resistors, with voltage and current always in phase, are *resistive*; they have *resistance* (*R*). In general, in a circuit that combines resistive and reactive components, the voltage and current at some place will have some in-between phase relationship, described by a complex impedance: impedance = resistance + reactance, or  $\mathbf{Z} = R + jX$  (more about this later).<sup>36</sup> However, you’ll see statements like “the impedance of the capacitor at this frequency is . . .” The reason you don’t have to use the word “reactance” in such a case is that impedance covers everything. In fact, you frequently use the word “impedance” even when you know it’s a resistance you’re talking about; you say “the source impedance” or “the output impedance” when you mean the Thévenin equivalent resistance of some source. The same holds for “input impedance.”

In all that follows, we will be talking about circuits driven by sinewaves at a single frequency. Analysis of circuits driven by complicated waveforms is more elaborate, involving the methods we used earlier (differential equations) or decomposition of the waveform into sinewaves (Fourier analysis). Fortunately, these methods are seldom necessary.

### 1.7.1 Frequency analysis of reactive circuits

Let’s start by looking at a capacitor driven by a sinewave voltage source  $V(t) = V_0 \sin \omega t$  (Figure 1.88). The current

<sup>36</sup> But, in a nutshell, the magnitude of **Z** gives the ratio of amplitudes of voltage to current, and the polar angle of **Z** gives the phase angle between current and voltage.

is

$$I(t) = C \frac{dV}{dt} = C\omega V_0 \cos \omega t,$$

i.e., a current of amplitude  $\omega CV_0$ , with its phase leading the input voltage by  $90^\circ$ . If we consider amplitudes only, and disregard phases, the current is

$$I = \frac{V}{1/\omega C}.$$

(Recall that  $\omega = 2\pi f$ .) It behaves like a frequency-dependent resistance  $R = 1/\omega C$ , but in addition the current is  $90^\circ$  out of phase with the voltage (Figure 1.89).

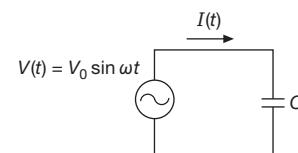


Figure 1.88. A sinusoidal ac voltage drives a capacitor.

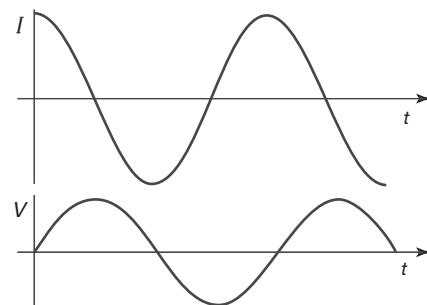


Figure 1.89. Current in a capacitor leads the sinusoidal voltage by  $90^\circ$ .

For example, a  $1 \mu\text{F}$  capacitor put across the 115 volts (rms) 60 Hz powerline draws a current of rms amplitude:

$$I = \frac{115}{1/(2\pi \times 60 \times 10^{-6})} = 43.4 \text{ mA (rms)}.$$

Soon enough we will complicate matters by explicitly worrying about *phase shifts* and the like – and that will get us into some complex algebra that terrifies beginners (often) and mathophobes (always). Before we do that, though, this is a good time to develop intuition about the frequency-dependent behavior of some basic and important circuits that use capacitors, ignoring for the time being the troublesome fact that, when driven with a sinusoidal signal, currents and voltages in a capacitor are not in phase.

As we just saw, the ratio of *magnitudes* of voltage to

current, in a capacitor driven at a frequency  $\omega$ , is just

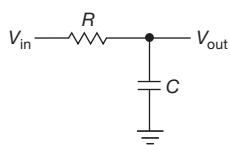
$$\frac{|V|}{|I|} = \frac{1}{\omega C},$$

which we can think of as a sort of “resistance” – the magnitude of the current is proportional to the magnitude of the applied voltage. The official name for this quantity is *reactance*, with the symbol  $X$ , thus  $X_C$  for the reactance of a capacitor.<sup>37</sup> So, for a capacitor,

$$X_C = \frac{1}{\omega C}. \quad (1.26)$$

This means that a larger capacitance has a smaller reactance. And this makes sense, because, for example, if you double the value of a capacitor, it takes twice as much current to charge and discharge it through the same voltage swing in the same time (recall  $I = CdV/dt$ ). For the same reason the reactance decreases as you increase the frequency – doubling the frequency (while holding  $V$  constant) doubles the rate of change of voltage and therefore requires that you double the current, thus half the reactance.

So, roughly speaking, we can think of a capacitor as a “frequency-dependent resistor.” Sometimes that’s good enough, sometimes it isn’t. We’ll look at a few circuits in which this simplified view gets us reasonably good results, and provides nice intuition; later we’ll fix it up, using the correct complex algebra, to get a precise result. (Keep in mind that the results we are about to get are approximate; we’re *lying* to you – but it’s a small lie, and anyway we’ll tell the truth later. In the meantime we’ll use the weird symbol  $\asymp$  instead of  $=$  in all such “approximate equations,” and we’ll flag the equation as approximate.)

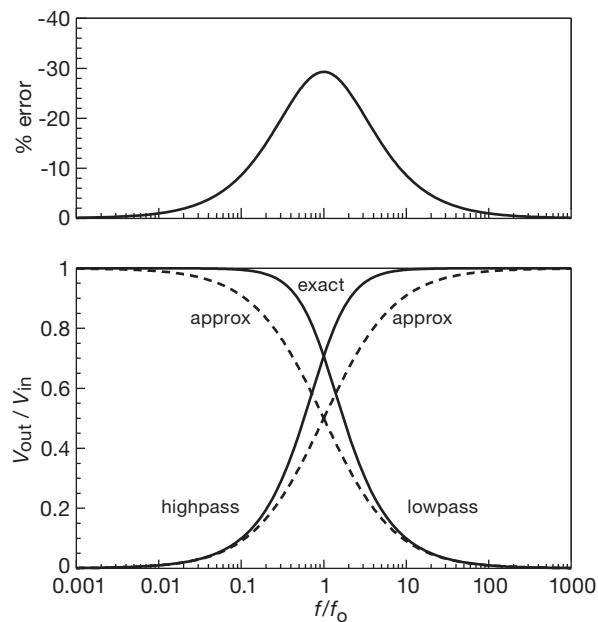


**Figure 1.90.** Lowpass filter.

#### A. RC lowpass filter (approximate)

The circuit in Figure 1.90 is called a *lowpass filter*, because it passes low frequencies and blocks high frequencies. If you think of it as a frequency-dependent voltage divider, this makes sense: the lower leg of the divider (the capacitor) has a decreasing reactance with increasing frequency, so the ratio of  $V_{\text{out}}/V_{\text{in}}$  decreases accordingly:

<sup>37</sup> Later we’ll see *inductors*, which also have a  $90^\circ$  phase shift (though of the opposite sign), and so are likewise characterized by a reactance  $X_L$ .



**Figure 1.91.** Frequency response of single-section *RC* filters, showing the results both of a simple approximation that ignores phase (dashed curve), and the exact result (solid curve). The fractional error (i.e., dashed/solid) is plotted above.

$$\frac{V_{\text{out}}}{V_{\text{in}}} \asymp \frac{X_C}{R + X_C} = \frac{1/\omega C}{R + 1/\omega C} = \frac{1}{1 + \omega RC} \quad (\text{approximate!}) \quad (1.27)$$

We’ve plotted that ratio in Figure 1.91 (and also that of its cousin, the *highpass filter*), along with their exact results that we’ll understand soon enough in §1.7.8.

You can see that the circuit passes low frequencies fully (because at low frequencies the capacitor’s reactance is very high, so it’s like a divider with a smaller resistor atop a larger one) and that it blocks high frequencies. In particular, the crossover from “pass” to “block” (often called the *breakpoint*) occurs at a frequency  $\omega_0$  at which the capacitor’s reactance ( $1/\omega_0 C$ ) is equal to the resistance  $R$ :  $\omega_0 = 1/RC$ . At frequencies well beyond the crossover (where the product  $\omega RC \gg 1$ ), the output decreases inversely with increasing frequency; that makes sense because the reactance of the capacitor, already much smaller than  $R$ , continues dropping as  $1/\omega$ . It’s worth noting that, even with our “ignorance of phase shifts,” the equation (and graph) for the ratio of voltages is quite accurate at both low and high frequencies and is only modestly in error around the crossover

frequency, where the correct ratio is  $V_{\text{out}}/V_{\text{in}} = 1/\sqrt{2} \approx 0.7$ , rather than the 0.5 that we got.<sup>38</sup>

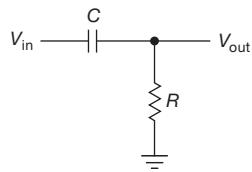


Figure 1.92. Highpass filter.

### B. RC highpass filter (approximate)

You get the reverse behavior (pass high frequencies, block low) by interchanging  $R$  and  $C$ , as in Figure 1.92. Treating it as a frequency-dependent voltage divider, and ignoring phase shifts once again, we get (see Figure 1.91)

$$\frac{V_{\text{out}}}{V_{\text{in}}} \asymp \frac{R}{R+X_C} = \frac{R}{R+1/\omega C} = \frac{\omega RC}{1+\omega RC} \quad (\text{approximate!}) \quad (1.28)$$

High frequencies (above the same crossover frequency as before,  $\omega \gg \omega_0 = 1/RC$ ) pass through (because the capacitor's reactance is much smaller than  $R$ ), whereas frequencies well below the crossover are blocked (the capacitor's reactance is much larger than  $R$ ). As before, the equation and graph are accurate at both ends, and only modestly in error at the crossover, where the correct ratio is, again,  $V_{\text{out}}/V_{\text{in}} = 1/\sqrt{2}$ .

### C. Blocking capacitor

Sometimes you want to let some band of signal frequencies pass through a circuit, but you want to block any steady dc voltage that may be present (we'll see how this can happen when we learn about amplifiers in the next chapter). You can do the job with an  $RC$  highpass filter if you choose the crossover frequency correctly: a highpass filter always blocks dc, so what you do is choose component values so that the crossover frequency is *below* all frequencies of interest. This is one of the more frequent uses of a capacitor and is known as a dc *blocking capacitor*.

For instance, every stereo audio amplifier has all its inputs capacitively coupled, because it doesn't know what dc level its input signals might be riding on. In such a coupling application you always pick  $R$  and  $C$  so that all frequencies of interest (in this case, 20 Hz to 20 kHz) are passed

without loss (attenuation). That determines the product  $RC$ :  $RC > 1/\omega_{\min}$ , where for this case you might choose  $f_{\min} \approx 5 \text{ Hz}$ , and so  $RC = 1/\omega_{\min} = 1/2\pi f_{\min} \approx 30 \text{ ms}$ .

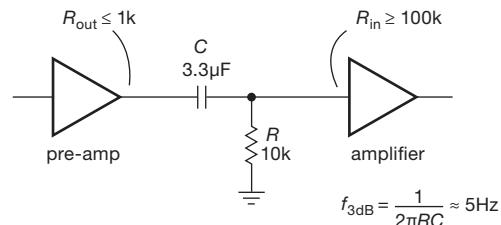


Figure 1.93. "Blocking capacitor": a highpass filter for which all signal frequencies of interest are in the passband.

You've got the product, but you still have to choose individual values for  $R$  and  $C$ . You do this by noticing that the input signal sees a load equal to  $R$  at signal frequencies (where  $C$ 's reactance is small – it's just a piece of wire there), so you choose  $R$  to be a reasonable load, i.e., not so small that it's hard to drive, and not so large that the circuit is prone to signal pickup from other circuits in the box. In the audio business it's common to see a value of  $10 \text{ k}\Omega$ , so we might choose that value, for which the corresponding  $C$  is  $3.3 \mu\text{F}$  (Figure 1.93). The circuit connected to the output should have an input resistance much greater than  $10 \text{ k}\Omega$ , to avoid loading effects on the filter's output; and the driving circuit should be able to drive a  $10 \text{ k}\Omega$  load without significant attenuation (loss of signal amplitude) to prevent circuit loading effects by the filter on the signal source. It's worth noting that our approximate model, ignoring phase shifts, is perfectly adequate for blocking capacitor design; that is because the signal band is fully in the passband, where the effects of phase shifts are negligible.

In this section we've been thinking in the frequency domain (sinewaves of frequency  $f$ ). But it's useful to think in the time domain, where, for example, you might use a blocking capacitor to couple pulses, or square waves. In such situations you encounter waveform *distortion*, in the form of "droop" and overshoot (rather than the simple amplitude attenuation and phase shift you get with sinusoidal waves). Thinking in the time domain, the criterion you use to avoid waveform distortion in a pulse of duration  $T$  is that the time constant  $\tau = RC \gg T$ . The resulting droop is approximately  $T/\tau$  (followed by a comparable overshoot at the next transition).

You often need to know the reactance of a capacitor at a given frequency (e.g., for design of filters). Figure 1.100 in §1.7.8 provides a very useful graph covering large ranges

<sup>38</sup> Of course, it fails to predict anything about phase shifts in this circuit.

As we'll see later, the output signal's phase lags the input by  $90^\circ$  at high frequencies, going smoothly from  $0^\circ$  at low frequencies, with a  $45^\circ$  lag at  $\omega_0$  (see Figure 1.104 in §1.7.9).

of capacitance and frequency, giving the value of  $X_C = 1/(2\pi fC)$ .

#### D. Driving and loading RC filters

This example of an audio-blocking capacitor raised the issue of driving and loading *RC* filter circuits. As we discussed in §1.2.5A, in the context of voltage dividers, you generally like to arrange things so that the circuit being driven does not significantly load the driving resistance (Thévenin equivalent source resistance) of the signal source.

The same logic applies here, but with a generalized kind of resistance that includes the reactance of capacitors (and inductors), known as *impedance*. So a signal source's impedance should generally be small compared with the impedance of the thing being driven.<sup>39</sup> We'll have a precise way of talking about impedance shortly, but it's correct to say that, apart from phase shifts, the impedance of a capacitor is equal to its reactance.

What we want to know, then, are the input and output impedances of the two simple *RC* filters (lowpass and highpass). This sounds complicated, because there are four impedances, and they all vary with frequency. However, if you ask the question the right way, the answer is simple, and the same in all cases!

First, assume that in each case the right thing is being done to the other end of the filter: when we ask the input impedance, we assume the output drives a high impedance (compared with its own); and when we ask the output impedance, we assume the input is driven by a signal source of low internal (Thévenin) impedance. Second, we dispose of the variation of impedances with frequency by asking only for the *worst-case* value; that is, we care what only the *maximum* output impedance of a filter circuit may be (because that is the worst for driving a load), and we care about only the *minimum* input impedance (because that is the hardest to drive).

Now the answer is astonishingly simple: in all cases the worst-case impedance is just  $R$ .

**Exercise 1.23.** Show that the preceding statement is correct.

So, for example, if you want to hang an *RC* lowpass filter onto the output of an amplifier whose output resistance is  $100\Omega$ , start with  $R = 1k$ , then choose  $C$  for the breakpoint you want. Be sure that whatever loads the output has an input impedance of at least  $10k$ . You can't go wrong.

**Exercise 1.24.** Design a two-stage “bandpass” *RC* filter, in which

the first stage is highpass with a breakpoint of  $100\text{Hz}$ , and the second stage is lowpass with a breakpoint of  $10\text{kHz}$ . Assume the input signal source has an impedance of  $100\Omega$ . What is the worst-case output impedance of your filter, and therefore what is the minimum recommended load impedance?

#### 1.7.2 Reactance of inductors

Before we embark on a fully correct treatment of impedance, replete with complex exponentials and the like, let's use our approximation tricks to figure out the reactance of an inductor.

It goes as before: we imagine an inductor  $L$  driven by a sinusoidal voltage source of angular frequency  $\omega$  such that a current  $I(t) = I_0 \sin \omega t$  is flowing.<sup>40</sup> Then the voltage across the inductor is

$$V(t) = L \frac{dI(t)}{dt} = L\omega I_0 \cos \omega t.$$

And so the ratio of *magnitudes* of voltage to current – the resistance-like quantity called *reactance* – is just

$$\frac{|V|}{|I|} = \frac{L\omega I_0}{I_0} = \omega L.$$

So, for an inductor,

$$X_L = \omega L. \quad (1.29)$$

Inductors, like capacitors, have a frequency-dependent reactance; however, here the reactance *increases* with increasing frequency (the opposite of capacitors, where it *decreases* with increasing frequency). So, in the simplest view, a series inductor can be used to pass dc and low frequencies (where its reactance is small) while blocking high frequencies (where its reactance is high). You often see inductors used this way, particularly in circuits that operate at radio frequencies; in that application they're sometimes called *chokes*.

#### 1.7.3 Voltages and currents as complex numbers

At this point it is necessary to get into some complex algebra; you may wish to skip over the math in some of the following sections, taking note of the results as we derive them. A knowledge of the detailed mathematics is not necessary for understanding the remainder of the book. Very little mathematics will be used in later chapters. The section ahead is easily the most difficult for the reader with little mathematical preparation. *Don't be discouraged!*

As we've just seen, there can be phase shifts between

<sup>39</sup> With two important exceptions, namely, transmission lines and current sources.

<sup>40</sup> We take the easy path here by specifying the *current*, rather than the voltage; we are rewarded with a simple derivative (rather than a simple integral!).

the voltage and current in an ac circuit being driven by a sinewave at some frequency. Nevertheless, as long as the circuit contains only *linear* elements (resistors, capacitors, inductors), the magnitudes of the currents everywhere in the circuit are still proportional to the magnitude of the driving voltage, so we might hope to find some generalization of voltage, current, and resistance in order to rescue Ohm's law. Evidently a single number won't suffice to specify the current, say, at some point in the circuit, because we must somehow have information about both the magnitude and phase shift.

Although we can imagine specifying the magnitudes and phase shifts of voltages and currents at any point in the circuit by writing them out explicitly, e.g.,  $V(t) = 23.7 \sin(377t + 0.38)$ , it turns out that we can meet our requirements more simply by using the algebra of complex numbers to *represent* voltages and currents. Then we can simply add or subtract the complex number representations, rather than laboriously having to add or subtract the actual sinusoidal functions of time themselves. Because the actual voltages and currents are real quantities that vary with time, we must develop a rule for converting from actual quantities to their representations, and vice versa. Recalling once again that we are talking about a single sinewave frequency,  $\omega$ , we agree to use the following rules.

1. Voltages and currents are *represented* by the complex quantities  $\mathbf{V}$  and  $\mathbf{I}$ . The voltage  $V_0 \cos(\omega t + \phi)$  is to be represented by the complex number  $V_0 e^{j\phi}$ . Recall that  $e^{j\phi} = \cos \phi + j \sin \phi$ , where  $j = \sqrt{-1}$ .
2. We obtain *actual* voltages and currents by multiplying their complex number representations by  $e^{j\omega t}$  and then taking the real part:  $V(t) = \Re e(\mathbf{V} e^{j\omega t})$ ,  $I(t) = \Re e(\mathbf{I} e^{j\omega t})$ .

In other words,

circuit voltage versus time	$\longrightarrow$	complex number representation
$V_0 \cos(\omega t + \phi)$	$\longleftarrow$	$V_0 e^{j\phi} = a + jb$
multiply by $e^{j\omega t}$ and take real part		

(In electronics, the symbol  $j$  is used instead of  $i$  in the exponential in order to avoid confusion with the symbol  $i$  meaning small-signal current.) Thus, in the general case, the actual voltages and currents are given by

$$\begin{aligned} V(t) &= \Re e(\mathbf{V} e^{j\omega t}) \\ &= \Re e(\mathbf{V}) \cos \omega t - \Im m(\mathbf{V}) \sin \omega t \end{aligned}$$

$$\begin{aligned} I(t) &= \Re e(\mathbf{I} e^{j\omega t}) \\ &= \Re e(\mathbf{I}) \cos \omega t - \Im m(\mathbf{I}) \sin \omega t. \end{aligned}$$

For example, a voltage whose complex representation is

$$\mathbf{V} = 5j$$

corresponds to a (real) voltage versus time of

$$\begin{aligned} V(t) &= \Re e[5j \cos \omega t + 5j(j) \sin \omega t] \\ &= -5 \sin \omega t \text{ volts.} \end{aligned}$$

### 1.7.4 Reactance of capacitors and inductors

With this convention we can apply complex Ohm's law correctly to circuits containing capacitors and inductors, just as for resistors, once we know the reactance of a capacitor or inductor. Let's find out what these are. We begin with a simple (co)sinusoidal voltage  $V_0 \cos \omega t$  applied across a capacitor:

$$V(t) = \Re e(V_0 e^{j\omega t}).$$

Then, using  $I = C(dV(t)/dt)$ , we obtain

$$\begin{aligned} I(t) &= -V_0 C \omega \sin \omega t = \Re e\left(\frac{V_0 e^{j\omega t}}{-j/\omega C}\right), \\ &= \Re e\left(\frac{V_0 e^{j\omega t}}{\mathbf{Z}_C}\right) \end{aligned}$$

i.e., for a capacitor

$$\mathbf{Z}_C = -j/\omega C \quad (= -jX_C);$$

$\mathbf{Z}_C$  is the impedance of a capacitor at frequency  $\omega$ ; it is equal in magnitude to the reactance  $X_C = 1/\omega C$  that we found earlier, but with a factor of  $-j$  that accounts for the  $90^\circ$  leading phase shift of current versus voltage. As an example, a  $1 \mu\text{F}$  capacitor has an impedance of  $-2653j\Omega$  at 60 Hz, and  $-0.16j\Omega$  at 1 MHz. The corresponding reactances are  $2653\Omega$  and  $0.16\Omega$ .<sup>41</sup> Its reactance (and also its impedance) at dc is infinite.

If we did a similar analysis for an inductor, we would find

$$\mathbf{Z}_L = j\omega L \quad (= jX_L).$$

A circuit containing only capacitors and inductors always has a purely imaginary impedance, meaning that the voltage and current are always  $90^\circ$  out of phase – it is purely reactive. When the circuit contains resistors, there is also

<sup>41</sup> Note the convention that the reactance  $X_C$  is a real number (the  $90^\circ$  phase shift is implicit in the term “reactance”), but the corresponding impedance is purely imaginary:  $\mathbf{Z} = R - jX$ .

a real part to the impedance. The term “reactance” in that case means the imaginary part only.

### 1.7.5 Ohm's law generalized

With these conventions for representing voltages and currents, Ohm's law takes a simple form. It reads simply

$$\mathbf{I} = \mathbf{V}/\mathbf{Z},$$

$$\mathbf{V} = \mathbf{I}\mathbf{Z},$$

where the voltage represented by  $\mathbf{V}$  is applied across a circuit of impedance  $\mathbf{Z}$ , giving a current represented by  $\mathbf{I}$ . The complex impedance of devices in series or parallel obeys the same rules as resistance:

$$\mathbf{Z} = \mathbf{Z}_1 + \mathbf{Z}_2 + \mathbf{Z}_3 + \dots \quad (\text{series}), \quad (1.30)$$

$$\mathbf{Z} = \frac{1}{\frac{1}{\mathbf{Z}_1} + \frac{1}{\mathbf{Z}_2} + \frac{1}{\mathbf{Z}_3} + \dots} \quad (\text{parallel}). \quad (1.31)$$

Finally, for completeness we summarize here the formulas for the impedance of resistors, capacitors, and inductors:

$$\mathbf{Z}_R = R \quad (\text{resistor}),$$

$$\mathbf{Z}_C = -j/\omega C = 1/j\omega C \quad (\text{capacitor}), \quad (1.32)$$

$$\mathbf{Z}_L = j\omega L \quad (\text{inductor}).$$

With these rules we can analyze many ac circuits by the same general methods we used in handling dc circuits, i.e., application of the series and parallel formulas and Ohm's law. Our results for circuits such as voltage dividers will look nearly the same as before. For multiply-connected networks we may have to use Kirchhoff's laws, just as with dc circuits, in this case using the complex representations for  $V$  and  $I$ : the sum of the (complex) voltage drops around a closed loop is zero, and the sum of the (complex) currents into a point is zero. The latter rule implies, as with dc circuits, that the (complex) current in a series circuit is the same everywhere.

**Exercise 1.25.** Use the preceding rules for the impedance of devices in parallel and in series to derive the formulas (1.17) and (1.18) for the capacitance of two capacitors (a) in parallel and (b) in series. Hint: in each case, let the individual capacitors have capacitances  $C_1$  and  $C_2$ . Write down the impedance of the parallel or series combination; then equate it to the impedance of a capacitor with capacitance  $C$ . Then find  $C$ .

Let's try out these techniques on the simplest circuit imaginable, an ac voltage applied across a capacitor, which we looked at earlier, in §1.7.1. Then, after a brief look at power in reactive circuits (to finish laying the groundwork),

we'll analyze (correctly, this time) the simple but extremely important and useful  $RC$  lowpass and highpass filter circuits.

Imagine putting a  $1 \mu\text{F}$  capacitor across a 115 volts (rms) 60 Hz powerline. What current flows? Using complex Ohm's law, we have

$$\mathbf{Z} = -j/\omega C.$$

Therefore, the current is given by

$$\mathbf{I} = \mathbf{V}/\mathbf{Z}.$$

The phase of the voltage is arbitrary, so let us choose  $\mathbf{V} = A$ , i.e.,  $V(t) = A \cos \omega t$ , where the amplitude  $A = 115\sqrt{2} \approx 163$  volts. Then

$$\mathbf{I} = j\omega CA \approx -0.061 \sin \omega t.$$

The resulting current has an amplitude of 61 mA (43 mA rms) and leads the voltage by  $90^\circ$ . This agrees with our previous calculation. More simply, we could have noticed that the impedance of the capacitor is negative imaginary, so whatever the absolute phase of  $V$ , the phase of  $I_{\text{cap}}$  must lead by  $90^\circ$ . And in general the phase angle between current and voltage, for any two-terminal  $RLC$  circuit, is equal to the angle of the (complex) impedance of that circuit.

Note that if we wanted to know just the magnitude of the current, and didn't care what the relative phase was, we could have avoided doing any complex algebra: if

$$\mathbf{A} = \mathbf{B}/\mathbf{C},$$

then

$$A = B/C,$$

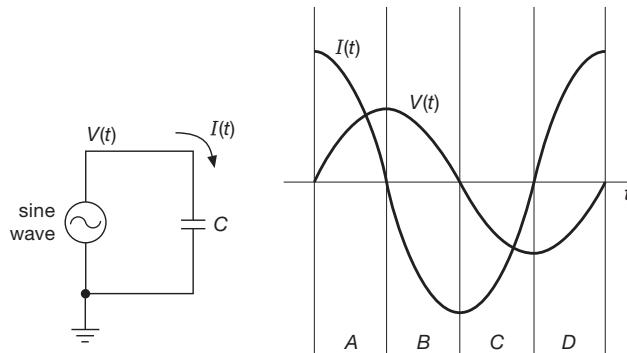
where  $A$ ,  $B$ , and  $C$  are the magnitudes of the respective complex numbers; this holds for multiplication, also (see Exercise 1.18). Thus, in this case,

$$I = V/Z = \omega CV.$$

This trick, which we used earlier (because we didn't know any better), is often useful.

Surprisingly, there is no power dissipated by the capacitor in this example. Such activity won't increase your electric bill; you'll see why in the next section. Then we will go on to look at circuits containing resistors and capacitors with our complex Ohm's law.

**Exercise 1.26.** Show that, if  $\mathbf{A} = \mathbf{BC}$ , then  $A = BC$ , where  $A$ ,  $B$ , and  $C$  are magnitudes. Hint: represent each complex number in polar form, i.e.,  $\mathbf{A} = Ae^{j\theta}$ .



**Figure 1.94.** The power delivered to a capacitor is zero over a full sinusoidal cycle, owing to the  $90^\circ$  phase shift between voltage and current.

### 1.7.6 Power in reactive circuits

The instantaneous power delivered to any circuit element is always given by the product  $P = VI$ . However, in reactive circuits where  $V$  and  $I$  are not simply proportional, you can't just multiply their amplitudes together. Funny things can happen; for instance, the sign of the product can reverse over one cycle of the ac signal. Figure 1.94 shows an example. During time intervals A and C, power is being delivered to the capacitor (albeit at a variable rate), causing it to charge up; its stored energy is increasing (power is the rate of change of energy). During intervals B and D, the power delivered to the capacitor is negative; it is discharging. The average power over a whole cycle for this example is in fact exactly zero, a statement that is always true for any purely reactive circuit element (inductors, capacitors, or any combination thereof). If you know your trigonometric integrals, the next exercise will show you how to prove this.

**Exercise 1.27.** Optional exercise: prove that a circuit whose current is  $90^\circ$  out of phase with the driving voltage consumes no power, averaged over an entire cycle.

How do we find the average power consumed by an arbitrary circuit? In general, we can imagine adding up little pieces of  $VI$  product, then dividing by the elapsed time. In other words,

$$P = \frac{1}{T} \int_0^T V(t)I(t) dt, \quad (1.33)$$

where  $T$  is the time for one complete cycle. Luckily, that's almost never necessary. Instead, it is easy to show that the average power is given by

$$P = \Re(\mathbf{V}\mathbf{I}^*) = \Re(\mathbf{V}^*\mathbf{I}), \quad (1.34)$$

where  $\mathbf{V}$  and  $\mathbf{I}$  are complex rms amplitudes (and an asterisk means *complex conjugate* – see the math review, Appendix A, if this is unfamiliar).

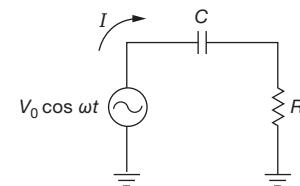
Let's take an example. Consider the preceding circuit, with a 1 volt (rms) sinewave driving a capacitor. We'll do everything with rms amplitudes, for simplicity. We have

$$\mathbf{V} = 1,$$

$$\mathbf{I} = \frac{\mathbf{V}}{-j/\omega C} = j\omega C,$$

$$P = \Re(\mathbf{V}\mathbf{I}^*) = \Re(-j\omega C) = 0.$$

That is, the average power is zero, as stated earlier.



**Figure 1.95.** Power and power factor in a series  $RC$  circuit.

As another example, consider the circuit shown in Figure 1.95. Our calculations go like this:

$$\mathbf{Z} = R - \frac{j}{\omega C},$$

$$\mathbf{V} = V_0,$$

$$\mathbf{I} = \frac{\mathbf{V}}{\mathbf{Z}} = \frac{V_0}{R - (j/\omega C)} = \frac{V_0[R + (j/\omega C)]}{R^2 + (1/\omega^2 C^2)},$$

$$P = \Re(\mathbf{V}\mathbf{I}^*) = \frac{V_0^2 R}{R^2 + (1/\omega^2 C^2)}.$$

(In the third line we multiplied numerator and denominator by the complex conjugate of the denominator in order to make the denominator real.) The calculated power<sup>42</sup> is less than the product of the magnitudes of  $\mathbf{V}$  and  $\mathbf{I}$ . In fact, their ratio is called the *power factor*:

$$|\mathbf{V}||\mathbf{I}| = \frac{V_0^2}{[R^2 + (1/\omega^2 C^2)]^{1/2}},$$

$$\text{power factor} = \frac{\text{power}}{|\mathbf{V}||\mathbf{I}|}$$

$$= \frac{R}{[R^2 + (1/\omega^2 C^2)]^{1/2}}$$

<sup>42</sup> It's always a good idea to check limiting values: here we see that  $P \rightarrow V^2/R$  for large  $C$ ; and for small  $C$  the magnitude of the current  $|I| \rightarrow V_0/X_C$ , or  $V_0\omega C$ , thus  $P \rightarrow I^2R = V_0^2\omega^2C^2R$ , in agreement at both limits.

in this case. The power factor is the cosine of the phase angle between the voltage and the current, and it ranges from 0 (purely reactive circuit) to 1 (purely resistive). A power factor of less than 1 indicates some component of reactive current.<sup>43</sup> It's worth noting that the power factor goes to unity, and the dissipated power goes to  $V^2/R$ , in the limit of large capacitance (or of high frequency), where the reactance of the capacitor becomes much less than  $R$ .

**Exercise 1.28.** Show that all the average power delivered to the preceding circuit winds up in the resistor. Do this by computing the value of  $V_R^2/R$ . What is that power, in watts, for a series circuit of a  $1\ \mu\text{F}$  capacitor and a  $1.0\text{k}$  resistor placed across the 115 volt (rms), 60 Hz powerline?

Power factor is a serious matter in large-scale electrical power distribution, because reactive currents don't result in useful power being delivered to the load, but cost the power company plenty in terms of  $I^2R$  heating in the resistance of generators, transformers, and wiring. Although residential users are billed only for "real" power [ $\Re(\mathbf{V}\mathbf{I}^*)$ ], the power company charges industrial users according to the power factor. This explains the capacitor yards that you see behind large factories, built to cancel the inductive reactance of industrial machinery (i.e., motors).

**Exercise 1.29.** Show that adding a series capacitor of value  $C = 1/\omega^2L$  makes the power factor equal 1.0 in a series  $RL$  circuit. Now do the same thing, but with the word "series" changed to "parallel."

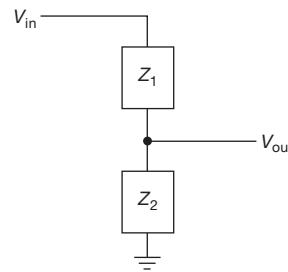
### 1.7.7 Voltage dividers generalized

Our original voltage divider (Figure 1.6) consisted of a pair of resistors in series to ground, input at the top and output at the junction. The generalization of that simple resistive divider is a similar circuit in which either or both resistors are replaced with a capacitor or inductor (or a more complicated network made from  $R$ ,  $L$ , and  $C$ ), as in Figure 1.96. In general, the division ratio  $V_{\text{out}}/V_{\text{in}}$  of such a divider is not constant, but depends on frequency (as we have already seen, in our approximate treatment of the lowpass and highpass filters in §1.7.1). The analysis is straightforward:

$$\mathbf{I} = \frac{\mathbf{V}_{\text{in}}}{\mathbf{Z}_{\text{total}}},$$

$$\mathbf{Z}_{\text{total}} = \mathbf{Z}_1 + \mathbf{Z}_2$$

<sup>43</sup> Or, for nonlinear circuits, it indicates that the current waveform is not proportional to the voltage waveform. More on this in §9.7.1.



**Figure 1.96.** Generalized voltage divider: a pair of arbitrary impedances.

$$\mathbf{V}_{\text{out}} = \mathbf{I}\mathbf{Z}_2 = \mathbf{V}_{\text{in}} \frac{\mathbf{Z}_2}{\mathbf{Z}_1 + \mathbf{Z}_2}.$$

Rather than worrying about this result in general, let's look at some simple (but very important) examples, beginning with the  $RC$  highpass and lowpass filters we approximated earlier.

### 1.7.8 RC highpass filters

We've seen that by combining resistors with capacitors it is possible to make frequency-dependent voltage dividers, owing to the frequency dependence of a capacitor's impedance  $\mathbf{Z}_C = -j/\omega C$ . Such circuits can have the desirable property of passing signal frequencies of interest while rejecting undesired signal frequencies. In this subsection and the next we revisit the simple lowpass and highpass  $RC$  filters, correcting the approximate analysis of §1.7.1; though simple, these circuits are important and widely used. Chapter 6 and Appendix E describe filters of greater sophistication.

Referring back to the classic  $RC$  highpass filter (Figure 1.92), we see that the complex Ohm's law (or the complex voltage-divider equation) gives

$$\mathbf{V}_{\text{out}} = \mathbf{V}_{\text{in}} \frac{R}{R - j/\omega C} = \mathbf{V}_{\text{in}} \frac{R(R + j/\omega C)}{R^2 + (1/\omega^2 C^2)}.$$

(For the last step, multiply top and bottom by the complex conjugate of the denominator.) Most often we don't care about the phase of  $V_{\text{out}}$ , just its amplitude:

$$\begin{aligned} V_{\text{out}} &= (\mathbf{V}_{\text{out}} \mathbf{V}_{\text{out}}^*)^{1/2} \\ &= \frac{R}{[R^2 + (1/\omega^2 C^2)]^{1/2}} V_{\text{in}}. \end{aligned}$$

Note the analogy to a resistive divider, where

$$V_{\text{out}} = \frac{R_2}{R_1 + R_2} V_{\text{in}}.$$

Here the impedance of the series  $RC$  combination

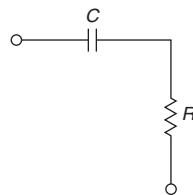


Figure 1.97. Input impedance of unloaded highpass filter.

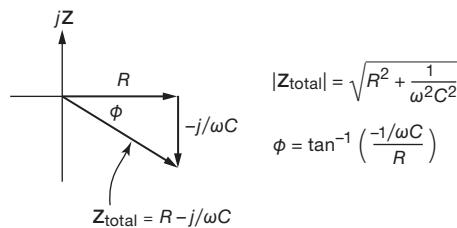


Figure 1.98. Impedance of series  $RC$ .

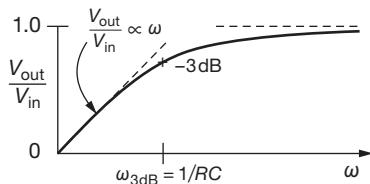


Figure 1.99. Frequency response of highpass filter. The corresponding phase shift goes smoothly from  $+90^\circ$  (at  $\omega = 0$ ), through  $+45^\circ$  (at  $\omega_{3dB}$ ), to  $0^\circ$  (at  $\omega = \infty$ ), analogous to the lowpass filter's phase shift (Figure 1.104).

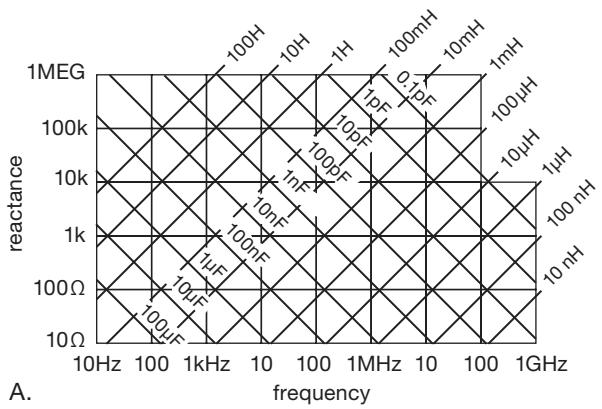
(Figure 1.97) is as shown in Figure 1.98. So the “response” of this circuit, ignoring phase shifts by taking magnitudes of the complex amplitudes, is given by

$$\begin{aligned} V_{\text{out}} &= \frac{R}{[R^2 + (1/\omega^2 C^2)]^{1/2}} V_{\text{in}} \\ &= \frac{2\pi f RC}{[1 + (2\pi f RC)^2]^{1/2}} V_{\text{in}} \end{aligned} \quad (1.35)$$

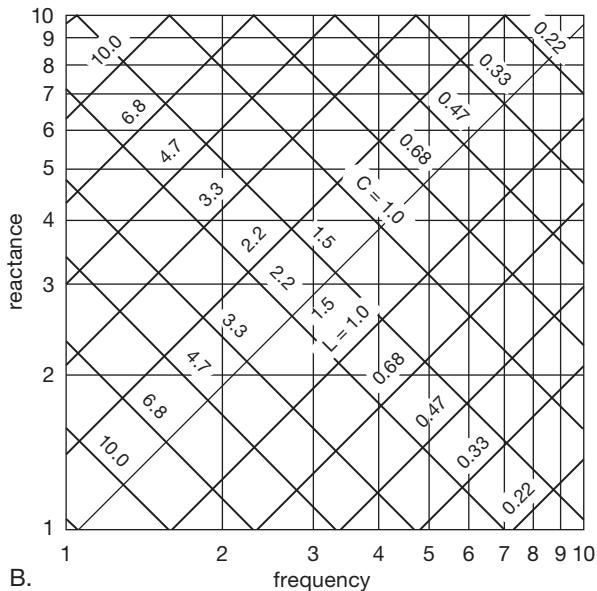
and looks as shown in Figure 1.99 (and earlier in Figure 1.91).

Note that we could have gotten this result immediately by taking the ratio of the *magnitudes* of impedances, as in Exercise 1.26 and the example immediately preceding it; the numerator is the magnitude of the impedance of the lower leg of the divider ( $R$ ), and the denominator is the magnitude of the impedance of the series combination of  $R$  and  $C$ .

As we noted earlier, the output is approximately equal to the input at high frequencies (how high?  $\omega \gtrsim 1/RC$ ) and goes to zero at low frequencies. The highpass filter is



A.



B.

Figure 1.100. A: Reactance of inductors and capacitors versus frequency; all decades are identical, except for scale. B: A single decade from part A expanded, with standard 20% component values (EIA “E6”) shown.

very common; for instance, the input to the oscilloscope can be switched to “ac coupling.” That’s just an  $RC$  highpass filter with the bend at about 10 Hz (you would use ac coupling if you wanted to look at a small signal riding on a large dc voltage). Engineers like to refer to the  $-3\text{ dB}$  “breakpoint” of a filter (or of any circuit that behaves like a filter). In the case of the simple  $RC$  high-pass filter, the  $-3\text{ dB}$  breakpoint is given by

$$f_{3\text{dB}} = 1/2\pi RC.$$

You often need to know the impedance of a capacitor at a given frequency (e.g., for the design of filters).

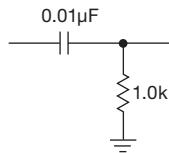


Figure 1.101. Highpass filter example.

Figure 1.100 provides a very useful graph covering large ranges of capacitance and frequency, giving the value of  $|Z| = 1/2\pi fC$ .

As an example, consider the filter shown in Figure 1.101. It is a highpass filter with the 3 dB point<sup>44</sup> at 15.9 kHz. The impedance of a load driven by it should be much larger than 1.0k in order to prevent circuit loading effects on the filter's output, and the driving source should be able to drive a 1.0k load without significant attenuation (loss of signal amplitude) in order to prevent circuit loading effects by the filter on the signal source (recall §1.7.1D for worst-case source and load impedances of  $RC$  filters).

### 1.7.9 RC lowpass filters

Revisiting the lowpass filter, in which you get the opposite frequency behavior by interchanging  $R$  and  $C$  (Figure 1.90, repeated here as Figure 1.102), we find the accurate result

$$V_{\text{out}} = \frac{1}{(1 + \omega^2 R^2 C^2)^{1/2}} V_{\text{in}} \quad (1.36)$$

as seen in Figure 1.103 (and earlier in Figure 1.91). The 3 dB point is again at a frequency<sup>45</sup>  $f = 1/2\pi RC$ . Lowpass filters are quite handy in real life. For instance, a lowpass filter can be used to eliminate interference from nearby radio and television stations (0.5–800 MHz), a problem that plagues audio amplifiers and other sensitive electronic equipment.

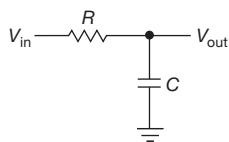


Figure 1.102. Lowpass filter.

**Exercise 1.30.** Show that the preceding expression for the response of an  $RC$  lowpass filter is correct.

<sup>44</sup> One often omits the minus sign when referring to the  $-3$  dB point.

<sup>45</sup> As mentioned in §1.7.1A, we often like to define the breakpoint frequency  $\omega_0 = 1/RC$ , and work with frequency ratios  $\omega/\omega_0$ . Then a useful form for the denominator in eq'n 1.36 is  $\sqrt{1 + (\omega/\omega_0)^2}$ . The same applies to eq'n 1.35, where the numerator becomes  $\omega/\omega_0$ .

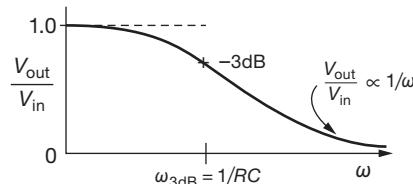


Figure 1.103. Frequency response of lowpass filter.

The lowpass filter's output can be viewed as a signal source in its own right. When driven by a perfect ac voltage (zero source impedance), the filter's output looks like  $R$  at low frequencies (the perfect signal source can be replaced with a short, i.e., by its small-signal source impedance, for the purpose of impedance calculations). It drops to zero impedance at high frequencies, where the capacitor dominates the output impedance. The signal driving the filter sees a load of  $R$  plus the load resistance at low frequencies, dropping to just  $R$  at high frequencies. As we remarked in §1.7.1D, the worst-case source impedance and the worst-case load impedance of an  $RC$  filter (lowpass or highpass) are both equal to  $R$ .

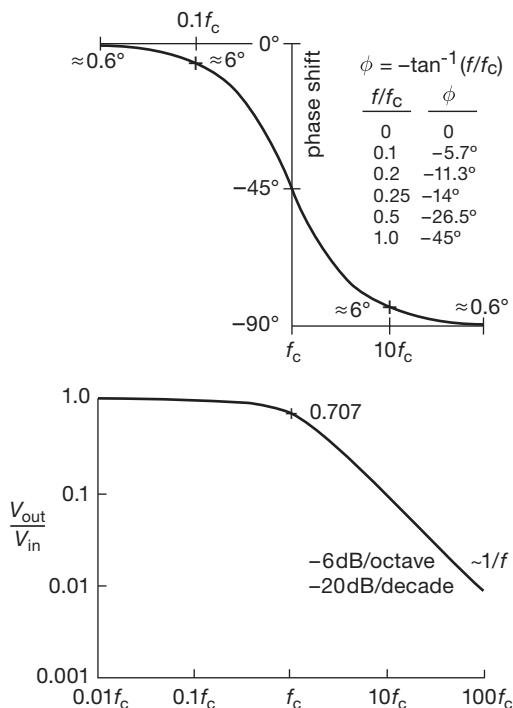


Figure 1.104. Frequency response (phase and amplitude) of lowpass filter plotted on logarithmic axes. Note that the phase shift is  $-45^\circ$  at the  $-3$  dB point and is within  $6^\circ$  of its asymptotic value for a decade of frequency change.

In Figure 1.104, we've plotted the same lowpass filter response with logarithmic axes, which is a more common way that it's done. You can think of the vertical axis as decibels, and the horizontal axis as octaves (or decades). On such a plot, equal distances correspond to equal ratios. We've also plotted the phase shift, using a linear vertical axis (degrees) and the same logarithmic frequency axis. This sort of plot is good for seeing the detailed response even when it is greatly attenuated (as at right); we'll see a number of such plots in Chapter 6, when we treat active filters. Note that the filter curve plotted here becomes a straight line at large attenuations, with a slope of  $-20 \text{ dB/decade}$  (engineers prefer to say " $-6 \text{ dB/octave}$ "). Note also that the phase shift goes smoothly from  $0^\circ$  (at frequencies well below the breakpoint) to  $-90^\circ$  (well above it), with a value of  $-45^\circ$  at the  $-3 \text{ dB}$  point. A rule of thumb for single-section  $RC$  filters is that the phase shift is  $\approx 6^\circ$  from its asymptotic value at  $0.1f_{3\text{dB}}$  and at  $10f_{3\text{dB}}$ .

**Exercise 1.31.** Prove the last assertion.

An interesting question is the following: is it possible to make a filter with some arbitrary specified amplitude response and some other arbitrary specified phase response? Surprisingly, the answer is no: the demands of causality (i.e., that response must follow cause, not precede it) force a relationship between phase and amplitude response of realizable analog filters (known officially as the Kramers–Kronig relation).

### 1.7.10 $RC$ differentiators and integrators in the frequency domain

The  $RC$  differentiator that we saw in §1.4.3 is exactly the same circuit as the highpass filter in this section. In fact, it can be considered as either, depending on whether you're thinking of waveforms in the time domain or response in the frequency domain. We can restate the earlier time-domain condition for its proper operation ( $dV_{\text{out}} \ll dV_{\text{in}}$ ) in terms of the frequency response: for the output to be small compared with the input, the signal frequency (or frequencies) must be well below the  $3 \text{ dB}$  point. This is easy to check: suppose we have the input signal  $V_{\text{in}} = \sin \omega t$ . Then, using the equation we obtained earlier for the differentiator output, we have

$$V_{\text{out}} = RC \frac{d}{dt} \sin \omega t = \omega RC \cos \omega t,$$

and so  $dV_{\text{out}} \ll dV_{\text{in}}$  if  $\omega RC \ll 1$ , i.e.,  $RC \ll 1/\omega$ . If the input signal contains a range of frequencies, this must hold for the highest frequencies present in the input.

The  $RC$  integrator (§1.4.4) is the same circuit as the low-pass filter; by similar reasoning, the criterion for a good integrator is that the lowest signal frequencies must be well above the  $3 \text{ dB}$  point.

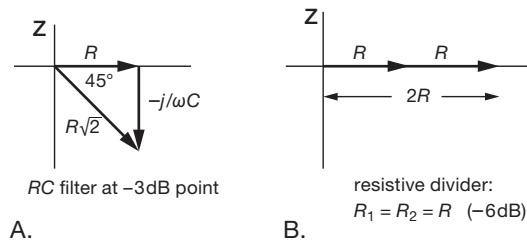
### 1.7.11 Inductors versus capacitors

Instead of capacitors, inductors can be combined with resistors to make lowpass (or highpass) filters. In practice, however, you rarely see  $RL$  lowpass or highpass filters. The reason is that inductors tend to be more bulky and expensive and perform less well (i.e., they depart further from the ideal) than capacitors (see Chapter 1x). If you have a choice, use a capacitor. One important exception to this general statement is the use of ferrite beads and chokes in high-frequency circuits. You just string a few beads here and there in the circuit; they make the wire interconnections slightly inductive, raising the impedance at very high frequencies and preventing oscillations, without the added series resistance you would get with an  $RC$  filter. An *RF choke* is an inductor, usually a few turns of wire wound on a ferrite core, used for the same purpose in RF circuits. Note, however, that inductors are *essential* components in (a)  $LC$  tuned circuits (§1.7.14), and (b) switch-mode power converters (§9.6.4).

### 1.7.12 Phasor diagrams

There's a nice graphical method that can be helpful when we are trying to understand reactive circuits. Let's take an example, namely, the fact that an  $RC$  filter attenuates  $3 \text{ dB}$  at a frequency  $f = 1/2\pi RC$ , which we derived in §1.7.8. This is true for both highpass and lowpass filters. It is easy to get a bit confused here, because at that frequency the reactance of the capacitor equals the resistance of the resistor; so you might at first expect  $6 \text{ dB}$  attenuation (a factor of  $1/2$  in voltage). That is what you would get, for example, if you were to replace the capacitor with a resistor of the same impedance magnitude. The confusion arises because the capacitor is reactive, but the matter is clarified by a phasor diagram (Figure 1.105). The axes are the real (resistive) and imaginary (reactive) components of the impedance. In a series circuit like this, the axes also represent the (complex) voltage, because the current is the same everywhere. So for this circuit (think of it as an  $RC$  voltage divider) the input voltage (applied across the series  $RC$  pair) is proportional to the length of the hypotenuse, and the output voltage (across  $R$  only) is proportional to the length of the  $R$  leg of the triangle. The diagram represents the situation at the frequency where the capacitor's reactance equals  $R$ ,

i.e.,  $f = 1/2\pi RC$ , and shows that the ratio of output voltage to input voltage is  $1/\sqrt{2}$ , i.e.,  $-3$  dB.



**Figure 1.105.** Phasor diagram for lowpass filter at 3 dB point.

The angle between the vectors gives the phase shift from input to output. At the 3 dB point, for instance, the output amplitude equals the input amplitude divided by the square root of 2, and it leads by  $45^\circ$  in phase. This graphical method makes it easy to read off amplitude and phase relationships in  $RLC$  circuits. You can use it, for example, to get the response of the highpass filter that we previously derived algebraically.

**Exercise 1.32.** Use a phasor diagram to derive the response of an  $RC$  high-pass filter:  $V_{out} = V_{in}R / \sqrt{R^2 + (1/\omega^2 C^2)}$ .

**Exercise 1.33.** At what frequency does an  $RC$  lowpass filter attenuate by 6 dB (output voltage equal to half the input voltage)? What is the phase shift at that frequency?

**Exercise 1.34.** Use a phasor diagram to obtain the lowpass filter response previously derived algebraically.

In the next chapter (§2.2.8) we'll see a nice example of phasor diagrams in connection with a constant-amplitude phase-shifting circuit.

### 1.7.13 “Poles” and decibels per octave

Look again at the response of the  $RC$  lowpass filter (Figures 1.103 and 1.104). Far to the right of the “knee” the output amplitude is dropping proportional to  $1/f$ . In one octave (as in music, one octave is twice the frequency) the output amplitude will drop to half, or  $-6$  dB; so a simple  $RC$  filter has a 6 dB/octave falloff. You can make filters with several  $RC$  sections; then you get 12 dB/octave (two  $RC$  sections), 18 dB/octave (three sections), and so on. This is the usual way of describing how a filter behaves beyond the cutoff. Another popular way is to say a “three-pole filter,” for instance, meaning a filter with three  $RC$  sections (or one that behaves like one). (The word “pole” derives from a method of analysis that is beyond the scope of this book and that involves complex

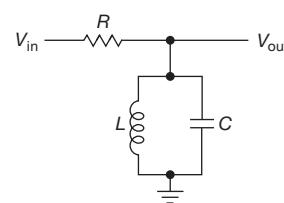
transfer functions in the complex frequency plane, known by engineers as the “ $s$ -plane.” This is discussed in the advanced volume, in Chapter 1x.)

A caution on multistage filters: you can't simply cascade several identical filter sections in order to get a frequency response that is the concatenation of the individual responses. The reason is that each stage will load the previous one significantly (since they're identical), changing the overall response. Remember that the response function we derived for the simple  $RC$  filters was based on a zero-impedance driving source and an infinite-impedance load. One solution is to make each successive filter section have much higher impedance than the preceding one. A better solution involves active circuits like transistor or operational amplifier (op-amp) interstage “buffers,” or active filters. These subjects will be treated in Chapters 2–4, 6, and 13.

### 1.7.14 Resonant circuits

When capacitors are combined with inductors or are used in special circuits called active filters, it is possible to make circuits that have very sharp frequency characteristics (e.g., a large peak in the response at a particular frequency), compared with the gradual characteristics of the  $RC$  filters we've seen so far. These circuits find applications in various audio and RF devices. Let's now take a quick look at  $LC$  circuits (there will be more on them, and active filters, in Chapter 6 and in Appendix E).

#### A. Parallel and series $LC$ circuits



**Figure 1.106.**  $LC$  resonant circuit: bandpass filter.

First, consider the circuit shown in Figure 1.106. The impedance of the  $LC$  combination at frequency  $f$  is just

$$\begin{aligned}\frac{1}{Z_{LC}} &= \frac{1}{Z_L} + \frac{1}{Z_C} = \frac{1}{j\omega L} - \frac{\omega C}{j} \\ &= j \left( \omega C - \frac{1}{\omega L} \right),\end{aligned}$$

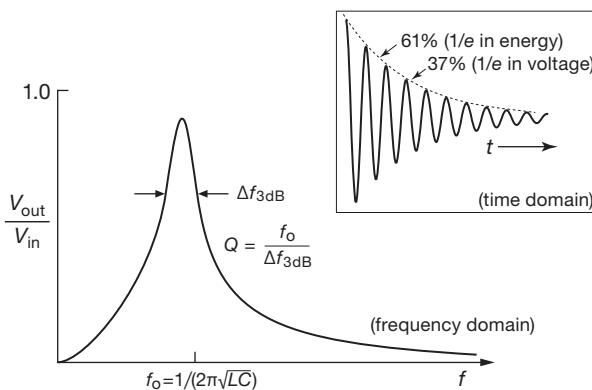
i.e.,

$$Z_{LC} = \frac{j}{(1/\omega L) - \omega C}.$$

In combination with  $R$  it forms a voltage divider. Because of the opposite behaviors of inductors and capacitors, the impedance of the parallel  $LC$  goes to infinity at the *resonant frequency*

$$f_0 = 1/2\pi\sqrt{LC} \quad (1.37)$$

(i.e.,  $\omega_0 = 1/\sqrt{LC}$ ), giving a peak in the response there. The overall response is as shown in Figure 1.107.



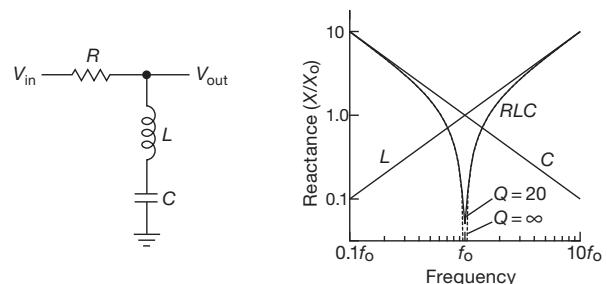
**Figure 1.107.** Frequency response of parallel  $LC$  “tank” circuit. The inset shows the time-domain behavior: a damped oscillation (“ringing”) waveform following an input voltage step or pulse.

In practice, losses in the inductor and capacitor limit the sharpness of the peak, but with good design these losses can be made very small. Conversely, a *Q*-spoiling resistor is sometimes added intentionally to reduce the sharpness of the resonant peak. This circuit is known simply as a parallel  $LC$  resonant circuit (or “tuned circuit,” or “tank”) and is used extensively in RF circuits to select a particular frequency for amplification (the  $L$  or  $C$  can be variable, so you can tune the resonant frequency). The higher the driving impedance, the sharper the peak; it is not uncommon to drive them with something approaching a current source, as you will see later. The *quality factor*  $Q$  is a measure of the sharpness of the peak. It equals the resonant frequency divided by the width at the  $-3\text{ dB}$  points. For a parallel  $RLC$  circuit,  $Q = \omega_0 RC$ .<sup>46</sup>

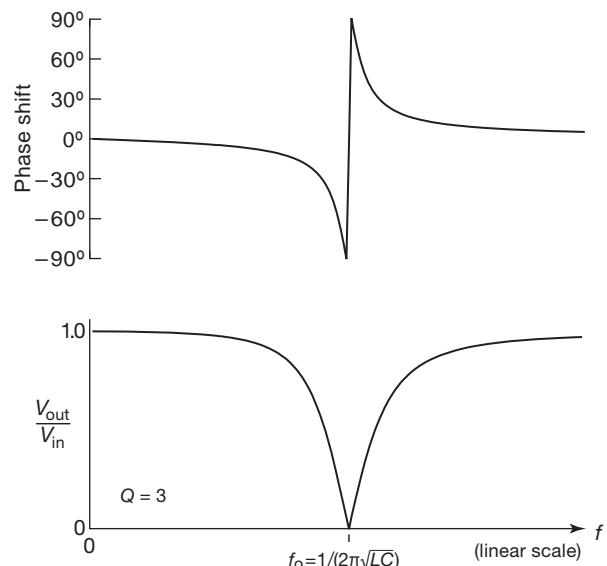
Another variety of  $LC$  circuit is the series  $LC$  (Figure 1.108). By writing down the impedance formulas involved, and assuming that both the capacitor and inductor are ideal, i.e., that they have no resistive losses,<sup>47</sup> you can convince yourself that the impedance of the  $LC$  goes to zero

<sup>46</sup> Or, equivalently,  $Q = R/X_C = R/X_L$ , where  $X_L = X_C$  are the reactances at  $\omega_0$ .

<sup>47</sup> We’ll see in Chapter 1x that real components depart from the ideal, often expressed in terms of an effective series resistance, ESR.



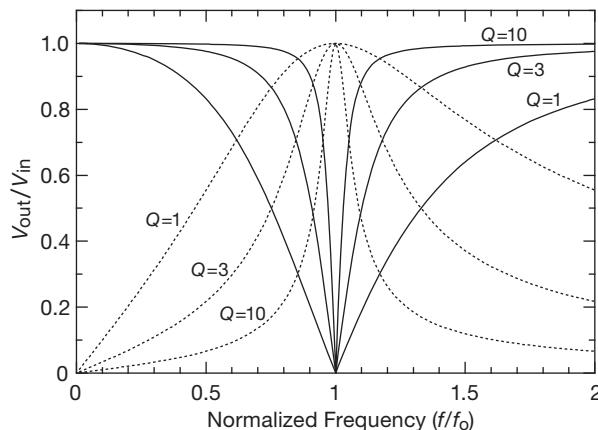
**Figure 1.108.**  $LC$  notch filter (“trap”). The inductive and capacitive reactances behave as shown, but the opposite sign of their complex impedances causes the series impedance to plummet. For ideal components the reactance of the series  $LC$  goes completely to zero at resonance; for real-world components the minimum is non-zero, and usually dominated by the inductor.



**Figure 1.109.** Frequency and phase response of the series  $LC$  trap. The phase changes abruptly at resonance, an effect seen in other resonator types (see for example Figure 7.36).

at resonance ( $f_0 = 1/2\pi\sqrt{LC}$ ). Such a circuit is a “trap” for signals at or near the resonant frequency, shorting them to ground. Again, this circuit finds application mainly in RF circuits. Figure 1.109 shows what the response looks like. The  $Q$  of a series  $RLC$  circuit is  $Q = \omega_0 L/R$ .<sup>48</sup> To see the impact of increasing  $Q$ , look at the accurate plots of tank and notch response in Figure 1.110.

<sup>48</sup> Or, equivalently,  $Q = X_L/R = X_C/R$ , where  $X_L = X_C$  are the reactances at  $\omega_0$ .



**Figure 1.110.** Response of *LC* tank (dotted curves) and trap (solid curves) for several values of quality factor,  $Q$ .

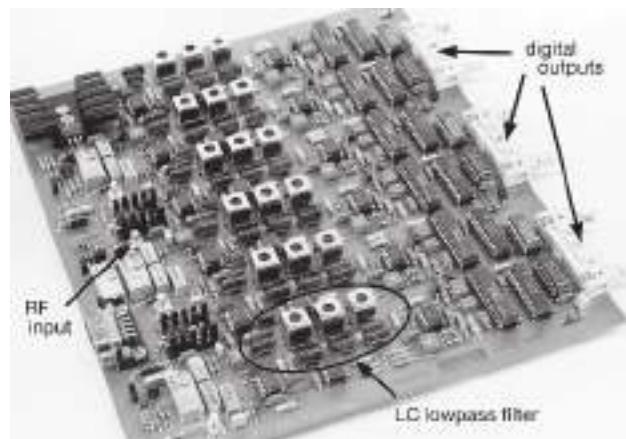
**Exercise 1.35.** Find the response ( $V_{\text{out}}/V_{\text{in}}$  versus frequency) for the series *LC* trap circuit in Figure 1.108.

These descriptions of *LC* resonant circuits are phrased in terms of frequency response, i.e., in the frequency domain. In the time domain you're generally interested in a circuit's response to pulses, or steps; there you see the sort of behavior shown in the inset of Figure 1.107, an *LC* circuit with  $Q=20$ . The signal voltage falls to  $1/e$  (37%) in  $Q/\pi$  cycles; the stored energy (proportional to  $v^2$ ) falls to  $1/e$  (61% in amplitude) in  $Q/2\pi$  cycles. You may prefer to think in radians: the energy falls to  $1/e$  in  $Q$  radians, and the voltage falls to  $1/e$  in  $2Q$  radians. *LC* resonant circuits are not unique in providing highly frequency-selective circuit behavior; alternatives include quartz-crystal, ceramic, and surface acoustic-wave (SAW) resonators; transmission lines; and resonant cavities.

### 1.7.15 *LC* filters

By combining inductors with capacitors you can produce filters (lowpass, highpass, bandpass) with far sharper behavior in frequency response than you can with a filter made from a simple *RC*, or from any number of cascaded *RC* sections. We'll see more of this, and the related topic of active filters, in Chapter 6. But it's worth admiring now how well this works, to appreciate the virtue of the humble inductor (an often-maligned circuit component).

As an example, look at Figure 1.111, a photograph of a "mixer-digitizer" circuit board that we built for a project some years back (specifically, a radio receiver with 250 million simultaneous channels). There's lots of stuff on the board, which has to frequency-shift and digitize three



**Figure 1.111.** There are six *LC* lowpass filters on this circuit board, part of the process of frequency conversion and digitizing for which this "mixer-digitizer" was designed.

RF bands; its design could occupy a book chapter. For now just gaze at the lumpy filter in the oval (there are five more on the board), comprising three inductors (the square metal cans) and four capacitors (the pairs of shiny oblongs). It's a lowpass filter, designed to cut off at 1.0 MHz; it prevents "aliases" in the digitized output, a subject we'll visit in Chapter 13.

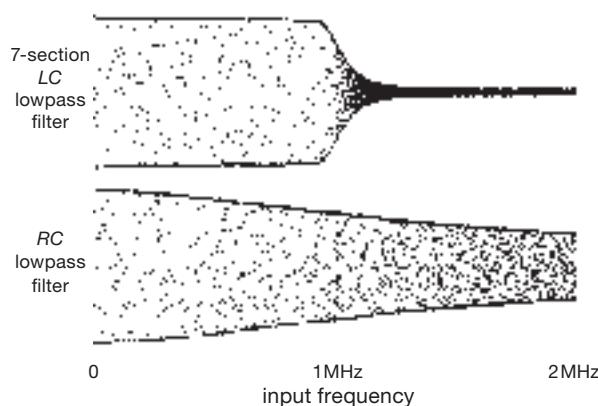
How well does it work? Figure 1.112 shows a "frequency sweep," in which a sinewave input goes from 0 Hz to 2 MHz as the trace goes from left to right across the screen. The sausage shapes are the "envelope" of the sinewave output, here comparing the *LC* filter with an *RC* lowpass filter with the same 1 MHz cutoff ( $1\text{ k}\Omega$  and  $160\text{ pF}$ ). The *LC* wins, hands down. The *RC* is pathetic by comparison. It's not even good English to call 1 MHz its "cutoff": it hardly cuts anything off.

### 1.7.16 Other capacitor applications

In addition to their uses in filters, resonant circuits, differentiators, and integrators, capacitors are needed for several other important applications. We treat these in detail later in the book, mentioning them here only as a preview.

#### A. Bypassing

The impedance of a capacitor decreases with increasing frequency. This is the basis of another important application: *bypassing*. There are places in circuits where you want to allow a dc voltage, but you don't want signals present. Placing a capacitor across that circuit element (usually a resistor) will help to kill any signals there. You choose the (noncritical) capacitor value so that its



**Figure 1.112.** Frequency sweep of the *LC* lowpass filter shown in Figure 1.111 compared with an *RC* lowpass filter with the same 1 MHz cutoff frequency. The dark outline is the amplitude envelope of the fast swept sinewave, which achieves a sandpaper appearance in this digital 'scope capture.

impedance at signal frequencies is small compared with what it is bypassing. You will see much more of this in later chapters.

## B. Power-supply filtering

We saw this application in §1.6.3, to filter the ripple from rectifier circuits. Although circuit designers often call them *filter* capacitors, this is really a form of bypassing, or energy storage, with large-value capacitors; we prefer the term *storage* capacitor. And these capacitors really are large – they're the big shiny round things you see inside most electronic instruments. We'll get into dc power-supply design in detail in Chapter 9.

## C. Timing and waveform generation

As we've seen, a capacitor supplied with a constant current charges up with a ramp waveform. This is the basis of ramp and sawtooth generators, used in analog function generators, oscilloscope sweep circuits, analog–digital converters, and timing circuits. *RC* circuits are also used for timing, and they form the basis of delay circuits (monostable multivibrators). These timing and waveform applications are important in many areas of electronics and will be covered in Chapters 3, 6, 10, and 11.

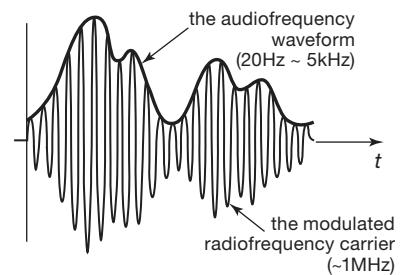
### 1.7.17 Thévenin's theorem generalized

When capacitors and inductors are included, Thévenin's theorem must be restated: any two-terminal network of resistors, capacitors, inductors, and signal sources is equivalent to a single complex impedance in series with a single signal source. As before, you find the (complex) impedance

and the signal source (waveform, amplitude, and phase) from the open-circuit output voltage and the short-circuit output current.

## 1.8 Putting it all together – an AM radio

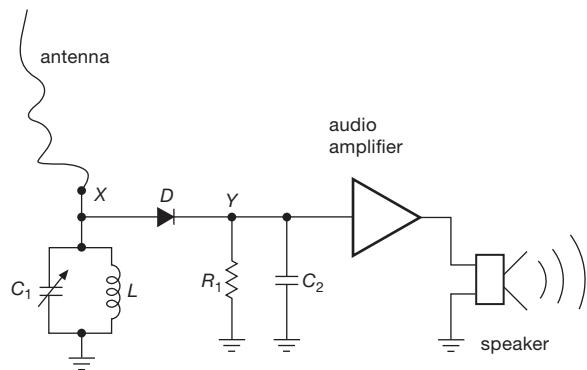
In our circuit course we tie together the topics of this chapter by hooking up a simple AM radio. The signal that's transmitted is a sinewave at the station's frequency in the AM band (520–1720 kHz), with its amplitude varied ("modulated") according to the audio waveform (Figure 1.113). In other words, an audio waveform described by some function  $f(t)$  would be transmitted as a RF signal  $[A + f(t)] \sin 2\pi f_c t$ ; here  $f_c$  is the station's "carrier" frequency, and the constant  $A$  is added to the audio waveform so that the coefficient  $[A + f(t)]$  is never negative.



**Figure 1.113.** An AM signal consists of an RF carrier (~1 MHz) whose amplitude is varied by the audio-frequency signal (speech or music; audible frequencies up to ~5 kHz). The audio waveform is dc offset so that the envelope does not cross zero.

At the receiver end (that's *us!*) the task is to select this station (among many) and somehow extract the modulating *envelope*, which is the desired audio signal. Figure 1.114 shows the simplest AM radio; it is the "crystal set" of yesteryear. It's really quite straightforward: the parallel *LC* resonant circuit is tuned to the station's frequency by the variable capacitor  $C_1$  (§1.7.14); the diode  $D$  is a half-wave rectifier (§1.6.2), which (if ideal) would pass only the positive half-cycles of the modulated carrier; and  $R_1$  provides a light load, so that the rectified output follows the half-cycles back down to zero. We're almost done. We add small capacitor  $C_2$  to prevent the output from following the fast half-cycles of carrier (it's a storage capacitor, §1.7.16B), choosing the time constant  $R_1 C_2$  to be long compared with a carrier period (~1  $\mu$ s), but short compared with the period of the highest audio frequency (~200  $\mu$ s).

Figures 1.115 and 1.116 show what you see when you probe around with a 'scope. The *bare* antenna shows plenty of low-frequency pickup (mostly 60 Hz ac powerline), and a tiny bit of signal from all the AM stations at once. But



**Figure 1.114.** The simplest AM receiver. Variable capacitor  $C_1$  tunes the desired station, diode  $D$  picks off the positive envelope (smoothed by  $R_1C_2$ ), and the resulting weak audio signal is amplified to drive the loudspeaker, loudly.

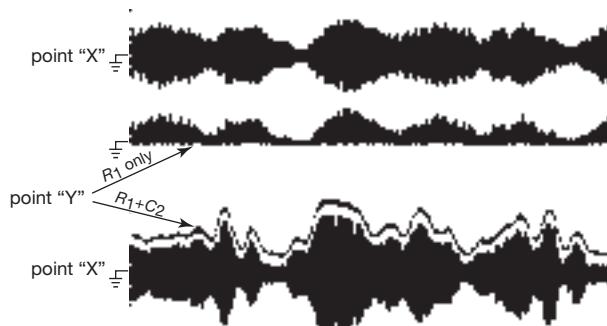
when you connect it to the  $LC$  resonant circuit, all the low-frequency stuff disappears (because the  $LC$  looks like a very low impedance, Figure 1.107) and it sees only the selected AM station. What's interesting here is that the amplitude of the selected station is much larger with the  $LC$  attached than with nothing connected to the antenna: that's because the resonant circuit's high  $Q$  is storing energy from multiple cycles of the signal.<sup>49</sup>



**Figure 1.115.** Observed waveforms at point "X" from the bare antenna (top) and with the  $LC$  connected. Note that the low-frequency junk disappears and that the radio signal gets *larger*. These are single-shot traces, in which the  $\sim 1$  MHz radiofrequency carrier appears as a solid filled area. Vertical: 1 V/div; horizontal: 4 ms/div.

The audio amplifier is fun, too, but we're not ready for it. We'll see how to make one of those in Chapter 2 (with discrete transistors), and again in Chapter 4 (with operational amplifiers, the Lego<sup>TM</sup> block of analog design).

<sup>49</sup> There are more complicated ways of framing this, but you don't really want to know just yet...



**Figure 1.116.** Observed waveforms at point "Y" with  $R_1$  only (top) and with smoothing capacitor  $C_2$  included (bottom). The upper pair is a single-shot capture (with the  $\sim 1$  MHz carrier appearing as solid black), and the lower pair is a separate single-shot capture, in which we have offset the rectified wave for clarity. Vertical: 1 V/div; horizontal: 1 ms/div.

And one amusing final note: in our class, we like to show the effect of probing point "X" with a length of BNC (bayonet Neill-Concelman) cable going to a 'scope input (that's how we start out, in the first week). When we do that, the cable's capacitance (about 30 pF/ft) adds to  $C_1$ , lowering the resonant frequency and so tuning to a different station. If we choose right, it changes *languages* (from English to Spanish)! The students howl with laughter – a language-translating electronic component. Then we use an ordinary 'scope probe, with its  $\sim 10$  pF of capacitance: no change of station, nor of language.

## 1.9 Other passive components

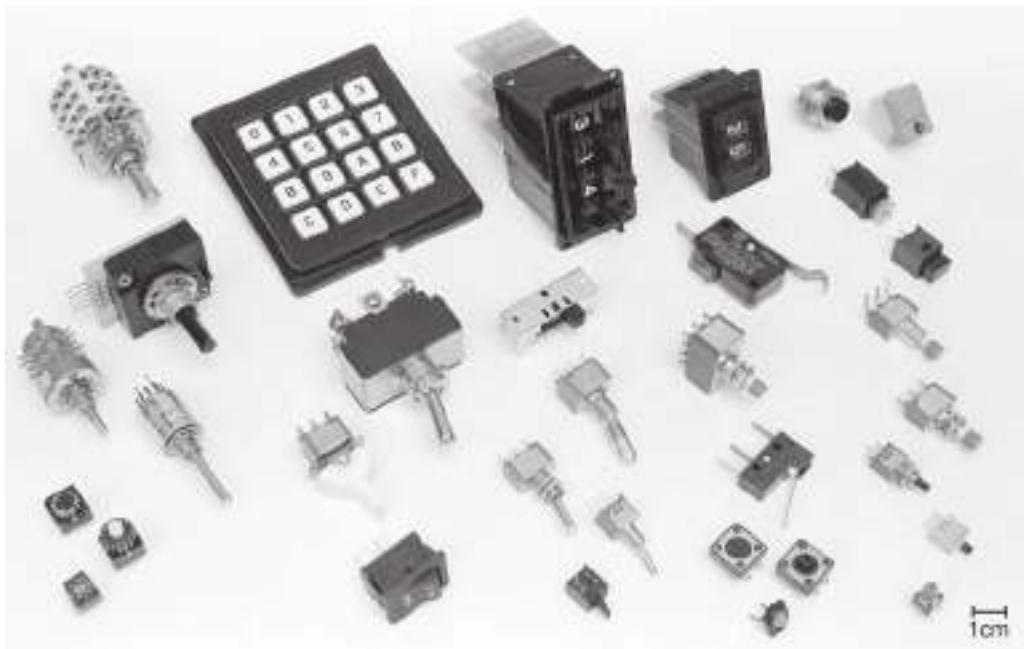
In the following subsections we would like to introduce briefly an assortment of miscellaneous but essential components. If you are experienced in electronic construction, you may wish to proceed to the next chapter.

### 1.9.1 Electromechanical devices: switches

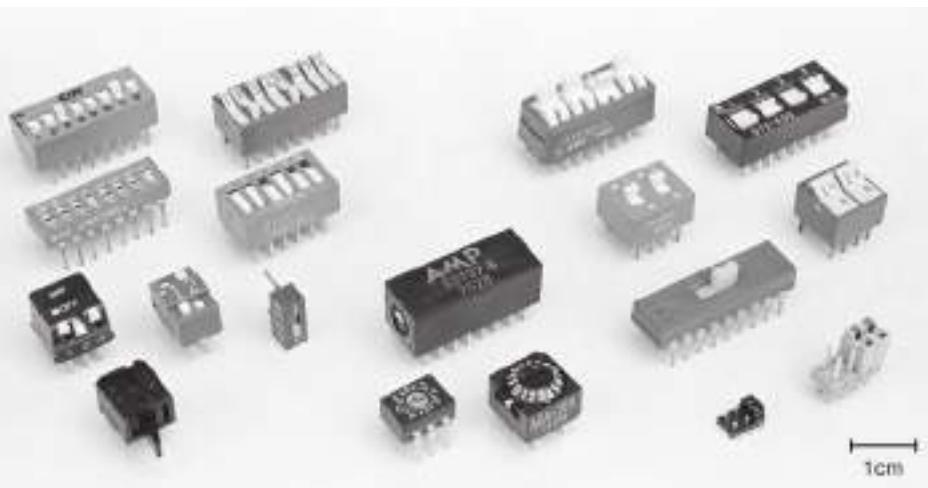
These mundane but important devices seem to wind up in most electronic equipment. It is worth spending a few paragraphs on the subject (and there's more in Chapter 1x). Figures 1.117 and 1.118 show some common switch types.

#### A. Toggle switches

The simple toggle switch is available in various configurations, depending on the number of poles; Figure 1.119 shows the usual ones (SPST indicates a single-pole single-throw switch, SPDT indicates a single-pole double-throw



**Figure 1.117.** Switch Smorgasbord. The nine switches at right are momentary-contact (“pushbutton”) switches, including both panel-mounting and PCB-mounting types (PCB, printed-circuit board). To their left are additional types, including lever-actuated and multipole styles. Above them are a pair of panel-mounting binary-coded *thumbwheel* switches, to the left of which is a matrix-encoded hexadecimal keypad. The switches at center foreground are toggle switches, in both panel-mounting and PCB-mounting varieties; several actuator styles are shown, including a locking variety (fourth from front) that must be pulled before it will switch. The rotary switches in the left column illustrate binary-coded types (the three in front and the larger square one), and the traditional multipole—multiposition configurable wafer switches.



**Figure 1.118.** Board-mounted “DIP switches.” Left group, front to back and left to right (all are SPST): single station side-action toggle; three-station side-action, two-station rocker, and single-station slide; eight-station slide (low-profile) and six-station rocker; eight-station slide and rocker. Middle group (all are hexadecimal coded): six-pin low-profile, six-pin with top or side adjust; 16-pin with true and complement coding. Right group: 2 mm×2 mm surface-mount header block with movable jumper (“shunt”), 0.1”×0.1” (2.54 mm×2.54 mm) through-hole header block with shunts; 18-pin SPDT (common actuator); eight-pin dual SPDT slide and rocker; 16-pin quad SPDT slide (two examples).

switch, and DPDT indicates a double-pole double-throw switch). Toggle switches are also available with “center OFF” positions and with up to four poles switched simultaneously. Toggle switches are always “break before make,” e.g., the moving contact never connects to both terminals in an SPDT switch.

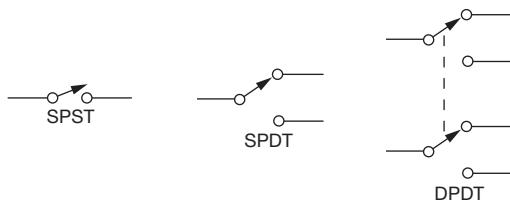


Figure 1.119. Fundamental switch types.

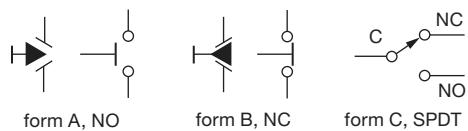


Figure 1.120. Momentary-contact (pushbutton) switches.

## B. Pushbutton switches

Pushbutton switches are useful for momentary-contact applications; they are drawn schematically as shown in Figure 1.120 (NO and NC mean normally open and normally closed). For SPDT momentary-contact switches, the terminals must be labeled NO and NC, whereas for SPST types the symbol is self-explanatory. Momentary-contact switches are always “break before make.” In the electrical (as opposed to electronic) industry, the terms form A, form B, and form C are used to mean SPST (NO), SPST (NC), and SPDT, respectively.

## C. Rotary switches

Rotary switches are available with many poles and many positions, often as kits with individual wafers and shaft hardware. Both *shorting* (make-before-break) and *non-shorting* (break-before-make) types are available, and they can be mixed on the same switch. In many applications the shorting type is useful to prevent an open circuit between switch positions, because circuits can go amok with unconnected inputs. Nonshorting types are necessary if the separate lines being switched to one common line must not ever be connected to each other.

Sometimes you don’t really want all those poles, you just want to know how many clicks (detents) the shaft has been turned. For that a common form of rotary switch en-

codes its position as a 4-bit binary quantity, thereby saving lots of wires (only five are needed: the four bits, and a common line). An alternative is the use of a *rotary encoder*, an electromechanical panel-mounting device that creates a sequence of  $N$  pulse pairs for each full rotation of the knob. These come in two flavors (internally using either mechanical contacts or electro-optical methods), and typically provide from 16 to 200 pulse pairs per revolution. The optical varieties cost more, but they last forever.

## D. PC-mounting switches

It’s common to see little arrays of switches on printed-circuit (PC) boards, like the ones shown in Figure 1.118. They’re often called *DIP switches*, referring to the integrated circuit dual in-line package that they borrow, though contemporary practice increasingly uses the more compact *surface-mount technology* (SMT) package. As the photograph illustrates, you can get coded rotary switches; and because these are used for set-and-forget internal settings, you can substitute a multipin *header* block, with little slide-on “shunts” to make the connections.

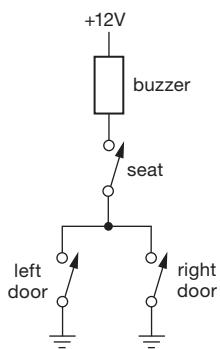
## E. Other switch types

In addition to these basic switch types, there are available various exotic switches such as Hall-effect switches, reed switches, proximity switches, etc. All switches carry maximum current and voltage ratings; a small toggle switch might be rated at 150 volts and 5 amps. Operation with inductive loads drastically reduces switch life because of arcing during turn-off. It’s always OK to operate a switch *below* its maximum ratings, with one notable exception: since many switches rely on substantial current flow to clean away contact oxides, it’s important to use a switch that is designed for “dry switching” when switching low-level signals;<sup>50</sup> otherwise you’ll get noisy and intermittent operation (see Chapter 1x).

## F. Switch examples

As an example of what can be done with simple switches, let’s consider the following problem: suppose you want to sound a warning buzzer if the driver of a car is seated and one of the car doors is open. Both doors and the driver’s seat have switches, all normally open. Figure 1.121 shows a circuit that does what you want. If one OR the other door is open (switch closed) AND the seat switch is closed, the buzzer sounds. The words OR and AND are used in a logic sense here, and we will see this example again in

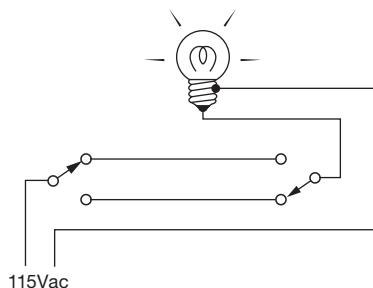
<sup>50</sup> These use gold contact plating.



**Figure 1.121.** Switch circuit example: open door warning.

Chapters 2, 3, and 10 when we talk about transistors and digital logic.

Figure 1.122 shows a classic switch circuit used to turn a ceiling lamp on or off from a switch at either of two entrances to a room.



**Figure 1.122.** Electrician's "three-way" switch wiring.

**Exercise 1.36.** Although few electronic circuit designers know how, every *electrician* can wire up a light fixture so that any of  $N$  switches can turn it on or off. See if you can figure out this generalization of Figure 1.122. It requires two SPDT switches and  $N-2$  DPDT switches.

## 1.9.2 Electromechanical devices: relays

Relays are electrically controlled switches. In the traditional electromechanical relay, a coil pulls in an armature (to close the contacts) when sufficient coil current flows. Many varieties are available, including "latching" and "stepping" relays.<sup>51</sup> Relays are available with dc or ac

<sup>51</sup> In an amusing historical footnote, the stepping relay used for a century as the cornerstone of telephone exchanges (the "Strowger selector") was invented by a Topeka undertaker, Almon Strowger, evidently because he suspected that telephone calls intended for his business were being routed (by the switchboard operators in his town) to a funeral home competitor.

excitation, and coil voltages from 3 volts up to 115 volts (ac or dc) are common. "Mercury-wetted" and "reed" relays are intended for high-speed ( $\sim 1$  ms) applications, and giant relays intended to switch thousands of amps are used by power companies.

The *solid-state relay* (SSR) – consisting of a semiconductor electronic switch that is turned on by a LED – provides better performance and reliability than mechanical relays, though at greater cost. SSRs operate rapidly, without contact "bounce," and usually provide for smart switching of ac power (they turn on at the moment of zero voltage, and they turn off at the moment of zero current). Much more on these useful devices in Chapter 12.

As we'll learn, electrically controlled switching of signals within a circuit can be accomplished with transistor switches, without having to use relays of any sort (Chapters 2 and 3). The primary uses of relays are in remote switching and high-voltage (or high-current) switching, where it is important to have complete electrical isolation between the control signal and the circuit being switched.

## 1.9.3 Connectors

Bringing signals in and out of an instrument, routing signal and dc power around between the various parts of an instrument, providing flexibility by permitting circuit boards and larger modules of the instrument to be unplugged (and replaced) – these are the functions of the *connector*, an essential ingredient (and usually the most unreliable part) of any piece of electronic equipment. Connectors come in a bewildering variety of sizes and shapes.<sup>52</sup> Figures 1.123, 1.124, and 1.125 give some idea of the variety.

### A. Single-wire connectors

The simplest kind of connector is the simple pin jack or banana jack used on multimeters, power supplies, etc. It is handy and inexpensive, but not as useful as the shielded-cable or multiwire connectors you often need. The humble binding post is another form of single-wire connector, notable for the clumsiness it inspires in those who try to use it.

### B. Shielded-cable connectors

To prevent capacitive pickup, and for other reasons we'll go into in Appendix H, it is usually desirable to pipe signals around from one instrument to another in shielded coaxial cable. The most popular connector is the BNC type that

<sup>52</sup> A search for "connector" on the DigiKey website returns 116 categories, with approximately 43,000 individual varieties in stock.



**Figure 1.123.** Rectangular connectors. The variety of available multipin connectors is staggering. Here is a collection of common specimens: the five connectors at lower left are multipin nylon power connectors (sometimes called *Molex-type* for historical reasons). Above them are four dual-row box headers (0.1" spacing, shown with and without latch ejectors, and also with Wire-Wrap® and right-angle tails), and to their right an open (“unshrouded”) 0.1" dual-row header, along with a pair of dual-row headers of finer pitch (2 mm and 1.27 mm). These dual-row male connectors mate with *insulation displacement* connectors (IDC) such as the one shown attached to a short length of ribbon cable (just above the unshrouded header). Just below the ribbon are shown single-row 0.1" headers, with mating shells (AMP MODU) that accept individual wire leads. At bottom right are several terminal blocks used for power wiring, and four “Faston”-type crimpable spade lugs. Above them are USB connectors, and to their left are the common RJ-45 and RJ-11 modular telephone/data connectors. The popular and reliable D-subminiature connectors are at center, including (right to left) a pair of 50-pin micro-D (cable plug, PCB socket), the 9-pin D-sub, 26-pin high-density, and a pair of 25-pin D-subs (one IDC). Above them are (right to left) a 96-pin VME backplane connector, a 62-pin card-edge connector with solder tails, a “Centronics-type” connector with latching bail, and a card-edge connector with ribbon IDC. At top left is a miscellany – a mating pair of “GR-type” dual banana connectors, a mating pair of Cinch-type connectors, a mating pair of shrouded Winchester-type connectors with locking jackscrews, and (to their right) a screw-terminal barrier block. Not shown here are the *really* tiny connectors used in small portable electronics (smartphones, cameras, etc); you can see a fine example in Figure 1.131.

adorns most instrument front panels. It connects with a quarter-turn twist and completes both the shield (ground) circuit and inner conductor (signal) circuit simultaneously. Like all connectors used to mate a cable to an instrument, it comes in both panel-mounting and cable-terminating varieties.

Among the other connectors for use with coaxial cable are the TNC (“threaded Neill-Concelman,” a close cousin of the BNC, but with threaded outer shell), the high-performance but bulky type N, the miniature SMA and SMB, the subminiature LEMO and SMC, and the high-voltage MHV and SHV. The so-called phono jack used in audio equipment is a nice lesson in bad design, because the

inner (signal) conductor mates *before* the shield (ground) when you plug it in; furthermore, the design of the connector is such that both shield and center conductor tend to make poor contact. You’ve undoubtedly *heard* the results! Not to be outdone, the television industry has responded with its own bad standard, the type-F coax “connector,” which uses the unsupported inner wire of the coax as the pin of the male plug, and a shoddy arrangement to mate the shield.<sup>53</sup>

We hereby induct these losers into the Electronic

<sup>53</sup> Advocates of each would probably reply “This is our most *modestly* priced receptacle.”



**Figure 1.124.** Circular connectors. A selection of multipin and other “non-RF” connectors; the panel-mounting receptacle is shown to the left of each cable-mounted plug. Top row, left to right: “MS”-type (MIL-C-5015) rugged connector (available in hundreds of configurations), high-current (50 A) “Supericon,” multipin locking XLR. Middle row: weatherproof (Switchcraft EN3), 12 mm video (Hirose RM), circular DIN, circular mini-DIN, 4-pin microphone connector. Bottom row: locking 6-pin (Lemo), microminiature 7-pin shielded (Microtech EP-7S), miniature 2-pin shrouded (Litton SM), 2.5 mm power, banana, pin jack.

Components Hall of Infamy, some charter members of which are shown in Figure 1.126.

### C. Multipin connectors

Very frequently electronic instruments demand multiwire cables and connectors. There are literally dozens of different kinds. The simplest example is a three-wire “IEC” powerline cord connector. Among the more popular are the excellent type-D subminiature, the Winchester MRA series, the venerable MS type, and the flat ribbon-cable mass-termination connectors. These and others are shown in Figure 1.123.

Beware of connectors that can’t tolerate being dropped on the floor (the miniature hexagon connectors are classic) or that don’t provide a secure locking mechanism (e.g., the Jones 300 series).

### D. Card-edge connectors

The most common method used to make connection to printed-circuit cards is the card-edge connector, which mates to a row of gold-plated contacts at the edge of the card; common examples are the motherboard connectors that accept plug-in computer memory modules. Card-edge connectors may have from 15 to 100 or more connections,

and they come with different lug styles according to the method of connection. You can solder them to a “motherboard” or “backplane,” which is itself just another PCB containing the interconnecting wiring between the individual circuit cards. Alternatively, you may want to use edge connectors with standard solder-lug terminations, particularly in a system with only a few cards. A more reliable (though more costly) solution is the use of “two-part” PCB connectors, in which one part (soldered onto the board) mates with the other part (on a backplane, etc); an example is the widely used VME (VersaModule Eurocard) connector (upper right-hand corner of Figure 1.123).

## 1.9.4 Indicators

### A. Meters

To read out the value of some voltage or current, you have a choice between the time-honored moving-pointer type of meter and digital-readout meters. The latter are more expensive and more accurate. Both types are available in a variety of voltage and current ranges. There are, in addition, exotic panel meters that read out such things as VUs (volume units, an audio dB scale), expanded-scale ac volts (e.g., 105 to 130 V), temperature (from a thermocouple),



**Figure 1.125.** RF and shielded connectors. The panel-mounting receptacle is shown to the left of each cable-mounted plug. Top row, left to right: stereo phone jack, audio “XLR” type; N and UHF (RF connectors). Second row down: BNC, TNC, type F; MHV and SHV (high voltage). Third row down: 2.5 mm (3/32”) audio, 3.5 mm stereo, improved 3.5 mm stereo, phono (“RCA type”), LEMO coaxial. Bottom row: SMA (panel jack, flexible coax plug), SMA (board-mount jack, rigid coax plug), SMB; SC and ST (optical fiber).

percentage motor load, frequency, etc. Digital panel meters often provide the option of logic-level outputs, in addition to the visible display, for internal use by the instrument.

As a substitute for a dedicated meter (whether analog or digital), you increasingly see an LCD (liquid-crystal display) or LED panel with a meter-like pattern. This is flexible and efficient: with a graphic LCD display module (§12.5.3) you can offer the user a choice of “meters,” according to the quantity being displayed, all under the control of an embedded controller (a built-in microprocessor; see Chapter 15).

### B. Lamps, LEDs, and displays

Flashing lights, screens full of numbers and letters, eerie sounds – these are the stuff of science fiction movies, and except for the last, they form the subject of lamps and displays (see §12.5.3). Small incandescent lamps used to be standard for front-panel indicators, but they have been re-

placed with LEDs. The latter behave electrically like ordinary diodes, but with a forward voltage drop in the range of 1.5 to 2 volts (for red, orange, and some green LEDs; 3.6 V for blue<sup>54</sup> and high-brightness green; see Figure 2.8). When current flows in the forward direction, they light up. Typically, 2 mA to 10 mA produces adequate brightness. LEDs are cheaper than incandescent lamps, they last pretty much forever, and they come in four standard colors as well as “white” (which is usually a blue LED with a yellow fluorescent coating). They come in convenient panel-mounting packages; some even provide built-in current limiting.<sup>55</sup>

LEDs can also be used for digital displays, for example

<sup>54</sup> The invention of the gallium nitride blue LED was the breakthrough product of a lone and unappreciated employee of Nichia Chemical Industries, Shuji Nakamura.

<sup>55</sup> And of course, for both residential and commercial area lighting, LEDs have now largely relegated to the dustbin of history the century-old hot-filament incandescent lamp.



**Figure 1.126.** Components to avoid. We advise against using components like these, if you have a choice (see text if you need convincing!). Top row, left to right: low-value wirewound pot, type UHF connector, electrical tape (“just say no!”). Middle row: “cinch-type” connectors, microphone connector, hexagon connectors. Bottom row: slide switch, cheap IC socket (not “screw-machined”), type-F connector, open-element trimmer pot, phono connector.

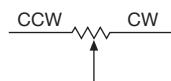
as 7-segment numeric displays or (for displaying letters as well as numbers – “alphanumeric”) 16-segment displays or dot-matrix displays. However, if more than a few digits or characters need to be displayed, LCDs are generally preferred. These come in line-oriented arrays (e.g., 16 characters by 1 line, up to 40 characters by 4 lines), with a simple interface that permits sequential or addressable entry of alphanumeric characters and additional symbols. They are inexpensive, low power, and visible even in sunlight. Back-lighted versions work well even in subdued light, but are not low power. Much more on these (and other) *optoelectronic* devices in §12.5.

## 1.9.5 Variable components

### A. Variable resistors

Variable resistors (also called volume controls, potentiometers, pots, or trimmers) are useful as panel controls or internal adjustments in circuits. A classic panel type is the 2-watt-type AB potentiometer; it uses the same basic material as the fixed carbon-composition resistor, with a rotatable “wiper” contact. Other panel types are available with ceramic or plastic resistance elements, with improved characteristics. Multiturn types (3, 5, or 10 turns) are available, with counting dials, for improved resolution and linearity. “Ganged” pots (several independent sections on one shaft) are also manufactured, although in limited variety, for applications that demand them. Figure 1.8 shows a representative selection of pots and trimmers.

For use inside an instrument, rather than on the front panel, *trimmer pots* come in single-turn and multiturn styles, most intended for printed-circuit mounting. These are handy for calibration adjustments of the “set-and-forget” type. Good advice: resist the temptation to use lots of trimmers in your circuits. Use good design instead.



**Figure 1.127.** Potentiometer (three-terminal variable resistor).

The symbol for a variable resistor, or pot, is shown in Figure 1.127. Sometimes the symbols CW and CCW are used to indicate the clockwise and counterclockwise ends.

An all-electronic version of a potentiometer can be made with an array of electronic (transistor) switches that select a tap in a long chain of fixed resistors. As awkward as that may sound, it is a perfectly workable scheme when implemented as an IC. For example, Analog Devices, Maxim/Dallas Semiconductor, and Xicor make a series of “digital potentiometers” with up to 1024 steps; they come as single or dual units, and some of them are “nonvolatile,” meaning that they remember their last setting even if power has been turned off. These find application in consumer electronics (televisions, stereos) where you want to adjust the volume from your infrared remote control, rather than by turning a knob; see §3.4.3E.

One important point about variable resistors: don’t attempt to use a potentiometer as a substitute for a precise resistor value somewhere within a circuit. This is tempting, because you can trim the resistance to the value you want. The trouble is that potentiometers are not as stable as good (1%) resistors, and in addition they may not have good resolution (i.e., they can’t be set to a precise value). If you must have a precise and settable resistor value somewhere, use a combination of a 1% (or better) precision resistor and a potentiometer, with the fixed resistor contributing most of the resistance. For example, if you need a 23.4k resistor, use a 22.6k 1% fixed resistor (a standard value) in series with a 2k trimmer pot. Another possibility is to use a series combination of several precision resistors, selecting the last (and smallest) resistor to give the desired series resistance.

As we’ll see later (§3.2.7), it is possible to use FETs as voltage-controlled variable resistors in some applications. Another possibility is an “*optophotoresistor*” (§12.7). Transistors can be used as variable-gain amplifiers, again controlled by a voltage. Keep an open mind when design brainstorming.



Figure 1.128. Variable capacitor.

### B. Variable capacitors

Variable capacitors are primarily confined to the smaller capacitance values (up to about 1000 pF) and are commonly used in RF circuits. Trimmers are available for in-circuit adjustments, in addition to the panel type for user tuning. Figure 1.128 shows the symbol for a variable capacitor.

Diodes operated with applied reverse voltage can be used as voltage-variable capacitors; in this application they're called *varactors*, or sometimes *varicaps* or *epi-caps*. They're very important in RF applications, especially phase-locked loops, automatic frequency control (AFC), modulators, and parametric amplifiers.

### C. Variable inductors

Variable inductors are usually made by arranging to move a piece of core material in a fixed coil. In this form they're available with inductances ranging from microhenrys to henrys, typically with a 2:1 tuning range for any given inductor. Also available are rotary inductors (coreless coils with a rolling contact).<sup>56</sup>

### D. Variable transformers

Variable transformers are handy devices, especially the ones operated from the 115 volt ac line. They're usually configured as "autotransformers," which means that they have only one winding, with a sliding contact. They're also commonly called Variacs (the name given to them by General Radio), and they are made by Technipower, Superior Electric, and others. Figure 1.129 shows a classic unit from General Radio. Typically they provide 0 to 135 volts ac output when operated from 115 volts, and they come in current ratings from 1 amp to 20 amps or more. They're good for testing instruments that seem to be affected by power-line variations, and in any case to verify worst-case performance. *Important Warning:* don't forget that the output is not electrically isolated from the powerline, as it would be with a transformer!



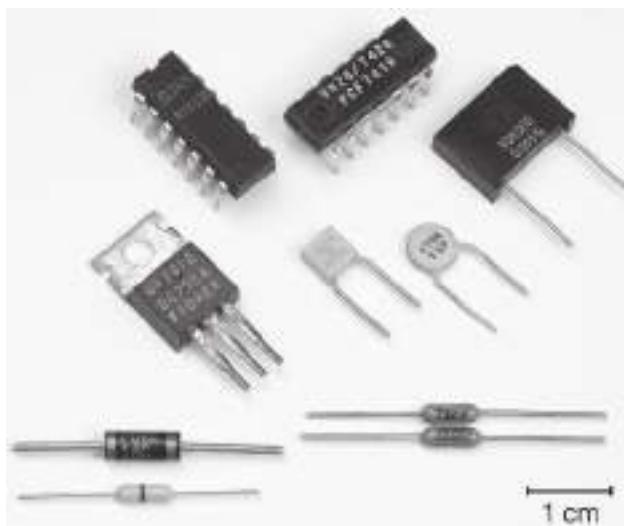
Figure 1.129. A powerline variable transformer ("Variac") lets you adjust the ac input voltage to something you are testing. Here a 5 A unit is shown, both clothed and undressed.

### 1.10 A parting shot: confusing markings and itty-bitty components

In our electronics course,<sup>57</sup> and indeed in day-to-day electronics on the bench, we encounter a wonderful confusion of component markings. Capacitors in particular are just, well, perverse: they rarely bother specifying *units* (even though they span 12 orders of magnitude, picofarads to farads), and for ceramic SMT varieties they dispense with any markings whatsoever! Even worse, they are still caught up in the transition from printing the value as an integer (e.g., "470" meaning 470 pF) versus using exponent notation (e.g., "470" meaning  $47 \times 10^0$ , i.e., 47 pF). Figure 1.130 shows exactly that case! Another trap for the unwary (and sometimes the wary, as well) is the date-code *gotcha*: the 4-digit code (yydd) can masquerade as a part number, as in the four examples in the photo. And, as components become smaller and smaller, there's precious little room for all but the briefest of markings; so, following the pharmaceutical industry, manufacturers invent a short

<sup>57</sup> Physics 123 ("Laboratory Electronics") at Harvard University: "Half course (fall term; repeated spring term). A lab-intensive introduction to electronic circuit design. Develops circuit intuition and debugging skills through daily hands-on lab exercises, each preceded by class discussion, with minimal use of mathematics and physics. Moves quickly from passive circuits, to discrete transistors, then concentrates on operational amplifiers, used to make a variety of circuits including integrators, oscillators, regulators, and filters. The digital half of the course treats analog-digital interfacing, emphasizing the use of microcontrollers and programmable logic devices (PLDs)." See <http://webdocs.registrar.fas.harvard.edu/courses/Physics.html>.

<sup>56</sup> An interesting form of variable inductor of yesteryear was the *variometer*, a rotatable coil positioned within a fixed outer coil and connected in series with it. As the inner coil was rotated, the total inductance went from maximum (four times the inductance of either coil alone) all the way down to zero. These things were *consumer* items, listed for example in the 1925 Sears Roebuck catalog.

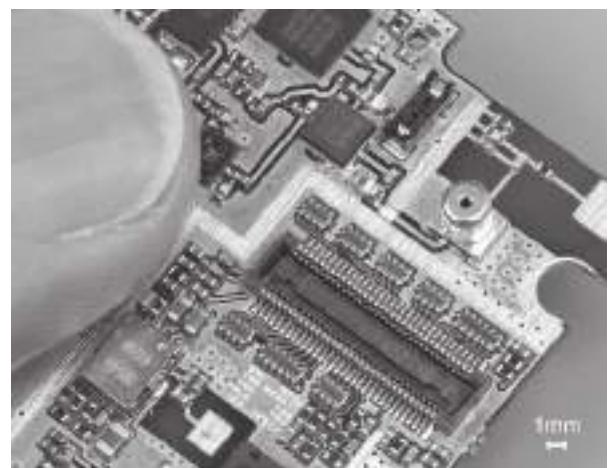


**Figure 1.130.** Confusion Central! The three ICs are each marked with both a part number (e.g., UA7812) and a “date code” (e.g., UC7924, signifying the 24th week of 1979). Unfortunately, both are perfectly valid part numbers (a +12V or a -24V regulator). The resistor pair (actually two views of identically marked resistors) suffers from the same problem: it could be  $7.32\text{ k}\Omega \pm 1\%$ , or it could be  $85.0\text{ k}\Omega \pm 5\%$  (it’s the former, but who would know?). The pair of ceramic capacitors are both marked 470K (470,000 of something?), but, surprise, the “K” means 10% tolerance; and, bigger surprise, the square cap is 47 pF, the round one is 470 pF. And what is one to make of a black box labeled 80K000 (pronounced “eighty-koooh”), or a diode with two cathodes (and no anode?), or a resistor with a single black band in the center?

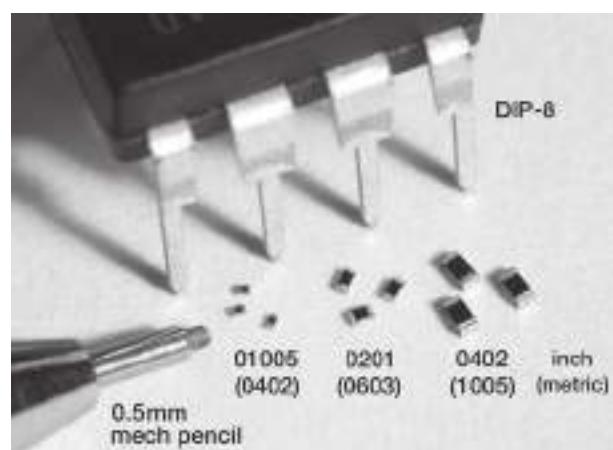
alphanumeric code for each component. And that’s all you get. For example, National’s LMOV81 op-amp comes in several 6-pin packages: the SOT23 is marked “A78A,” the smaller SC70 says “A77,” and the really tiny microSMD blurts out a single letter “A” (or “H” if it’s free of lead). Not much to go on.

### 1.10.1 Surface-mount technology: the joy and the pain

While we’re complaining, let’s whine just a bit about the difficulty of prototyping circuits with tiny surface-mount components. From an *electrical* point of view they are excellent: low inductance, and compact. But they are nearly impossible to wire up in prototype breadboard fashion, in the way that was easy with “through-hole” (or “leaded” – pronounced lee’-ded) components, such as resistors with axial leads (a wire sticking out each end), or integrated circuits in DIP (dual in-line) cases. Figure 1.131 gives



**Figure 1.131.** We’re “all thumbs” when working with surface-mount technology (SMT). This is a corner of a cellphone circuit board, showing small ceramic resistors and capacitors, integrated circuits with ball-grid connecting dots on their undersides, and the Lilliputian connectors for the antenna and display panel. See also Figure 4.84.



**Figure 1.132.** How small can these things get?! The “01005”-size SMT ( $0.016'' \times 0.008''$ , or  $0.4\text{mm} \times 0.2\text{mm}$ ) represents the industry’s greatest insult to the experimenter.

a sense of the scale of these little components, and Figure 1.132 displays the true horror of the tiniest of these – the “01005”-size chip components (0402 metric) that measure  $200\mu\text{m} \times 400\mu\text{m}$ : not much thicker than a human hair, and indistinguishable from dust!

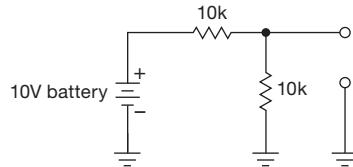
Sometimes you can use little adapter carriers (from companies like Bellin Dynamic Systems, Capital Advanced Technologies, or Aries) to convert an SMT integrated circuit to a fake DIP. But the densest surface-mount



**Figure 1.133.** A taste of the world of passive components in surface-mount packages: connectors, switches, trimmer pots, inductors, resistors, capacitors, crystals, fuses.... If you can name it, you can probably get it in SMT.

packages have no leads at all, just an array of bumps (up to several thousand!) on the underside; and these require serious “reflow” equipment before you can do anything with them. Sadly, we cannot ignore this disturbing trend, because the majority of new components are offered only in surface-mount packages. Woe to the lone basement experimenter-inventor! Figure 1.133 give a sense of the variety of passive component types that come in surface-mount configurations.

down at 10 kHz). Use the same source and load impedances as in Exercise 1.39.



**Figure 1.134.** Example for Norton equivalent circuit.

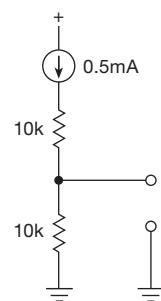
### Additional Exercises for Chapter 1

**Exercise 1.37.** Find the Norton equivalent circuit (a current source in parallel with a resistor) for the voltage divider in Figure 1.134. Show that the Norton equivalent gives the same output voltage as the actual circuit when loaded by a 5k resistor.

**Exercise 1.38.** Find the Thévenin equivalent for the circuit shown in Figure 1.135. Is it the same as the Thévenin equivalent for Exercise 1.37?

**Exercise 1.39.** Design a “rumble filter” for audio. It should pass frequencies greater than 20 Hz (set the  $-3\text{ dB}$  point at 10 Hz). Assume zero source impedance (perfect voltage source) and 10k (minimum) load impedance (that’s important so that you can choose  $R$  and  $C$  such that the load doesn’t affect the filter operation significantly).

**Exercise 1.40.** Design a “scratch filter” for audio signals (3 dB



**Figure 1.135.** Example for Thévenin equivalent circuit.

**Exercise 1.41.** How would you make a filter with  $R$ ’s and  $C$ ’s to give the response shown in Figure 1.136?

**Exercise 1.42.** Design a bandpass  $RC$  filter (as in Figure 1.137);

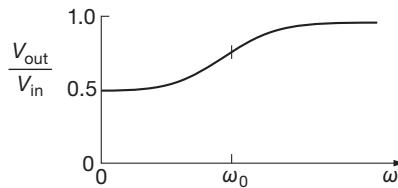


Figure 1.136. High-emphasis filter response.

$f_1$  and  $f_2$  are the 3 dB points. Choose impedances so that the first stage isn't much affected by the loading of the second stage.

**Exercise 1.43.** Sketch the output for the circuit shown in Figure 1.138.

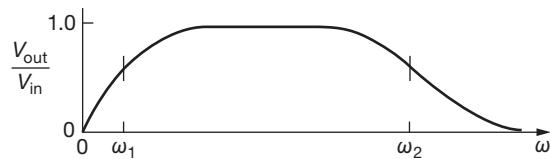


Figure 1.137. Bandpass filter response.

**Exercise 1.44.** Design an oscilloscope “ $\times 10$  probe” to use with a scope whose input impedance is  $1 \text{ M}\Omega$  in parallel with  $20 \text{ pF}$  by figuring out what goes inside the probe handle in Figure 1.139. Assume that the probe cable adds an additional  $100 \text{ pF}$  and that the probe components are placed at the tip end (rather than at the scope end) of the cable. The resultant network should have

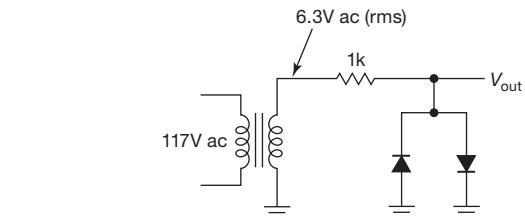
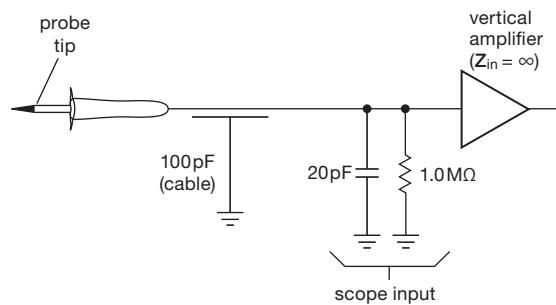


Figure 1.138. Circuit for Exercise 1.43.

Figure 1.139. Oscilloscope  $\times 10$  probe.

20 dB ( $\times 10$  voltage division ratio) attenuation at all frequencies, including dc. The reason for using a  $\times 10$  probe is to increase the load impedance seen by the circuit under test, which reduces loading effects. What input impedance ( $R$  in parallel with  $C$ ) does your  $\times 10$  probe present to the circuit under test when used with the scope?

## Review of Chapter 1

An A-to-H summary of what we have learned in Chapter 1. This summary reviews basic principles and facts in Chapter 1, but it does not cover application circuit diagrams and practical engineering advice presented there.

### ¶A. Voltage and Current.

Electronic circuits consist of components connected together with wires. *Current* ( $I$ ) is the rate of flow of charge through some point in these connections; it's measured in amperes (or millamps, microamps, etc.). *Voltage* ( $V$ ) between two points in a circuit can be viewed as an applied driving "force" that causes currents to flow between them; voltage is measured in volts (or kilovolts, millivolts, etc.); see §1.2.1. Voltages and currents can be steady (dc), or varying. The latter may be as simple as the sinusoidal alternating voltage (ac) from the wallplug, or as complex as a high-frequency modulated communications waveform, in which case it's usually called a *signal* (see ¶B below). The algebraic sum of currents at a point in a circuit (a *node*) is zero (Kirchhoff's current law, KCL, a consequence of conservation of charge), and the sum of voltage drops going around a closed loop in a circuit is zero (Kirchhoff's voltage law, KVL, a consequence of the conservative nature of the electrostatic field).

### ¶B. Signal Types and Amplitude.

See §1.3. In digital electronics we deal with *pulses*, which are signals that bounce around between two voltages (e.g., +5 V and ground); in the analog world it's *sinewaves* that win the popularity contest. In either case, a periodic signal is characterized by its frequency  $f$  (units of Hz, MHz, etc.) or, equivalently, period  $T$  (units of ms,  $\mu$ s, etc.). For sinewaves it's often more convenient to use *angular* frequency (radians/s), given by  $\omega=2\pi f$ .

Digital amplitudes are specified simply by the HIGH and LOW voltage levels. With sinewaves the situation is more complicated: the amplitude of a signal  $V(t)=V_0 \sin \omega t$  can be given as (a) *peak* amplitude (or just "amplitude")  $V_0$ , (b) *root-mean-square* (rms) amplitude  $V_{\text{rms}}=V_0/\sqrt{2}$ , or (c) *peak-to-peak* amplitude  $V_{\text{pp}}=2V_0$ . If unstated, a sinewave amplitude is usually understood to be  $V_{\text{rms}}$ . A signal of rms amplitude  $V_{\text{rms}}$  delivers power  $P=V_{\text{rms}}^2/R_{\text{load}}$  to a resistive load (regardless of the signal's waveform), which accounts for the popularity of rms amplitude measure.

*Ratios* of signal amplitude (or power) are commonly expressed in *decibels* (dB), defined as  $\text{dB} = 10 \log_{10}(P_2/P_1)$  or  $20 \log_{10}(V_2/V_1)$ ; see §1.3.2. An amplitude ratio of 10 (or power ratio of 100) is 20 dB; 3 dB is a doubling of

power; 6 dB is a doubling of amplitude (or quadrupling of power). Decibel measure is also used to specify amplitude (or power) directly, by giving a reference level: for example, -30 dBm (dB relative to 1 mW) is 1 microwatt; +3 dBVRms is a signal of 1.4 V rms amplitude (2 Vpeak, 4 Vpp).

Other important waveforms are square waves, triangle waves, ramps, noise, and a host of *modulation* schemes by which a simple "carrier" wave is varied in order to convey information; some examples are AM and FM for analog communication, and PPM (pulse-position modulation) or QAM (quadrature-amplitude modulation) for digital communication.

### ¶C. The Relationship Between Current and Voltage.

This chapter concentrated on the fundamental, essential, and ubiquitous *two-terminal linear devices*: resistors, capacitors, and inductors. (Subsequent chapters deal with *transistors* – three-terminal devices in which a signal applied to one terminal controls the current flow through the other pair – and their many interesting applications. These include amplification, filtering, power conversion, switching, and the like.) The simplest linear device is the *resistor*, for which  $I=V/R$  (Ohm's Law, see §1.2.2A). The term "linear" means that the response (e.g., current) to a combined sum of inputs (i.e., voltages) is equal to the sum of the responses that each input would produce:  $I(V_1+V_2)=I(V_1)+I(V_2)$ .

### ¶D. Resistors, Capacitors, and Inductors.

The resistor is clearly linear. But it is not the only linear two-terminal component, because linearity does not require  $I \propto V$ . The other two linear components are *capacitors* (§1.4.1) and *inductors* (§1.5.1), for which there is a time-dependent relationship between voltage and current:  $I=CdV/dt$  and  $V=LdI/dt$ , respectively. These are the *time domain* descriptions. Thinking instead in the *frequency domain*, these components are described by their *impedances*, the ratio of voltage to current (as a function of frequency) when driven with a sinewave (§1.7). A linear device, when driven with a sinusoid, responds with a sinusoid of the same frequency, but with changed amplitude and phase. Impedances are therefore complex, with the real part representing the amplitude of the response that is in-phase, and the imaginary part representing the amplitude of the response that is in quadrature (90° out of phase). Alternatively, in the polar representation of complex impedance ( $Z=|Z|e^{j\theta}$ ), the magnitude  $|Z|$  is the ratio of magnitudes ( $|Z|=|V|/|I|$ ) and the quantity  $\theta$  is the phase shift between  $V$  and  $I$ . The impedances of the three

linear 2-terminal components are  $Z_R=R$ ,  $Z_C=-j/\omega C$ , and  $Z_L=j\omega L$ , where (as always)  $\omega=2\pi f$ ; see §1.7.5. Sinewave current through a resistor is in phase with voltage, whereas for a capacitor it leads by  $90^\circ$ , and for an inductor it lags by  $90^\circ$ .

### ¶E. Series and Parallel.

The impedance of components connected in series is the sum of their impedances; thus  $R_{\text{series}}=R_1+R_2+\dots$ ,  $L_{\text{series}}=L_1+L_2+\dots$ , and  $1/C_{\text{series}}=1/C_1+1/C_2+\dots$ . When connected in parallel, on the other hand, it's the *admittances* (inverse of impedance) that add. Thus the formula for capacitors in parallel looks like the formula for resistors in series,  $C_{\text{parallel}}=C_1+C_2+\dots$ ; and vice versa for resistors and inductors, thus  $1/R_{\text{parallel}}=1/R_1+1/R_2+\dots$ . For a pair of resistors in parallel this reduces to  $R_{\text{parallel}}=(R_1 R_2)/(R_1+R_2)$ . For example, two resistors of value  $R$  have resistance  $R/2$  when connected in parallel, or resistance  $2R$  in series.

The power dissipated in a resistor  $R$  is  $P=I^2 R=V^2/R$ . There is no dissipation in an ideal capacitor or inductor, because the voltage and current are  $90^\circ$  out of phase. See §1.7.6.

### ¶F. Basic Circuits with $R$ , $L$ , and $C$ .

Resistors are everywhere. They can be used to set an operating current, as for example when powering an LED or biasing a zener diode (Figure 1.16); in such applications the current is simply  $I=(V_{\text{supply}}-V_{\text{load}})/R$ . In other applications (e.g., as a transistor's load resistor in an amplifier, Figure 3.29) it is the *current* that is known, and a resistor is used to convert it to a voltage. An important circuit fragment is the *voltage divider* (§1.2.3), whose unloaded output voltage (across  $R_2$ ) is  $V_{\text{out}}=V_{\text{in}}R_2/(R_1+R_2)$ .

If one of the resistors in a voltage divider is replaced with a capacitor, you get a simple *filter*: lowpass if the lower leg is a capacitor, highpass if the upper leg is a capacitor (§§1.7.1 and 1.7.7). In either case the  $-3$  dB transition frequency is at  $f_{3\text{dB}}=1/2\pi RC$ . The ultimate rolloff rate of such a “single-pole” lowpass filter is  $-6$  dB/octave, or  $-20$  dB/decade; i.e., the signal amplitude falls as  $1/f$  well beyond  $f_{3\text{dB}}$ . More complex filters can be created by combining inductors with capacitors, see Chapter 6. A capacitor in parallel with an inductor forms a *resonant circuit*; its impedance (for ideal components) goes to infinity at the resonant frequency  $f=1/(2\pi\sqrt{LC})$ . The impedance of a *series LC* goes to zero at that same resonant frequency. See §1.7.14.

Other important capacitor applications in this chapter (§1.7.16) include (a) *bypassing*, in which a capacitor's low

impedance at signal frequencies suppresses unwanted signals, e.g., on a dc supply rail; (b) *blocking* (§1.7.1C), in which a highpass filter blocks dc, but passes all frequencies of interest (i.e., the breakpoint is chosen below all signal frequencies); (c) *timing* (§1.4.2D), in which an *RC* circuit (or a constant current into a capacitor) generates a sloping waveform used to create an oscillation or a timing interval; and (d) *energy storage* (§1.7.16B), in which a capacitor's stored charge  $Q=CV$  smooths out the ripples in a dc power supply.

In later chapters we'll see some additional applications of capacitors: (e) *peak detection* and *sample-and-hold* (§§4.5.1 and 4.5.2), which capture the voltage peak or transient value of a waveform, and (f) the *integrator* (§4.2.6), which performs a mathematical integration of an input signal.

### ¶G. Loading; Thévenin Equivalent Circuit.

Connecting a load (e.g., a resistor) to the output of a circuit (a “signal source”) causes the unloaded output voltage to drop; the amount of such *loading* depends on the load resistance, and the signal source's ability to drive it. The latter is usually expressed as the *equivalent source impedance* (or *Thévenin impedance*) of the signal. That is, the signal source is modeled as a perfect voltage source  $V_{\text{sig}}$  in series with a resistor  $R_{\text{sig}}$ . The output of the resistive voltage divider driven from an input voltage  $V_{\text{in}}$ , for example, is modeled as a voltage source  $V_{\text{sig}}=V_{\text{in}}R_2/(R_1+R_2)$  in series with a resistance  $R_{\text{sig}}=R_1R_2/(R_1+R_2)$  (which is just  $R_1||R_2$ ). So the output of a  $1\text{k}\Omega$ – $1\text{k}\Omega$  voltage divider driven by a  $10\text{ V}$  battery looks like  $5\text{ V}$  in series with  $500\ \Omega$ .

Any combination of voltage sources, current sources, and resistors can be modeled perfectly by a single voltage source in series with a single resistor (its “Thévenin equivalent circuit”), or by a single current source in parallel with a single resistor (its “Norton equivalent circuit”); see Appendix D. The Thévenin equivalent source and resistance values are found from the open-circuit voltage and short-circuit current as  $V_{\text{Th}}=V_{\text{oc}}$ ,  $R_{\text{Th}}=V_{\text{oc}}/I_{\text{sc}}$ ; and for the Norton equivalent they are  $I_{\text{N}}=I_{\text{sc}}$ ,  $R_{\text{N}}=V_{\text{oc}}/I_{\text{sc}}$ .

Because a load impedance forms a voltage divider with the signal's source impedance, it's usually desirable for the latter to be small compared with any anticipated load impedance (§1.2.5A). However, there are two exceptions: (a) a *current source* has a high source impedance (ideally infinite), and should drive a load of much lower impedance; and (b) signals of *high frequency* (or fast risetime), traveling through a length of cable, suffer reflections unless the load impedance equals the so-called “characteristic impedance”  $Z_0$  of the cable (commonly  $50\ \Omega$ ), see Appendix H.

### ¶H. The Diode, a Nonlinear Component.

There are important two-terminal devices that are not linear, notably the *diode* (or *rectifier*), see §1.6. The ideal diode conducts in one direction only; it is a “one-way valve.” The onset of conduction in real diodes is roughly at 0.5 V in the “forward” direction, and there is some small leakage current in the “reverse” direction, see Figure 1.55. Useful diode circuits include power-supply *rectification* (conversion of ac to dc, §1.6.2), signal rectification (§1.6.6A), *clamping* (signal limiting, §1.6.6C), and *gating* (§1.6.6B). Diodes are commonly used to prevent polarity

reversal, as in Figure 1.84; and their exponential current versus applied voltage can be used to fashion circuits with logarithmic response (§1.6.6E).

Diodes specify a maximum safe reverse voltage, beyond which avalanche breakdown (an abrupt rise of current) occurs. You don’t go there! But you can (and should) with a *zener diode* (§1.2.6A), for which a reverse breakdown voltage (in steps, going from about 3.3 V to 100 V or more) is specified. Zeners are used to establish a voltage within a circuit (Figure 1.16), or to limit a signal’s swing.

# BIPOLAR TRANSISTORS

## CHAPTER 2

### 2.1 Introduction

The transistor is our most important example of an “active” component, a device that can amplify, producing an output signal with more power in it than the input signal. The additional power comes from an external source of power (the power supply, to be exact). Note that *voltage* amplification isn’t what matters, since, for example, a step-up transformer, a “passive” component just like a resistor or capacitor, has voltage gain but no power gain.<sup>1</sup> Devices with power gain are distinguishable by their ability to make oscillators, by feeding some output signal back into the input.

It is interesting to note that the property of power amplification seemed very important to the inventors of the transistor. Almost the first thing they did to convince themselves that they had really invented something was to power a loudspeaker from a transistor, observing that the output signal sounded louder than the input signal.

The transistor is the essential ingredient of every electronic circuit, from the simplest amplifier or oscillator to the most elaborate digital computer. Integrated circuits (ICs), which have largely replaced circuits constructed from discrete transistors, are themselves merely arrays of transistors and other components built from a single chip of semiconductor material.

A good understanding of transistors is very important, even if most of your circuits are made from ICs, because you need to understand the input and output properties of the IC in order to connect it to the rest of your circuit and to the outside world. In addition, the transistor is the single most powerful resource for interfacing, whether between ICs and other circuitry or between one subcircuit and another. Finally, there are frequent (some might say too frequent) situations in which the right IC just doesn’t exist, and you have to rely on discrete transistor circuitry to do the job. As you will see, transistors have an excitement all their own. Learning how they work can be great fun.

<sup>1</sup> It is even possible to achieve modest voltage gain in a circuit comprising only resistors and capacitors. To explore this idea, surprising even to seasoned engineers, look at Appendix J on SPICE.

There are two major species of transistors: in this chapter we will learn about bipolar junction transistors (BJTs), which historically came first with their Nobel Prize-winning invention in 1947 at Bell Laboratories. The next chapter deals with “field-effect” transistors (FETs), the now-dominant species in digital electronics. To give the coarsest comparison, BJTs excel in accuracy and low noise, whereas FETs excel in low power, high impedance, and high-current switching; there is, of course, much more to this complex subject.

Our treatment of bipolar transistors is going to be quite different from that of many other books. It is common practice to use the *h*-parameter model and equivalent circuit. In our opinion that is unnecessarily complicated and unintuitive. Not only does circuit behavior tend to be revealed to you as something that drops out of elaborate equations, rather than deriving from a clear understanding in your own mind as to how the circuit functions; you also have the tendency to lose sight of which parameters of transistor behavior you can count on and, more important, which ones can vary over large ranges.

In this chapter we will instead build up a very simple introductory transistor model and immediately work out some circuits with it. Its limitations will soon become apparent; then we will expand the model to include the respected Ebers–Moll conventions. With the Ebers–Moll equations and a simple three-terminal model, you will have a good understanding of transistors; you won’t need to do a lot of calculations, and your designs will be first rate. In particular, they will be largely independent of the poorly controlled transistor parameters such as current gain.

Some important engineering notation should be mentioned. Voltage at a transistor terminal (relative to ground) is indicated by a single subscript (C, B, or E):  $V_C$  is the collector voltage, for instance. Voltage between two terminals is indicated by a double subscript:  $V_{BE}$  is the base-to-emitter voltage drop, for instance. If the same letter is repeated, that means a power-supply voltage:  $V_{CC}$  is the (positive) power-supply voltage associated with the collector,

and  $V_{EE}$  is the (negative) supply voltage associated with the emitter.<sup>2</sup>

### Why transistor circuits are difficult

For those learning electronics for the first time, this chapter will be difficult. Here's why: all the circuits in the last chapter dealt with *two-terminal devices*, whether linear (resistors, capacitors, inductors) or nonlinear (diodes). So there was only one voltage (the voltage between the terminals) and only one current (the current flowing through the device) to think about. Transistors, by contrast, are *three-terminal devices*, which means there are two voltages and two currents to juggle.<sup>3</sup>

#### 2.1.1 First transistor model: current amplifier

Let's begin. A bipolar transistor is a three-terminal device (Figure 2.1), in which a small current applied to the base controls a much larger current flowing between the collector and emitter. It is available in two flavors (*npn* and *pnp*), with properties that meet the following rules for *npn* transistors (for *pnp* simply reverse all polarities):

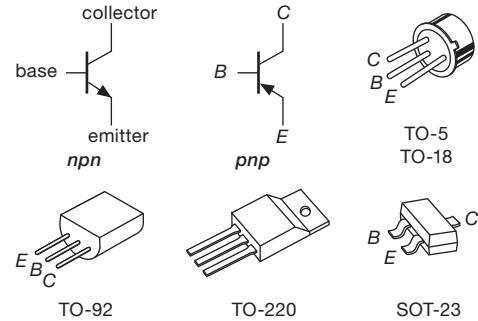
- 1. Polarity** The collector must be more positive than the emitter.
- 2. Junctions** The base-emitter and base-collector circuits behave like diodes (Figure 2.2) in which a small current applied to the base controls a much larger current flowing between the collector and emitter. Normally the base-emitter diode is conducting, whereas the base-collector diode is reverse-biased, i.e., the applied voltage is in the opposite direction to easy current flow.
- 3. Maximum ratings** Any given transistor has maximum values of  $I_C$ ,  $I_B$ , and  $V_{CE}$  that cannot be exceeded without costing the exceeding the price of a new transistor (for typical values, see the listing in Table 2.1 on page 74, Table 2.2 on page 106, and Table 8.1 on pages 501–502). There are also other limits, such as power dissipation ( $I_C V_{CE}$ ), temperature, and  $V_{BE}$ , that you must keep in mind.
- 4. Current amplifier** When rules 1–3 are obeyed,  $I_C$  is roughly proportional to  $I_B$  and can be written as

$$I_C = h_{FE} I_B = \beta I_B, \quad (2.1)$$

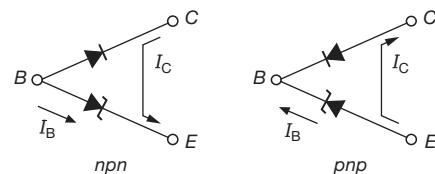
<sup>2</sup> In practice, circuit designers use  $V_{CC}$  to designate the positive supply and  $V_{EE}$  the negative supply, even though logically they should be interchanged for *pnp* transistors (where all polarities are reversed).

<sup>3</sup> You might think that there would be three voltages and three currents; but it's slightly less complicated than that, because there are only two independent voltages and two independent currents, thanks to Kirchhoff's voltage and current laws.

where  $\beta$ , the current gain (sometimes called<sup>4</sup>  $h_{FE}$ ), is typically about 100. Both  $I_B$  and  $I_C$  flow to the emitter. Note: the collector current is not due to forward conduction of the base-collector diode; that diode is reverse-biased. Just think of it as "transistor action."



**Figure 2.1.** Transistor symbols and small transistor package drawings (not to scale). A selection of common transistor packages are shown in Figure 2.3.



**Figure 2.2.** An ohmmeter's view of a transistor's terminals.

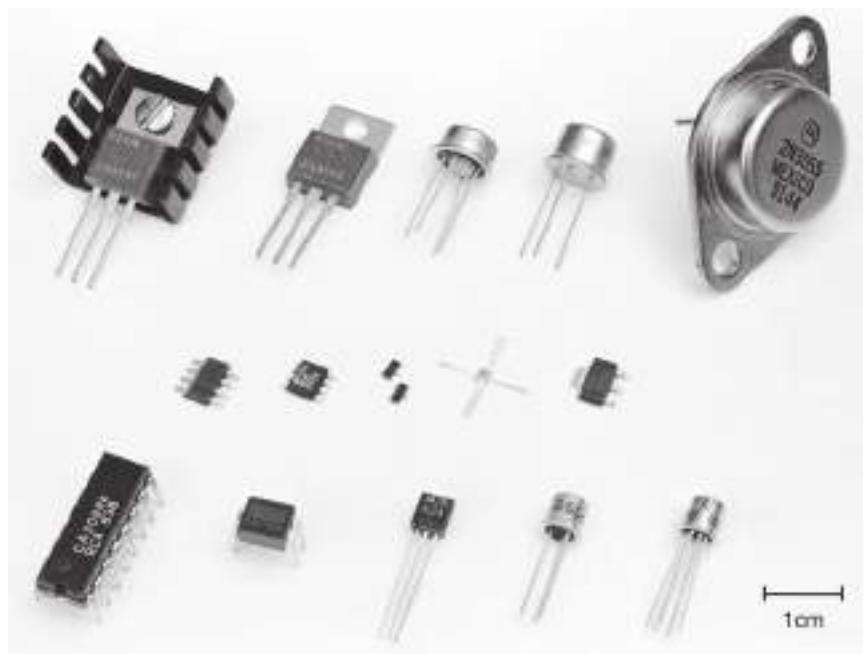
Rule 4 gives the transistor its usefulness: a small current flowing into the base controls a much larger current flowing into the collector.

An important warning: the current gain  $\beta$  is not a "good" transistor parameter; for instance, its value can vary from 50 to 250 for different specimens of a given transistor type. It also depends on the collector current, collector-to-emitter voltage, and temperature. *A circuit that depends on a particular value for beta is a bad circuit.*

Note particularly the effect of rule 2. This means you can't go sticking an arbitrary voltage across the base-emitter terminals, because an enormous current will flow if the base is more positive than the emitter by more than about 0.6 to 0.8 V (forward diode drop). This rule also implies that an operating transistor has  $V_B \approx V_E + 0.6$  V ( $V_B = V_E + V_{BE}$ ). Again, polarities are normally given for *npn* transistors; reverse them for *pnp*.

Let us emphasize again that you should not try to think of the collector current as diode conduction. It isn't,

<sup>4</sup> As the "h-parameter" transistor model has fallen out of popularity, you tend often to see  $\beta$  (instead of  $h_{FE}$ ) as the symbol for current gain.



**Figure 2.3.** Most of the common packages are shown here, for which we give the traditional designations. Top row (power), left to right: TO-220 (with and without heatsink), TO-39, TO-5, TO-3. Middle row (surface mount): SM-8 (dual), SO-8 (dual), SOT-23, ceramic SOE, SOT-223. Bottom row: DIP-16 (quad), DIP-4, TO-92, TO-18, TO-18 (dual).

because the collector–base diode normally has voltages applied across it in the reverse direction. Furthermore, collector current varies very little with collector voltage (it behaves like a not-too-great current source), unlike forward diode conduction, in which the current rises very rapidly with applied voltage.

Table 2.1 on the following page includes a selection of commonly used bipolar transistors, with the corresponding curves of current gain<sup>5</sup> in Figure 2.4, and a selection of transistors intended for power applications is listed in Table 2.2 on page 106. A more complete listing can be found in Table 8.1 on pages 501–502 and Figure 8.39 in Chapter 8.

## 2.2 Some basic transistor circuits

### 2.2.1 Transistor switch

Look at the circuit in Figure 2.5. This application, in which a small control current enables a much larger current to

flow in another circuit, is called a transistor switch. From the preceding rules it is easy to understand. When the mechanical switch is open, there is no base current. So, from rule 4, there is no collector current. The lamp is off.

When the switch is closed, the base rises to 0.6 V (base-emitter diode is in forward conduction). The drop across the base resistor is 9.4 V, so the base current is 9.4 mA. Blind application of Rule 4 gives  $I_C = 940$  mA (for a typical beta of 100). That is wrong. Why? Because rule 4 holds only if Rule 1 is obeyed: at a collector current of 100 mA the lamp has 10 V across it. To get a higher current you would have to pull the collector below ground. A transistor can't do this, and the result is what's called *saturation* – the collector goes as close to ground as it can (typical saturation voltages are about 0.05–0.2 V, see Chapter 2x.) and stays there. In this case, the lamp goes on, with its rated 10 V across it.

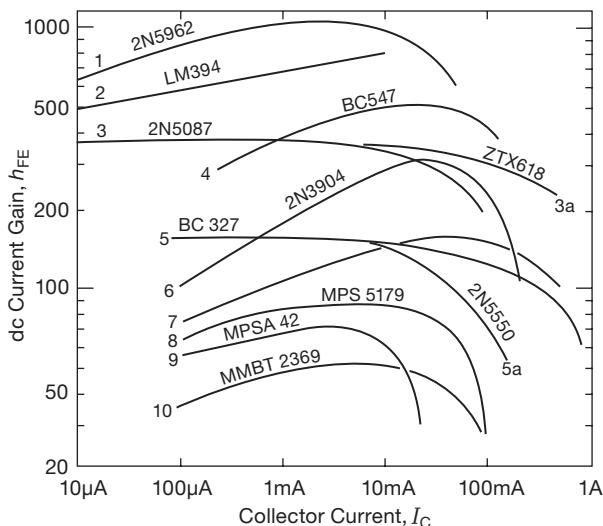
Overdriving the base (we used 9.4 mA when 1.0 mA would have barely sufficed) makes the circuit conservative; in this particular case it is a good idea, since a lamp draws more current when cold (the resistance of a lamp when cold is 5 to 10 times lower than its resistance at operating current). Also, transistor beta drops at low collector-to-base voltages, so some extra base current is necessary to bring

<sup>5</sup> In addition to listing typical betas ( $\beta_{FE}$ ) and maximum allowed collector-to-emitter voltages ( $V_{CEO}$ ), Table 2.1 includes the cutoff frequency ( $f_T$ , at which the beta has decreased to 1) and the feedback capacitance ( $C_{cb}$ ). These are important when dealing with fast signals or high frequencies; we'll see them in §2.4.5 and Chapter 2x.

**Table 2.1 Representative Bipolar Transistors**

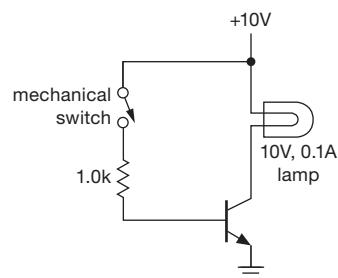
Part #											
<i>npn</i>		<i>pnp</i>		$V_{CEO}$ (V)	$I_C$ (max) (mA)	$h_{FE}$ @ mA (typ)	gain curve <sup>d</sup>	$C_{cb}^a$ (pF)	$f_T^a$ (MHz)	Comments	
TO-92	SOT-23	TO-92	SOT-23								
2N3904	MMBT3904	2N3906	MMBT3906	40	150	200	10	6	2.5	300	jellybean
2N4401	MMBT4401	2N4403	MMBT4403	40	500	150	150	7	7	300	'2222 and '2907 dies
BC337	BC817	BC327	BC807	45	750	350	40	5	10	150	jellybean
2N5089	MMBT5089	2N5087	MMBT5087	25	50	500	1	3	1.8	350	high beta
BC547C	BC847C	BC557C	BC857C	45	100	500	10	4	5	150	jellybean <sup>b</sup>
MPSA14	MMBTA14	MPSA64	MMBTA64	30	300	10000	50	-	7	125	Darlington
ZTX618	FMMT618	ZTX718	FMMT718	20	2500	320	3A	3a	-	120	high $I_C$ , small pkg
PN2369	MMBT2369	2N5771	MMBT5771	15	150	100	10	10	3	500	fast switch, gold doped
2N5551	MMBT5551	2N5401	MMBT5401	150	100	100	10	5a	2.5	100	SOT-223 available
MPSA42	MMBTA42	MPSA92	MMBTA92	300	30	75	10	9	1.5	50	HV small signal
MPS5179	BFS17	MPSH81	MMBTH81	15	25	90	20	8	0.9	900	RF amplifier
—	BFR93 <sup>c</sup>	—	BFT93 <sup>c</sup>	12	50	50	15	10	0.5	4000	RF amp
TIP142	—	TIP147	—	100	10A	>1000	5A	-	high	low	TO-220, Darlington

Notes: (a) see Chapter 2x for graphs of  $C_{cb}$  and  $f_T$ . (b) lower beta versions have an -A or -B suffix; low-noise versions are BC850 (*npn*) and BC860 (*pnp*). (c) also BFR25A and BFT25A. (d) see Figure 2.4.



**Figure 2.4.** Curves of typical transistor current gain,  $\beta$ , for a selection of transistors from Table 2.1. These curves are taken from manufacturers' literature. You can expect production spreads of  $\pm 100\%$ ,  $-50\%$  from the "typical" values graphed. See also Figure 8.39 for measured beta plots for 44 types of "low-noise" transistors.

a transistor into full saturation. Incidentally, in a real circuit you would probably put a resistor from base to ground (perhaps 10k in this case) to make sure the base is at ground with the switch open. It wouldn't affect the ON operation, because it would sink only 0.06 mA from the base circuit.



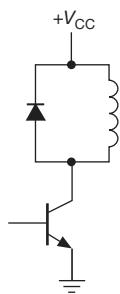
**Figure 2.5.** Transistor switch example.

There are certain cautions to be observed when designing transistor switches:

1. Choose the base resistor conservatively to get plenty of excess base current, especially when driving lamps, because of the reduced beta at low  $V_{CE}$ . This is also a good idea for high-speed switching, because of capacitive effects and reduced beta at very high frequencies (many megahertz).<sup>6</sup>
2. If the load swings below ground for some reason (e.g., it is driven from ac, or it is inductive), use a diode in series with the collector (or a diode in the reverse direction to ground) to prevent collector-base conduction on negative swings.
3. For inductive loads, protect the transistor with a diode

<sup>6</sup> A small "speed-up" capacitor – typically just a few picofarads – is often connected across the base resistor to improve high-speed performance.

across the load, as shown in Figure 2.6.<sup>7</sup> Without the diode the inductor will swing the collector to a large positive voltage when the switch is opened, most likely exceeding the collector-emitter breakdown voltage, as the inductor tries to maintain its “on” current from  $V_{CC}$  to the collector (see the discussion of inductors in §1.6.7).



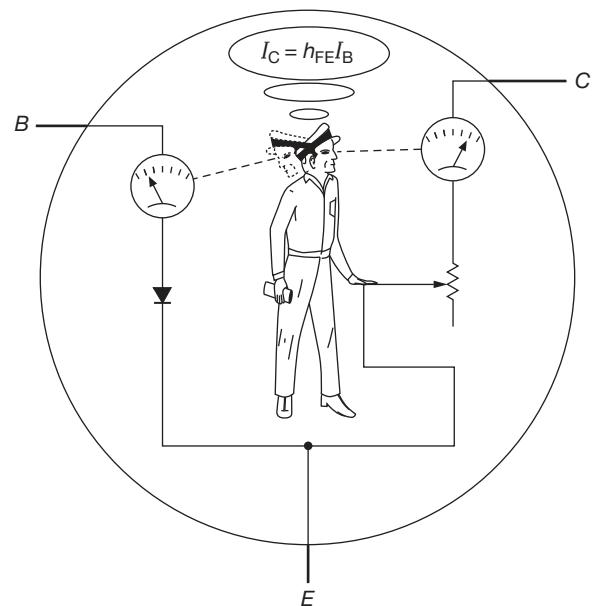
**Figure 2.6.** Always use a suppression diode when switching an inductive load.

You might ask why we are bothering with a transistor, and all its complexity, when we could just use that mechanical switch alone to control the lamp or other load. There are several good reasons: (a) a transistor switch can be driven *electrically* from some other circuit, for example a computer output bit; (b) transistor switches enable you to switch very rapidly, typically in a small fraction of a microsecond; (c) you can switch many different circuits with a single control signal; (d) mechanical switches suffer from wear, and their contacts “bounce” when the switch is activated, often making and breaking the circuit a few dozen times in the first few milliseconds after activation; and (e) with transistor switches you can take advantage of remote *cold switching*, in which only dc control voltages snake around through cables to reach front-panel switches, rather than the electronically inferior approach of having the signals themselves traveling through cables and switches (if you run lots of signals through cables, you’re likely to get capacitive pickup as well as some signal degradation).

#### A. “Transistor man”

The cartoon in Figure 2.7 may help you understand some limits of transistor behavior. The little man’s perpetual task in life is to try to keep  $I_C = \beta I_B$ ; however, he is only allowed to turn the knob on the variable resistor. Thus he can go from a short circuit (saturation) to an open circuit (transistor in the OFF state), or anything in between, but he isn’t allowed to use batteries, current sources, etc.

<sup>7</sup> Or, for faster turn-off, with a resistor, an  $RC$  network, or zener clamp; see §1.6.7.



**Figure 2.7.** “Transistor man” observes the base current, and adjusts the output rheostat in an attempt to maintain the output current  $\beta$  times larger;  $h_{FE}$  and  $\beta$  are used interchangeably.

One warning is in order here: don’t think that the collector of a transistor looks like a resistor. It doesn’t. Rather, it looks approximately like a poor-quality constant-current sink (the value of current depending on the signal applied to the base), primarily because of this little man’s efforts.

Another thing to keep in mind is that, at any given time, a transistor may be (a) cut off (no collector current), (b) in the active region (some collector current, and collector voltage more than a few tenths of a volt above the emitter), or (c) in saturation (collector within a few tenths of a volt of the emitter). See the discussion of transistor saturation in Chapter 2x for more details.

## 2.2.2 Switching circuit examples

The transistor switch is an example of a *nonlinear* circuit: the output is not proportional to the input;<sup>8</sup> instead it goes to one of two possible states (cut off, or saturated). Such two-state circuits are extremely common<sup>9</sup> and form the basis of digital electronics. But to the authors the subject of

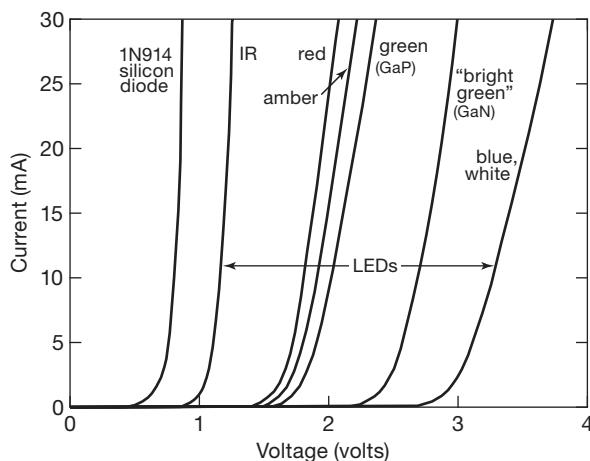
<sup>8</sup> A mathematician would define linearity by saying that the response to the sum of two inputs is the sum of the individual responses; this necessarily implies proportionality.

<sup>9</sup> If you took a census, asking the transistors of the world what they are doing, at least 95% would tell you they are switches.

*linear* circuits (such as amplifiers, current sources, and integrators) offers the most interesting challenges and the potential for great circuit creativity. We will move on to linear circuits in a moment, but this is a good time to enjoy a few circuit examples with transistors acting as switches – we like to give a feeling for the richness of electronics by showing real-world examples as soon as possible.

### A. LED driver

Light-emitting diode indicators – LEDs – have replaced the incandescent lamps of yesteryear for all electronic indicator and readout applications; they’re cheap, they come in lots of colors, and they last just about forever. Electrically they are similar to the ordinary silicon signal diodes we met in Chapter 1, but with a larger forward voltage drop (generally in the range of 1.5–3.5 V, rather than approximately<sup>10</sup> 0.6 V); that is, as you slowly increase the voltage across an LED’s terminals, you find that they start conducting current at, say, 1.5 V, and the current increases rapidly as you apply somewhat more voltage (Figure 2.8). They light up, too! Typical “high-efficiency” indicator LEDs look pretty good at a few millamps, and they’ll knock your eye out at 10–20 mA.

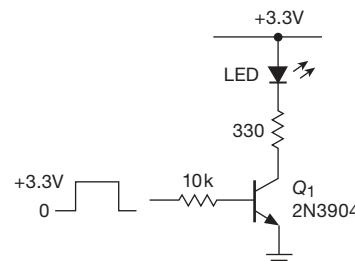


**Figure 2.8.** Like silicon diodes, LEDs have rapidly increasing current versus applied voltage, but with larger forward voltage drops.

We’ll show a variety of techniques for driving LEDs in Chapter 12; but we can drive them already, with what we know. The first thing to realize is that we can’t just switch a voltage across them, as in Figure 2.5, because of their steep  $I$  versus  $V$  behavior; for example, applying 5 V across an

LED is guaranteed to blow it out. We need instead to treat it gently, coaxing it to draw the right current.

Let’s assume that we want the LED to light in response to a digital signal line when it goes to a HIGH value of +3.3 V (from its normal resting voltage near ground). Let’s assume also that the digital line can provide up to 1 mA of current, if needed. The procedure goes like this: first, choose an LED operating current that will provide adequate brightness, say 5 mA (you might want to try a few samples, to make sure you like the color, brightness, and viewing angle). Then use an *n*p*n* transistor as a switch (Figure 2.9), choosing the collector resistor to provide the chosen LED current, realizing that the voltage drop across the resistor is the supply voltage minus the LED forward drop at its operating current. Finally, choose the base resistor to ensure saturation, assuming a conservatively low transistor beta ( $\beta \geq 25$  is pretty safe for a typical small-signal transistor like the popular 2N3904).



**Figure 2.9.** Driving an LED from a “logic-level” input signal, using an *n*p*n* saturated switch and series current-limiting resistor.

Note that the transistor is acting as a saturated switch, with the collector resistor setting the operating current. As we’ll see shortly, you can devise circuits that provide an accurate *current* output, largely independent of what the load does. Such a “current source” can also be used to drive LEDs. But our circuit is simple, and effective. There are other variations: we’ll see in the next chapter that a MOSFET-type<sup>11</sup> transistor is often a better choice. And in Chapters 10–12 we’ll see ways to drive LEDs and other optoelectronic devices directly from digital integrated circuits, without external discrete transistors.

**Exercise 2.1.** What is the LED current, approximately, in the circuit of Figure 2.9? What minimum beta is required for  $Q_1$ ?

### B. Variations on a theme

For these switch examples, one side of the load is connected to a positive supply voltage, and the other side is

<sup>10</sup> The larger drop is due to the use of different semiconductor materials such as GaAsP, GaAlAs, and GaN, with their larger bandgaps.

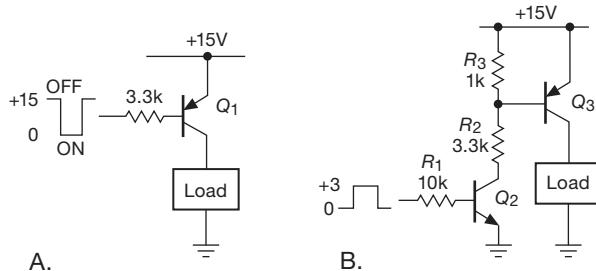
<sup>11</sup> metal-oxide semiconductor field-effect transistor.

switched to ground by the *npn* transistor switch. What if you want instead to ground one side of the load and switch the “high side” to a positive voltage?

It’s easy enough – but you’ve got to use the other polarity of transistor (*pnp*), with its emitter at the positive rail, and its collector tied to the load’s high side, as in Figure 2.10A. The transistor is cut off when the base is held at the emitter voltage (here +15 V), and switched into saturation by bringing the base toward the collector (i.e., toward ground). When the input is brought to ground, there’s about 4 mA of base current through the 3.3 k $\Omega$  base resistor, sufficient for switching loads up to about 200 mA ( $\beta > 50$ ).

An awkwardness of this circuit is the need to hold the input at +15 V to turn off the switch; it would be much better to use a lower control voltage, for example, +3 V and ground, commonly available in digital logic that we’ll be seeing in Chapters 10–15. Figure 2.10B shows how to do that: *npn* switch  $Q_2$  accepts the “logic-level” input of 0 V or +3 V, pulling its collector load to ground accordingly. When  $Q_2$  is cut off,  $R_3$  holds  $Q_3$  off; when  $Q_2$  is saturated (by a +3 V input),  $R_2$  sinks base current from  $Q_3$  to bring it into saturation.

The “divider” formed by  $R_2R_3$  may be confusing:  $R_3$ ’s job is to keep  $Q_3$  off when  $Q_2$  is off; and when  $Q_2$  pulls its collector low, most of its collector current comes from  $Q_3$ ’s base (because only  $\sim 0.6$  mA of the 4.4 mA collector current comes from  $R_3$  – make sure you understand why). That is,  $R_3$  does not have much effect on  $Q_3$ ’s saturation. Another way to say it is that the divider would sit at about +11.6 V (rather than +14.4 V), were it not for  $Q_3$ ’s base-emitter diode, which consequently gets most of  $Q_2$ ’s collector current. In any case, the value of  $R_3$  is not critical and could be made larger; the tradeoff is slower turn-off of  $Q_3$ , owing to capacitive effects.<sup>12</sup>

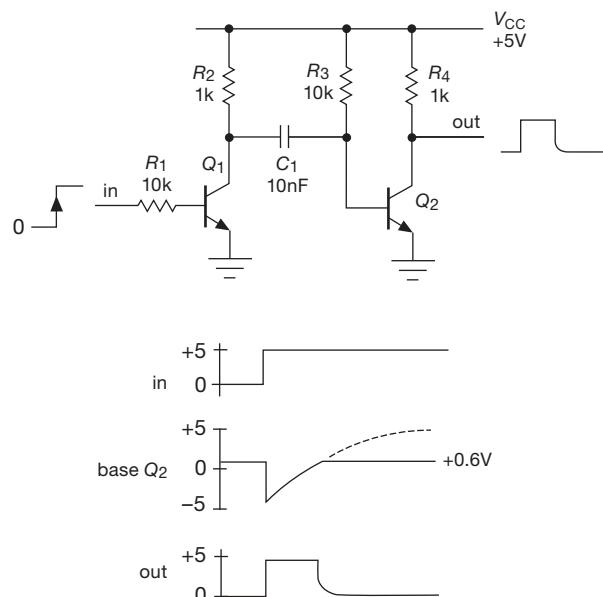


**Figure 2.10.** Switching the high side of a load returned to ground.

<sup>12</sup> But don’t make it too small:  $Q_3$  would not switch at all if  $R_3$  were reduced to 100  $\Omega$  (why?). We were surprised to see this basic error in an instrument, the rest of which displayed circuit design of the highest sophistication.

### C. Pulse generator – I

By including a simple  $RC$ , you can make a circuit that gives a pulse output from a step input; the time constant  $\tau = RC$  determines the pulse width. Figure 2.11 shows one way.  $Q_2$  is normally held in saturation by  $R_3$ , so its output is close to ground; note that  $R_3$  is chosen small enough to ensure  $Q_2$ ’s saturation. With the circuit’s input at ground,  $Q_1$  is cut off, with its collector at +5 V. The capacitor  $C_1$  is therefore charged, with +5 V on its left terminal and approximately +0.6 V on its right terminal; i.e., it has about 4.4 V across it. The circuit is waiting for something to happen.



**Figure 2.11.** Generating a short pulse from a step input waveform.

A +5 V positive input step brings  $Q_1$  into saturation (note the values of  $R_1$  and  $R_2$ ), forcing its collector to ground; because of the voltage across  $C_1$ , this brings the base of  $Q_2$  momentarily negative, to about -4.4 V.<sup>13</sup>  $Q_2$  is then cutoff, no current flows through  $R_4$ , and so its output jumps to +5 V; this is the beginning of the output pulse. Now for the  $RC$ :  $C_1$  can’t hold  $Q_2$ ’s base below ground forever, because current is flowing down through  $R_3$ , trying to pull it up. So the right-hand side of the capacitor charges toward +5 V, with a time constant  $\tau = R_3C_1$ , here equal to 100  $\mu$ s. The output pulse width is set by this time constant

<sup>13</sup> A caution here: this circuit should not be run from a supply voltage greater than +7 V, because the negative pulse can drive  $Q_2$ ’s base into reverse breakdown. This is a common oversight, even among experienced circuit designers.

and is proportional to  $\tau$ . To figure out the pulse width accurately you have to look in detail at the circuit operation. In this case it's easy enough to see that the output transistor  $Q_2$  will turn on again, terminating the output pulse, when the rising voltage on the base of transistor  $Q_2$  reaches the  $\approx 0.6$  V  $V_{BE}$  drop required for turn-on. Try this problem to test your understanding.

**Exercise 2.2.** Show that the output pulse width for the circuit of Figure 2.11 is approximately  $T_{pulse} = 0.76R_3C_1 = 76\mu s$ . A good starting point is to notice that  $C_1$  is charging exponentially from  $-4.4$  V toward  $+5$  V, with the time constant as above.

#### D. Pulse generator – II

Let's play with this circuit a bit. It works fine as described, but note that it requires that the input remain high throughout the duration of the output pulse, at least. It would be nice to eliminate that restriction, and the circuit in Figure 2.12 shows how. To the original circuit we've added a third transistor switch  $Q_3$ , whose job is to hold the collector of  $Q_1$  at ground once the output pulse begins, regardless of what the input signal does. Now any positive input pulse – whether longer or shorter than the desired output pulse width – produces the same output pulse width; look at the waveforms in the figure. Note that we've chosen  $R_5$  relatively large to minimize output loading while still ensuring full saturation of  $Q_3$ .

**Exercise 2.3.** Elaborate on this last statement: what is the output voltage during the pulse, slightly reduced owing to the loading effect of  $R_5$ ? What is the minimum required beta of  $Q_3$  to guarantee its saturation during the output pulse?

#### E. Pulse generator – III

For our final act, let's fix a deficiency of these circuits, namely a tendency for the output pulse to turn off somewhat slowly. That happens because  $Q_2$ 's base voltage, with its leisurely  $100\mu s$   $RC$  time constant, rises smoothly (and relatively slowly) through the turn-on voltage threshold of  $\approx 0.6$  V. Note, by the way, that this problem does not occur at the turn-on of the output pulse, because at that transition  $Q_2$ 's base voltage drops abruptly down to approximately  $-4.4$  V, owing to the sharp input step waveform, which is further sharpened by the switching action of  $Q_1$ .

The cure here is to add at the output a clever circuit known as a *Schmitt trigger*, shown in its transistor implementation<sup>14</sup> in Figure 2.13A. It works like this: imagine a time within the positive output pulse of the previous circuits, so the input to this new Schmitt circuit is high (near

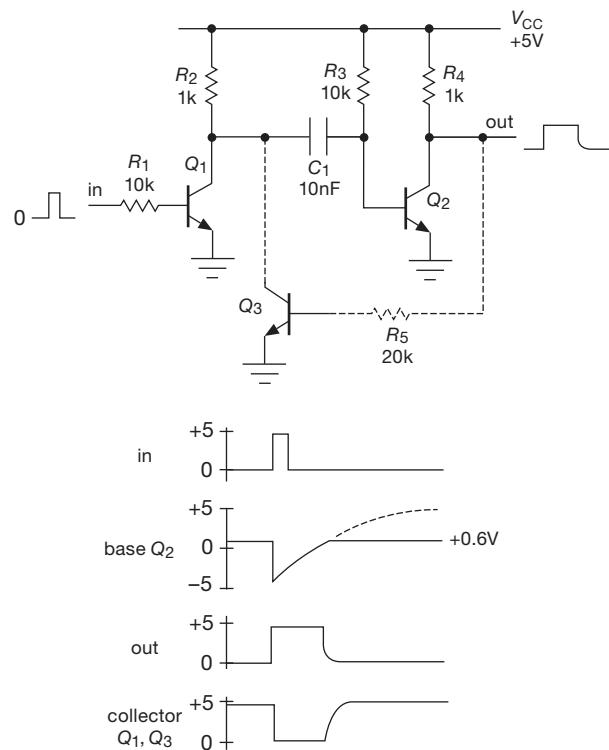


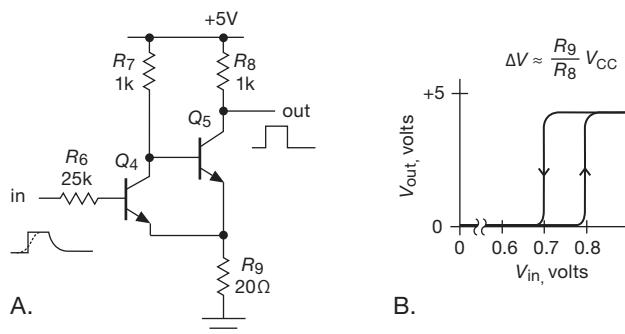
Figure 2.12. Generating a short pulse from a step or pulse input.

+5 V). That holds  $Q_4$  in saturation, and so  $Q_5$  is cut off, with the output at +5 V. The emitter current of  $Q_4$  is about 5 mA, so the emitter voltage is approximately +100 mV; the base is a  $V_{BE}$  higher, approximately +700 mV.

Now imagine the trailing edge of the input pulse waveform, whose voltage smoothly drops toward ground. As it drops below 700 mV,  $Q_4$  begins to turn off, so its collector voltage rises. If this were a simple transistor switch (i.e., if  $Q_5$  were absent) the collector would rise to +5 V; here, however, the collector resistor  $R_7$  instead supplies current to  $Q_5$ , putting it in saturation. So  $Q_5$ 's collector drops nearly to ground.

At this simple level of analysis the circuit appears to be pretty useless, because its output is the same as its input! Let's look a little closer, though: as the input voltage drops through the 700 mV threshold and  $Q_5$  turns on, the total emitter current rises to  $\approx 10$  mA (5 mA from  $Q_5$ 's collector current, and another  $\approx 5$  mA from its base current, both of which flow out the emitter). The drop across the emitter resistor is now 200 mV, which means that the input threshold has increased to about +800 mV. So the input voltage, which had just dropped below 700 mV, now finds itself well below the new threshold, causing the

<sup>14</sup> We'll see other ways of making a Schmitt trigger, using op-amps or comparators, in Chapter 4.



**Figure 2.13.** A “Schmitt trigger” produces an output with abrupt transitions, regardless of the speed of the input waveform.

output to switch abruptly. This “regenerative” action is how the Schmitt trigger turns a slowly moving waveform into an abrupt transition.

A similar action occurs as the input rises through this higher threshold; see Figure 2.13B, which illustrates how the output voltage changes as the input voltage passes through the two thresholds, an effect known as *hysteresis*. The Schmitt trigger produces rapid output transitions as the input passes through either threshold. We’ll see Schmitt triggers again in Chapters 4 and 10.

There are many enjoyable applications of transistor switches, including “signal” applications like this (combined with more complex digital logic circuits), as well as “power switching” circuits in which transistors operating at high currents, high voltages, or both, are used to control hefty loads, perform power conversion, and so on. Transistor switches can also be used as substitutes for mechanical switches when we are dealing with continuous (“linear” or “analog”) waveforms. We’ll see examples of these in the next chapter, when we deal with FETs, which are ideally suited to such switching tasks, and again in Chapter 12, where we deal with the control of signals and external loads from logic-level signals.

We now move on to consider the first of several *linear* transistor circuits.

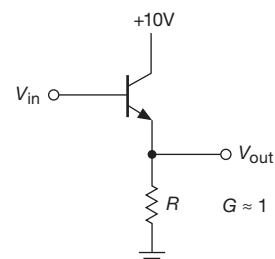
### 2.2.3 Emitter follower

Figure 2.14 shows an example of an *emitter follower*. It is called that because the output terminal is the emitter, which follows the input (the base), less one diode drop:

$$V_E \approx V_B - 0.6 \text{ volts.}$$

The output is a replica of the input, but 0.6 to 0.7 V less positive. For this circuit,  $V_{in}$  must stay at +0.6 V or more, or else the output will sit at ground. By returning the emit-

ter resistor to a negative supply voltage, you can permit negative voltage swings as well. Note that there is no collector resistor in an emitter follower.



**Figure 2.14.** Emitter follower.

At first glance this circuit may appear quite thoroughly useless, until you realize that the input impedance is much larger than the output impedance, as will be demonstrated shortly. This means that the circuit requires less power from the signal source to drive a given load than would be the case if the signal source were to drive the load directly. Or a signal of some internal impedance (in the Thévenin sense) can now drive a load of comparable or even lower impedance without loss of amplitude (from the usual voltage-divider effect). In other words, an emitter follower has current gain, even though it has no voltage gain. It has *power* gain. Voltage gain isn’t everything!

#### A. Impedances of sources and loads

This last point is very important and is worth some more discussion before we calculate in detail the beneficial effects of emitter followers. In electronic circuits, you’re always hooking the output of something to the input of something else, as suggested in Figure 2.15. The signal source might be the output of an amplifier stage (with Thévenin equivalent series impedance  $Z_{out}$ ), driving the next stage or perhaps a load (of some input impedance  $Z_{in}$ ). In general, the loading effect of the following stage causes a reduction of signal, as we discussed earlier in §1.2.5A. For this reason it is usually best to keep  $Z_{out} \ll Z_{in}$  (a factor of 10 is a comfortable rule of thumb).

In some situations it is OK to forgo this general goal of making the source stiff compared with the load. In particular, if the load is always connected (e.g., within a circuit) and if it presents a known and constant  $Z_{in}$ , it is not too serious if it “loads” the source. However, it is always nicer if signal levels don’t change when a load is connected. Also, if  $Z_{in}$  varies with signal level, then having a stiff source

$(Z_{\text{out}} \ll Z_{\text{in}})$  ensures linearity, where otherwise the level-dependent voltage divider would cause distortion.<sup>15</sup>

Finally, as we remarked in §1.2.5A, there are two situations in which  $Z_{\text{out}} \ll Z_{\text{in}}$  is actually the wrong thing to do: in radiofrequency circuits we usually *match* impedances ( $Z_{\text{out}} = Z_{\text{in}}$ ), for reasons we'll describe in Appendix H. A second exception applies if the signal being coupled is a *current* rather than a voltage. In that case the situation is reversed, and we strive to make  $Z_{\text{in}} \ll Z_{\text{out}}$  ( $Z_{\text{out}} = \infty$ , for a current source).

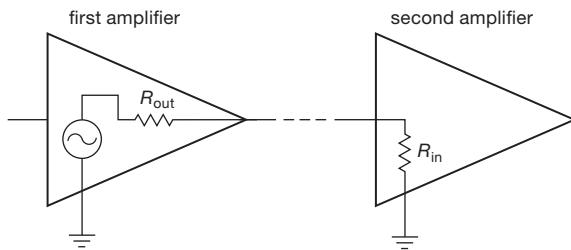


Figure 2.15. Illustrating circuit “loading” as a voltage divider.

## B. Input and output impedances of emitter followers

As we've stated, the emitter follower is useful for changing impedances of signals or loads. To put it starkly, that's really the whole point of an emitter follower.

Let's calculate the input and output impedances of the emitter follower. In the preceding circuit we consider  $R$  to be the load (in practice it sometimes *is* the load; otherwise the load is in parallel with  $R$ , but with  $R$  dominating the parallel resistance anyway). Make a voltage change  $\Delta V_B$  at the base; the corresponding change at the emitter is  $\Delta V_E = \Delta V_B$ . Then the change in emitter current is

$$\Delta I_E = \Delta V_B / R,$$

so

$$\Delta I_B = \frac{1}{\beta + 1} \Delta I_E = \frac{\Delta V_B}{R(\beta + 1)}$$

(using  $I_E = I_C + I_B$ ). The input resistance is  $\Delta V_B / \Delta I_B$ . Therefore

$$r_{\text{in}} = (\beta + 1)R. \quad (2.2)$$

The transistor small-signal (or “incremental”) current gain

<sup>15</sup> We use the boldface symbol  $\mathbf{Z}$  when the complex nature of impedance is important. In common usage the term “impedance” can refer loosely to the *magnitude* of impedance, or even to a purely real impedance (e.g., transmission-line impedance); for such instances we use the ordinary math-italic symbol  $Z$ .

$(\beta, \text{ or } h_{\text{fe}})$  is typically about 100, so a low-impedance load looks like a much higher impedance at the base; it is easier to drive.

In the preceding calculation we used the *changes* in the voltages and currents, rather than the steady (dc) values of those voltages (or currents), to arrive at our input resistance  $r_{\text{in}}$ . Such a “small-signal” analysis is used when the variations represent a possible signal, as in an audio amplifier, riding on a steady dc “bias” (see §2.2.7). Although we indicated changes in voltage and current explicitly (with “ $\Delta V$ ,” etc.), the usual practice is to use lowercase symbols for small-signal variations (thus  $\Delta V \leftrightarrow v$ ); with this convention the above equation for  $\Delta I_E$ , for example, would read  $i_E = v_B / R$ .

The distinction between dc current gain ( $h_{\text{FE}}$ ) and small-signal current gain ( $h_{\text{fe}}$ ) isn't always made clear, and the term beta is used for both. That's alright, since  $h_{\text{fe}} \approx h_{\text{FE}}$  (except at very high frequencies), and you never assume you know them accurately, anyway.

Although we used resistances in the preceding derivation, we could generalize to complex impedances by allowing  $\Delta V_B$ ,  $\Delta I_B$ , etc., to become complex numbers. We would find that the same transformation rule applies for impedances:

$$\mathbf{Z}_{\text{in}} = (\beta + 1)\mathbf{Z}_{\text{load}}. \quad (2.3)$$

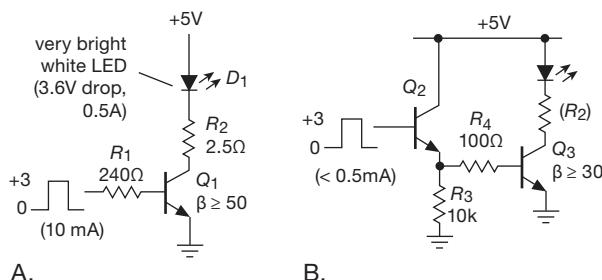
We could do a similar calculation to find that the output impedance  $\mathbf{Z}_{\text{out}}$  of an emitter follower (the impedance looking into the emitter) driven from a source of internal impedance  $\mathbf{Z}_{\text{source}}$  is given by

$$\mathbf{Z}_{\text{out}} = \frac{\mathbf{Z}_{\text{source}}}{\beta + 1}. \quad (2.4)$$

Strictly speaking, the output impedance of the circuit should also include the parallel resistance of  $R$ , but in practice  $\mathbf{Z}_{\text{out}}$  (the impedance looking into the emitter) dominates.

**Exercise 2.4.** Show that the preceding relationship is correct. Hint: hold the source voltage fixed and find the change in output current for a given forced change in output voltage. Remember that the source voltage is connected to the base through a series resistor.

Because of these nice properties, emitter followers find application in many situations, e.g., making low-impedance signal sources within a circuit (or at outputs), making stiff voltage references from higher-impedance references (formed from voltage dividers, say), and generally isolating signal sources from the loading effects of subsequent stages.



**Figure 2.16.** Putting an emitter follower in front of a switch makes it easy for a low-current control signal to switch a high-current load.

**Exercise 2.5.** Use a follower with the base driven from a voltage divider to provide a stiff source of +5 volts from an available regulated +15 V supply. Load current (max) = 25 mA. Choose your resistor values so that the output voltage doesn't drop more than 5% under full load.

### C. Follower drives switch

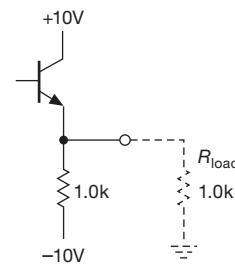
Figure 2.16 shows a nice example of an emitter follower rescuing an awkward circuit. We're trying to switch a really bright white LED (the kind you use for "area lighting"), which drops about 3.6 V at its desired 500 mA of forward current. And we've got a 0–3 V digital logic signal available to control the switch. The first circuit uses a single *npn* saturated switch, with a base resistor sized to produce 10 mA of base current, and a  $2.5\Omega$  current-limiting resistor in series with the LED.

This circuit is OK, sort of. But it draws an uncomfortably large current from the control input; and it requires  $Q_1$  to have plenty of current gain at the full load current of 0.5 A. In the second circuit (Figure 2.16B) an emitter follower has come to the rescue, greatly reducing the input current (because of its current gain), and at the same time relaxing the minimum beta requirement of the switch ( $Q_3$ ). To be fair, we should point out that a low-threshold MOSFET provides an even simpler solution here; we'll tell you how, in Chapters 3 and 12.

### D. Important points about followers

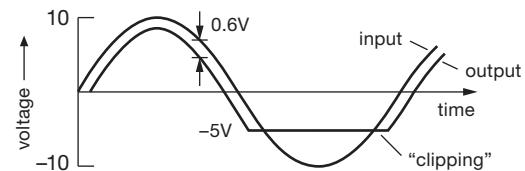
**Current flow in one direction only.** Notice (§2.1.1, rule

4) that in an emitter follower the *npn* transistor can only *source* (as opposed to *sink*) current. For instance, in the loaded circuit shown in Figure 2.17 the output can swing to within a transistor saturation voltage drop of  $V_{CC}$  (about +9.9 V), but it cannot go more negative than -5 volts. That is because on the extreme negative swing, the transistor can do no better than to turn off completely, which it does at -4.4 volts input (-5 V out-



**Figure 2.17.** An *npn* emitter follower can source plenty of current through the transistor, but can sink limited current only through its emitter resistor.

put, set by the divider formed by the load and emitter resistors). Further negative swing at the input results in back-biasing of the base-emitter junction, but no further change in output. The output, for a 10 volt amplitude sinewave input, looks as shown in Figure 2.18.



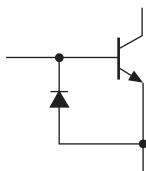
**Figure 2.18.** Illustrating the asymmetrical current drive capability of the *npn* emitter follower.

Another way to view the problem is to say that the emitter follower has a low value of *small-signal* output impedance, whereas its *large-signal* output impedance is much higher (as large as  $R_E$ ). The output impedance changes over from its *small-signal* value to its *large-signal* value at the point where the transistor goes out of the active region (in this case at an output voltage of -5 V). To put this point another way, a low value of *small-signal* output impedance doesn't necessarily mean that the circuit can generate large signal swings into a low resistance load. A low *small-signal* output impedance doesn't imply a large output current capability.

Possible solutions to this problem involve either decreasing the value of the emitter resistor (with greater power dissipation in resistor and transistor), using a *pnp* transistor (if all signals are negative only), or using a "push-pull" configuration, in which two complementary transistors (one *npn*, one *pnp*) are used (§2.4.1). This sort of problem can also come up when the load that an emitter follower is driving contains voltage or current sources of its own, and thus can force a current in the "wrong" direction. This happens most often with

regulated power supplies (the output is usually an emitter follower) driving a circuit that has other power supplies.

**Base-emitter breakdown.** Always remember that the base-emitter reverse breakdown voltage for silicon transistors is small, quite often as little as 6 volts. Input swings large enough to take the transistor out of conduction can easily result in breakdown (causing permanent degradation of current gain  $\beta$ ) unless a protective diode is added (Figure 2.19).



**Figure 2.19.** A diode prevents base-emitter reverse voltage breakdown.

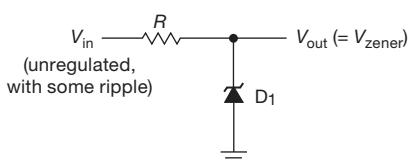
**Gain is slightly less than unity.** The voltage gain of an emitter follower is actually slightly less than 1.0, because the base-emitter voltage drop is not really constant, but depends slightly on collector current. You will see how to handle that later in the chapter, when we have the Ebers-Moll equation.

#### 2.2.4 Emitter followers as voltage regulators

The simplest regulated supply of voltage is simply a zener (Figure 2.20). Some current must flow through the zener, so you choose

$$\frac{V_{in}(\min) - V_{out}}{R} > I_{out}(\max).$$

Because  $V_{in}$  isn't regulated, you use the lowest value of  $V_{in}$  that might occur. Designing for satisfactory operation under the worst combination (here minimum  $V_{in}$  and maximum  $I_{out}$ ) is known as "worst-case" design. In practice, you would also worry about component tolerances, line-voltage limits, etc., designing to accommodate the worst possible combination that would ever occur.



**Figure 2.20.** Simple zener voltage regulator.

The zener must be able to dissipate

$$P_{zener} = \left( \frac{V_{in} - V_{out}}{R} - I_{out} \right) V_{zener}.$$

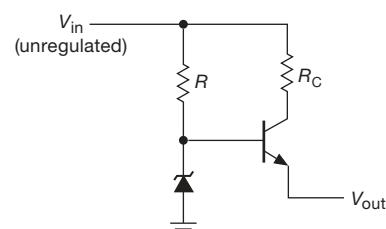
Again, for worst-case design, you would use  $V_{in}(\max)$  and  $I_{out}(\min)$ .

**Exercise 2.6.** Design a +10V regulated supply for load currents from 0 to 100 mA; the input voltage is +20 to +25 V. Allow at least 10 mA zener current under all (worst-case) conditions. What power rating must the zener have?

This simple zener-regulated supply is sometimes used for noncritical circuits or circuits using little supply current. However, it has limited usefulness, for several reasons:

- $V_{out}$  isn't adjustable or settable to a precise value.
- Zener diodes give only moderate ripple rejection and regulation against changes of input or load, owing to their finite dynamic impedance.
- For widely varying load currents a high-power zener is often necessary to handle the dissipation at low load current.<sup>16</sup>

By using an emitter follower to isolate the zener, you get the improved circuit shown in Figure 2.21. Now the situation is much better. Zener current can be made relatively independent of load current, since the transistor base current is small, and far lower zener power dissipation is possible (reduced by as much as a factor of  $\beta$ ). The collector resistor  $R_C$  can be added to protect the transistor from momentary output short circuits by limiting the current, even though it is not essential to the emitter follower function. Choose  $R_C$  so that the voltage drop across it is less than the drop across  $R$  for the highest normal load current (i.e., so that the transistor does not saturate at maximum load).



**Figure 2.21.** Zener regulator with follower, for increased output current.  $R_C$  protects the transistor by limiting maximum output current.

<sup>16</sup> This is a property shared by all *shunt regulators*, of which the zener is the simplest example.

**Exercise 2.7.** Design a +10 V supply with the same specifications as in Exercise 2.6. Use a zener and emitter follower. Calculate worst-case dissipation in transistor and zener. What is the percentage change in zener current from the no-load condition to full load? Compare with your previous circuit.

A nice variation of this circuit aims to eliminate the effect of ripple current (through  $R$ ) on the zener voltage by supplying the zener current from a current source, which is the subject of §2.2.6. An alternative method uses a lowpass filter in the zener bias circuit (Figure 2.22).  $R$  is chosen such that the series pair provides sufficient zener current. Then  $C$  is chosen large enough so that  $RC \gg 1/f_{\text{ripple}}$ .<sup>17</sup>

Later you will see better voltage regulators, ones in which you can vary the output easily and continuously by using feedback. They are also better voltage sources, with output impedances measured in milliohms, temperature coefficients of a few parts per million per degree centigrade, and other desirable features.

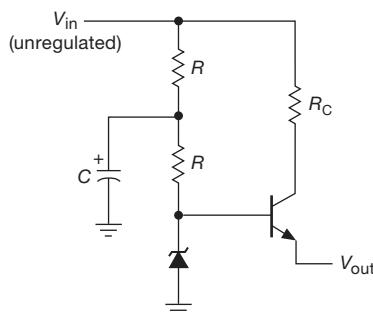


Figure 2.22. Reducing ripple in the zener regulator.

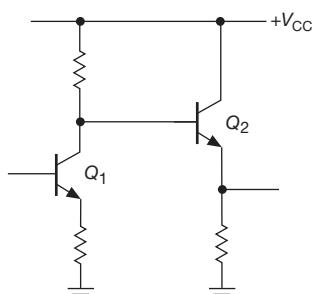


Figure 2.23. Biasing an emitter follower from a previous stage.

## 2.2.5 Emitter follower biasing

When an emitter follower is driven from a preceding stage in a circuit, it is usually OK to connect its base directly to the previous stage's output, as shown in Figure 2.23.

<sup>17</sup> In a variation of this circuit, the upper resistor is replaced with a diode.

Because the signal on  $Q_1$ 's collector is always within the range of the power supplies,  $Q_2$ 's base will be between  $V_{CC}$  and ground, and therefore  $Q_2$  is in the active region (neither cut off nor saturated), with its base-emitter diode in conduction and its collector at least a few tenths of a volt more positive than its emitter. Sometimes, though, the input to a follower may not be so conveniently situated with respect to the supply voltages. A typical example is a capacitively coupled (or ac-coupled) signal from some external source (e.g., an audio signal input to a stereo amplifier). In that case the signal's average voltage is zero, and direct coupling to an emitter follower will give an output like that in Figure 2.24.

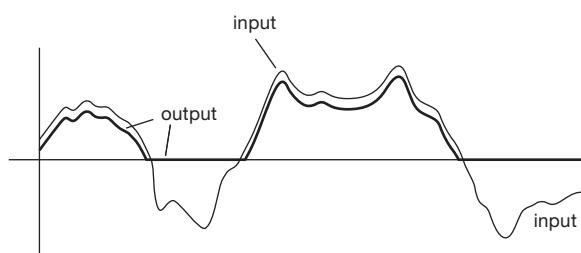
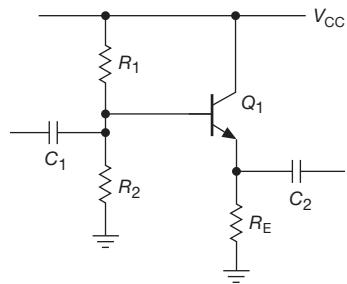


Figure 2.24. A transistor amplifier powered from a single positive supply cannot generate negative voltage swings at the transistor output terminal.

It is necessary to *bias* the follower (in fact, any transistor amplifier) so that collector current flows during the entire signal swing. In this case a voltage divider is the simplest way (Figure 2.25).  $R_1$  and  $R_2$  are chosen to put the base halfway between ground and  $V_{CC}$  when there is no input signal, i.e.,  $R_1$  and  $R_2$  are approximately equal. The process of selecting the operating voltages in a circuit, in the absence of applied signals, is known as setting the *quiescent point*. In this case, as in most cases, the quiescent point is chosen to allow maximum symmetrical signal swing of the output waveform without *clipping* (flattening of the top or bottom of the waveform). What values should  $R_1$  and  $R_2$  have? Applying our general principle (§1.2.5A, §2.2.3A), we make the impedance of the dc bias source (the impedance looking into the voltage divider) small compared with the load it drives (the dc impedance looking into the base of the follower). In this case,

$$R_1 \parallel R_2 \ll \beta R_E.$$

This is approximately equivalent to saying that the current flowing in the voltage divider should be large compared with the current drawn by the base.



**Figure 2.25.** An ac-coupled emitter follower. Note base bias voltage divider.

### A. Emitter follower design example

As an actual design example, let's make an emitter follower for audio signals (20 Hz to 20 kHz).  $V_{CC}$  is +15 V, and quiescent current is to be 1 mA.

**Step 1. Choose  $V_E$ .** For the largest possible symmetrical swing without clipping,  $V_E = 0.5V_{CC}$ , or +7.5 volts.

**Step 2. Choose  $R_E$ .** For a quiescent current of 1 mA,  $R_E = 7.5k$ .

**Step 3. Choose  $R_1$  and  $R_2$ .**  $V_B$  is  $V_E + 0.6$  V, or 8.1 V. This determines the ratio of  $R_1$  to  $R_2$  as 1:1.17. The preceding loading criterion requires that the parallel resistance of  $R_1$  and  $R_2$  be about 75k or less (one-tenth of  $7.5k \times \beta$ ). Suitable standard values are  $R_1 = 130k$ ,  $R_2 = 150k$ .

**Step 4. Choose  $C_1$ .** The capacitor  $C_1$  forms a high-pass filter with the impedance it sees as a load, namely the impedance looking into the base in parallel with the impedance looking into the base voltage divider. If we assume that the load this circuit will drive is large compared with the emitter resistor, then the impedance looking into the base is  $\beta R_E$ , about 750k. The divider looks like 70k. So the capacitor sees a load of about 63k, and it should have a value of at least  $0.15 \mu F$  so that the 3 dB point will be below the lowest frequency of interest, 20 Hz.

**Step 5. Choose  $C_2$ .** The capacitor  $C_2$  forms a high-pass filter in combination with the load impedance, which is unknown. However, it is safe to assume that the load impedance won't be smaller than  $R_E$ , which gives a value for  $C_2$  of at least  $1.0 \mu F$  to put the 3 dB point below 20 Hz. Because there are now two cascaded highpass filter sections, the capacitor values should be increased somewhat to prevent excessive attenuation (reduction of signal amplitude, in this case 6 dB) at the lowest frequency of interest.  $C_1 = 0.47 \mu F$  and  $C_2 = 3.3 \mu F$  might be good choices.<sup>18</sup>

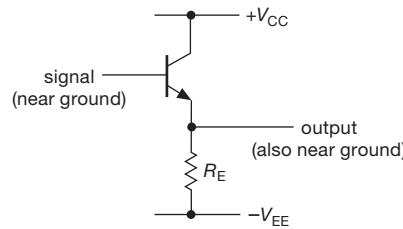
From our simple transistor model, the output impedance at the emitter is just  $Z_{out} = R_E \parallel [(\mathbf{Z}_{in} \parallel R_1 \parallel R_2) / \beta]$ , where  $\mathbf{Z}_{in}$  is the (Thévenin) output resistance of the signal that drives this circuit. So, taking  $\beta \approx 100$ , a signal source with  $10 k\Omega$  output resistance would result in an output impedance (at the emitter) of about  $87 \Omega$ . As we'll see later in the chapter (§2.3), there's an effect (the intrinsic emitter impedance,  $r_e$ ) that adds an additional resistance of  $0.025/I_E$  effectively in series with the emitter; so the output impedance here (with  $10 k\Omega$  source) would be about  $110 \Omega$ .

### B. Followers with split supplies

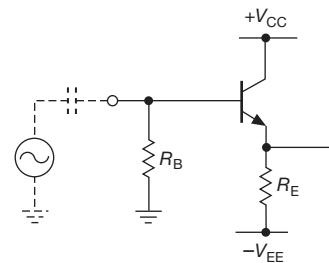
Because signals often are “near ground,” it is convenient to use symmetrical positive and negative supplies. This simplifies biasing and eliminates coupling capacitors (Figure 2.26).

*Warning:* you must always provide a dc path for base bias current, even if it goes only to ground. In this circuit it is assumed that the signal source has a dc path to ground. If not (e.g., if the signal is capacitively coupled), you must provide a resistor to ground (Figure 2.27).  $R_B$  could be about one-tenth of  $\beta R_E$ , as before.

**Exercise 2.8.** Design an emitter follower with  $\pm 15$  V supplies to operate over the audio range (20 Hz to 20 kHz). Use 5 mA quiescent current and capacitive input coupling.



**Figure 2.26.** A dc-coupled emitter follower with split supply.



**Figure 2.27.** Always provide a dc bias path.

<sup>18</sup> These values may seem curiously “unround.” But they are chosen from the widely available EIA “E6” decade values (see Appendix C); and in fact “round-number” values of  $0.5 \mu F$  and  $3.0 \mu F$  are harder to find.

### C. Bad biasing

You sometimes see sadness-inducing circuits like the disaster shown in Figure 2.28. The designer chose  $R_B$  by assuming a particular value for beta (100), estimating the base current, and then hoping for a 7 V drop across  $R_B$ . This is a bad design; beta is not a good parameter and will vary considerably. By using voltage biasing with a stiff voltage divider, as in the detailed example presented earlier, the quiescent point is insensitive to variations in transistor beta. For instance, in the previous design example the emitter voltage will increase by only 0.35 V (5%) for a transistor with  $\beta = 200$  instead of the nominal  $\beta = 100$ . And, as with this emitter follower example, it is just as easy to fall into this trap and design bad transistor circuits in the other transistor configurations (notably the common-emitter amplifier, which we will treat later in this chapter).

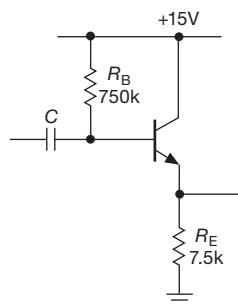


Figure 2.28. Don't do this!

### D. Cancelling the offset – I

Wouldn't it be nice if an emitter follower did not cause an offset of the output signal by the  $V_{BE} \approx 0.6$  V base–emitter drop? Figure 2.29 shows how to cancel the dc offset, by cascading a *pnp* follower (which has a positive  $V_{BE}$  offset) with an *npn* follower (which has a comparable negative  $V_{BE}$  offset). Here we've configured the circuit with  $\pm 10$  V symmetrical split supplies; and we've used equal-value emitter resistors so that the two transistors have a comparable quiescent current for an input signal near 0 V.

This is a nice trick, useful to know about and often helpful. But the cancellation isn't perfect, for reasons we'll see later in the chapter ( $V_{BE}$  depends somewhat on collector current, and on transistor size, §2.3), and again in Chapter 5. But, as we'll see in Chapter 4, it is in fact rather easy to make a follower, using *operational amplifiers*, with nearly perfect zero offset (10  $\mu$ V or less); and as a bonus you get input impedances in the gigaohms (or more), input currents in the nanoamps (or less), and output impedances

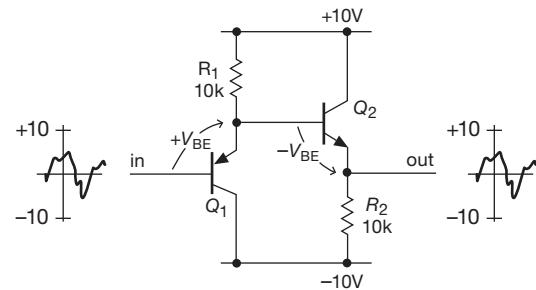


Figure 2.29. Cascading a *pnp* and an *npn* follower produces approximate cancellation of the  $V_{BE}$  offsets.

measured in fractions of an ohm. Take a look ahead at Chapter 4.

### 2.2.6 Current source

Current sources, although often neglected, are as important and as useful as voltage sources. They often provide an excellent way to bias transistors, and they are unequalled as “active loads” for super-gain amplifier stages and as emitter sources for differential amplifiers. Integrators, sawtooth generators, and ramp generators need current sources. They provide wide-voltage-range pullups within amplifier and regulator circuits. And, finally, there are applications in the outside world that require constant current sources, e.g., electrophoresis or electrochemistry.

#### A. Resistor plus voltage source

The simplest approximation to a current source is shown in Figure 2.30. As long as  $R_{load} \ll R$  (in other words,  $V_{load} \ll V$ ), the current is nearly constant and is approximately

$$I \approx V/R.$$

The load doesn't have to be resistive. A capacitor will charge at a constant rate, as long as  $V_{cap} \ll V$ ; this is just the first part of the exponential charging curve of an  $RC$ .

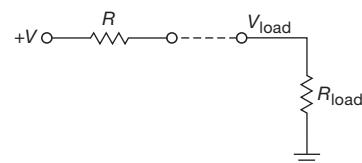


Figure 2.30. Current-source approximation.

There are several drawbacks to a simple resistor current source. To make a good approximation to a current source, you must use large voltages, with lots of power dissipation in the resistor. In addition, the current isn't easily

*programmable*, i.e., controllable over a large range by means of a voltage somewhere else in the circuit.

**Exercise 2.9.** If you want a current source constant to 1% over a load voltage range of 0 to +10 volts, how large a voltage source must you use in series with a single resistor?

**Exercise 2.10.** Suppose you want a 10 mA current in the preceding problem. How much power is dissipated in the series resistor? How much gets to the load?

### B. Transistor current source

Happily, it is possible to make a very good current source with a transistor (Figure 2.31). It works like this: applying  $V_B$  to the base, with  $V_B > 0.6$  V, ensures that the emitter is always conducting:

$$V_E = V_B - 0.6 \text{ volts.}$$

So

$$I_E = V_E / R_E = (V_B - 0.6 \text{ volts}) / R_E.$$

But, since  $I_E \approx I_C$  for large beta,

$$I_C \approx (V_B - 0.6 \text{ volts}) / R_E, \quad (2.5)$$

independent of  $V_C$ , as long as the transistor is not saturated ( $V_C \gtrsim V_E + 0.2$  volts).

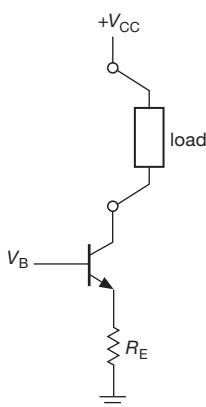


Figure 2.31. Transistor current source: basic concept.

### C. Current-source biasing

The base voltage can be provided in a number of ways. A voltage divider is OK, as long as it is stiff enough. As before, the criterion is that its impedance should be much less than the dc impedance looking into the base ( $\beta R_E$ ). Or you can use a zener diode (or a two-terminal IC reference like the LM385), biased from  $V_{CC}$ , or even a few forward-

biased diodes<sup>19</sup> in series from base to the corresponding emitter supply. Figure 2.32 shows some examples. In the last example (Figure 2.32C), a *pnp* transistor *sources* current to a load returned to ground. The other examples (using *npn* transistors) should properly be called current *sinks*, but the usual practice is to refer to them all loosely as “current sources.”<sup>20</sup> In the first circuit, the voltage-divider impedance of  $\sim 1.3\text{k}$  is stiff compared with the impedance looking into the base of about  $100\text{k}$  (for  $\beta = 100$ ), so any changes in beta with collector voltage will not much affect the output current by causing the base voltage to change. In the other two circuits the biasing resistors are chosen to provide several millamps to bring the diodes into conduction.

### D. Compliance

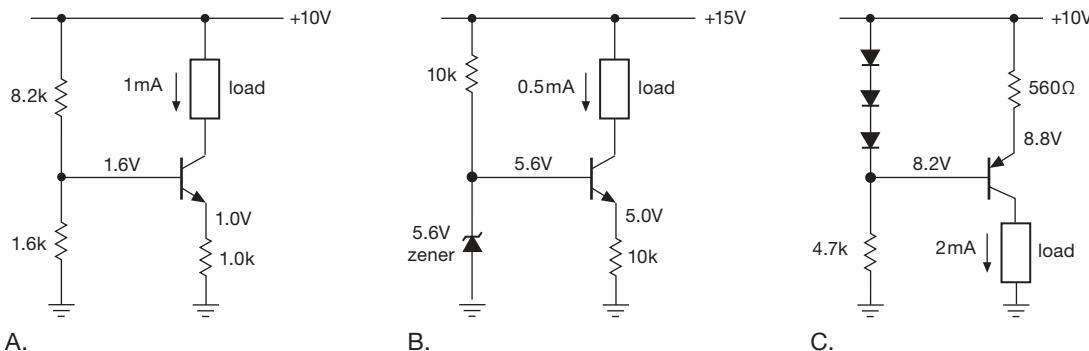
A current source can provide constant current to the load only over some finite range of load voltage. To do otherwise would be equivalent to providing infinite power. The output voltage range over which a current source behaves well is called its output *compliance*. For the preceding transistor current sources, the compliance is set by the requirement that the transistors stay in the active region. Thus in the first circuit the voltage at the collector can go down until the transistor is almost in saturation, perhaps +1.1 V at the collector. The second circuit, with its higher emitter voltage, can sink current down to a collector voltage of about +5.1 V.

In all cases the collector voltage can range from a value near saturation all the way up to the supply voltage. For example, the last circuit can source current to the load for any voltage between zero and about +8.6 V across the load. In fact, the load might even contain batteries or power supplies of its own, which could carry the collector beyond the supply voltage (Figure 2.32A,B) or below ground (Figure 2.32C). That’s OK, but you must watch out for transistor breakdown ( $V_{CE}$  must not exceed  $BV_{CEO}$ , the specified collector-emitter breakdown voltage) and also for excessive power dissipation (set by  $I_C V_{CE}$ ). As you will see in §§3.5.1B, 3.6.4C, and 9.4.2, there is an additional safe-operating-area constraint on power transistors.

**Exercise 2.11.** You have +5 and +15 V regulated supplies available in a circuit. Design a 5 mA *npn* current sink using the +5 V to bias the base. What is the output compliance?

<sup>19</sup> A red LED, with its forward voltage drop of  $\approx 1.6$  V, is a convenient substitute for a string of three diodes.

<sup>20</sup> “Sink” and “source” simply refer to the direction of current flow: if a circuit *supplies* (positive) current to a point, it is a *source*, and vice versa.

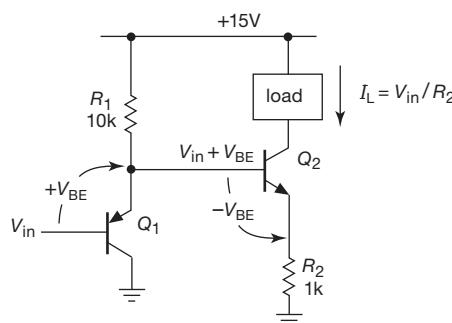


**Figure 2.32.** Transistor current-source circuits, illustrating three methods of base biasing; *npn* transistors sink current, whereas *pnp* transistors source current. The circuit in C illustrates a load returned to ground. See also Figure 3.26.

A current source doesn't have to have a fixed voltage at the base. By varying  $V_B$  you get a voltage-programmable current source. The input signal swing  $v_{in}$  (recall that lowercase symbols mean variations) must stay small enough so that the emitter voltage never drops to zero, if the output current is to reflect input-voltage variations smoothly. The result will be a current source with variations in output current proportional to the variations in input voltage,  $i_{out} = v_{in}/R_E$ . This is the basis of the amplifier we'll see next (§2.2.7).

### E. Cancelling the offset – II

It's a minor drawback of these current source circuits that you have to apply a base voltage that is offset by  $V_{BE} \approx 0.6$  V from the voltage that you want to appear across the emitter resistor; and it is of course the latter that sets the output current. It's the same offset issue as with an emitter follower; and you can use the same trick (§2.2.5D) to bring about approximate cancellation of the offset in situations in which that is a problem.



**Figure 2.33.** Compensating the  $V_{BE}$  drop in a current source.

Look at Figure 2.33. It has our standard current-source output stage  $Q_2$ , with the current set by the voltage across the emitter resistor:  $I_L = V_E/R_2$ . So  $Q_2$ 's base needs to be

a  $V_{BE}$  higher (the offset), but that's just what the *pnp* input follower does anyway. So, voilà, the voltage at  $Q_2$ 's emitter winds up being approximately equal to  $V_{in}$  that you apply; and so the output current is simply  $I_L = V_{in}/R_2$ , with no ifs, ands, or  $V_{BE}$  offsets. Cute!

We hasten to point out, though, that this is not a particularly accurate cancellation, because the two transistors will in general have different collector currents, and therefore somewhat different base–emitter drops (§2.3). But it's a first-order hack, and a lot better than nothing. And, once again, the magic of operational amplifiers (Chapter 4) will provide a way to make current sources in which the output current is accurately programmed by an input voltage, without that pesky  $V_{BE}$  offset.

### F. Deficiencies of current sources

These transistor current-source circuits perform well, particularly when compared with a simple resistor biased from a fixed voltage (Figure 2.30). When you look closely, though, you find that they do depart from the ideal at some level of scrutiny – that is, the load current does show some (relatively small) variation with voltage. Another way to say the same thing is that the current source has a finite ( $R_{Th} < \infty$ ) Thévenin equivalent resistance.

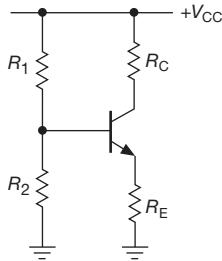
We discuss the causes of these deficiencies, and some very clever circuit fixes, later in the chapter, and also in Chapter 2x.

### 2.2.7 Common-emitter amplifier

Consider a current source with a resistor as load (Figure 2.34). The collector voltage is

$$V_C = V_{CC} - I_C R_C$$

We could capacitively couple a signal to the base to cause the collector voltage to vary. Consider the example in



**Figure 2.34.** Current source driving a resistor as load: an amplifier!

Figure 2.35. Blocking capacitor  $C$  is chosen so that all frequencies of interest are passed by the highpass filter it forms in combination with the parallel resistance of the base biasing resistors<sup>21</sup>; that is,

$$C \geq \frac{1}{2\pi f(R_1 \parallel R_2)}.$$

The quiescent collector current is 1.0 mA because of the applied base bias and the 1.0k emitter resistor. That current puts the collector at +10 volts (+20 V, minus 1.0 mA through 10k). Now imagine an applied wiggle in base voltage  $v_B$ . The emitter follows with  $v_E = v_B$ , which causes a wiggle in emitter current

$$i_E = v_E/R_E = v_B/R_E$$

and nearly the same change in collector current ( $\beta$  is large). So the initial wiggle in base voltage finally causes a collector voltage wiggle

$$v_C = -i_C R_C = -v_B (R_C/R_E)$$

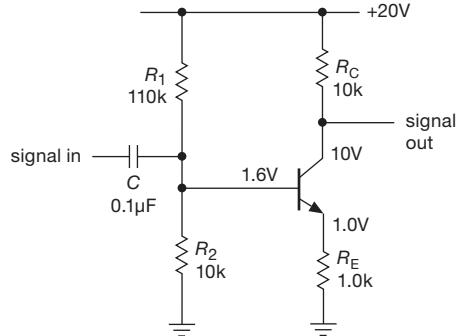
Aha! It's a *voltage amplifier*, with a voltage amplification (or "gain") given by

$$\text{gain} = v_{\text{out}}/v_{\text{in}} = -R_C/R_E \quad (2.6)$$

In this case the gain is  $-10,000/1000$ , or  $-10$ . The minus sign means that a positive wiggle at the input gets turned into a negative wiggle (10 times as large) at the output. This is called a *common-emitter amplifier* with emitter degeneration.

#### A. Input and output impedances of the common-emitter amplifier

We can easily determine the input and output impedances of the amplifier. The input signal sees, in parallel, 110k, 10k, and the impedance looking into the base. The latter is



**Figure 2.35.** An ac common-emitter amplifier with emitter degeneration. Note that the output terminal is the collector rather than the emitter.

about 100k ( $\beta$  times  $R_E$ ), so the input impedance (dominated by the 10k) is about 8k. The input coupling capacitor thus forms a highpass filter, with the 3 dB point at 200 Hz. The signal driving the amplifier sees  $0.1 \mu\text{F}$  in series with 8k, which to signals of normal frequencies (well above the 3 dB point) just looks like 8k.

The output impedance is 10k in parallel with the impedance looking into the collector. What is that? Well, remember that if you snip off the collector resistor, you're simply looking into a current source. The collector impedance is very large (measured in megohms), and so the output impedance is just the value of the collector resistor, 10k. It is worth remembering that the impedance looking into a transistor's collector is high, whereas the impedance looking into the emitter is low (as in the emitter follower). Although the output impedance of a common-emitter amplifier will be dominated by the collector load resistor, the output impedance of an emitter follower will not be dominated by the emitter load resistor, but rather by the impedance looking into the emitter.

#### 2.2.8 Unity-gain phase splitter

Sometimes it is useful to generate a signal and its inverse, i.e., two signals  $180^\circ$  out of phase. That's easy to do – just use an emitter-degenerated amplifier with a gain of  $-1$  (Figure 2.36). The quiescent collector voltage is set to  $0.75V_{\text{CC}}$ , rather than the usual  $0.5V_{\text{CC}}$ , in order to achieve the same result – maximum symmetrical output swing without clipping at either output. The collector can swing from  $0.5V_{\text{CC}}$  to  $V_{\text{CC}}$ , whereas the emitter can swing from ground to  $0.5V_{\text{CC}}$ .

Note that the phase-splitter outputs must be loaded with equal (or very high) impedances at the two outputs to maintain gain symmetry.

<sup>21</sup> The impedance looking into the base itself will usually be much larger because of the way the base resistors are chosen, and it can generally be ignored.

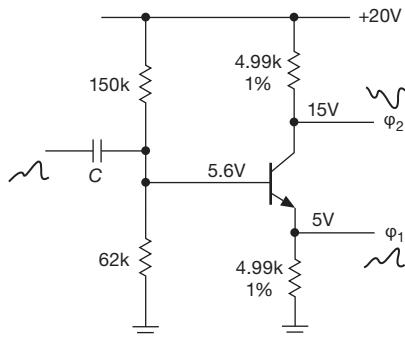


Figure 2.36. Unity-gain phase splitter.

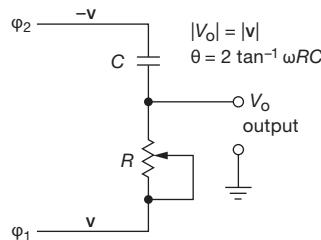


Figure 2.37. Constant-amplitude phase shifter.

### A. Phase shifter

A nice use of the phase splitter is shown in Figure 2.37. This circuit gives (for a sinewave input) an output sinewave of adjustable phase (from zero to  $180^\circ$ ) and with constant amplitude. It can be best understood with a phasor diagram of voltages (§1.7.12); with the input signal represented by a unit vector along the real axis, the signals look as shown in Figure 2.38.

Signal vectors  $v_R$  and  $v_C$  must be at right angles, and they must add to form a vector of constant length along the real axis. There is a theorem from geometry that says that the locus of such points is a circle. So the resultant vector (the output voltage) always has unit length, i.e., the same amplitude as the input, and its phase can vary from nearly zero to nearly  $180^\circ$  relative to the input wave as  $R$  is varied from nearly zero to a value much larger than  $X_C$  at the operating frequency. However, note that the phase shift depends on the frequency of the input signal for a given setting of the potentiometer  $R$ . It is worth noting that a simple  $RC$  highpass (or lowpass) network could also be used as an adjustable phase shifter. However, its output amplitude would vary over an enormous range as the phase shift was adjusted.

An additional concern here is the ability of the phase-splitter circuit to drive the  $RC$  phase shifter as a load. Ideally, the load should present an impedance that is large

compared with the collector and emitter resistors. As a result, this circuit is of limited utility where a wide range of phase shifts is required. You will see improved phase-splitter techniques in Chapter 4, where we use op-amps as impedance buffers, and in Chapter 7, where a cascade of several phase-shifter sections generates a set of “quadrature” signals that extends the phase-shifting range to a full  $0^\circ$  to  $360^\circ$ .

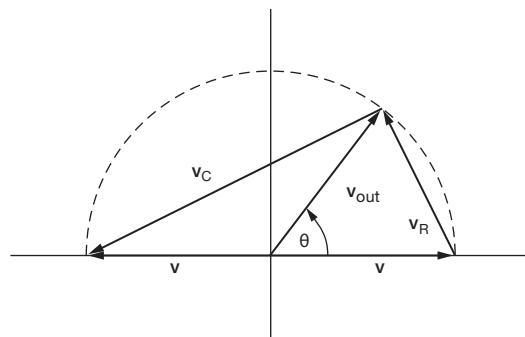
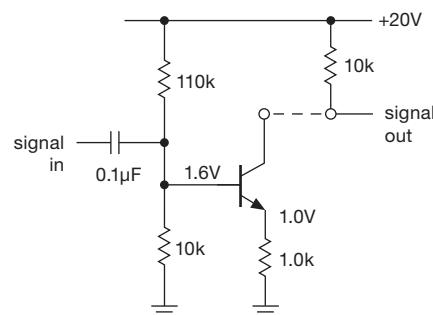
Figure 2.38. Phasor diagram for phase shifter, for which  $\theta = 2 \arctan(\omega RC)$ .

Figure 2.39. The common-emitter amplifier is a transconductance stage driving a (resistive) load.

### 2.2.9 Transconductance

In the preceding section we figured out the operation of the emitter-degenerated amplifier by (a) imagining an applied base voltage swing and seeing that the emitter voltage had the same swing, then (b) calculating the emitter current swing; then, ignoring the small base current contribution, we got the collector current swing and thus (c) the collector voltage swing. The voltage gain was then simply the ratio of collector (output) voltage swing to base (input) voltage swing.

There's another way to think about this kind of amplifier. Imagine breaking it apart, as in Figure 2.39. The first

part is a voltage-controlled current source, with quiescent current of 1.0 mA and gain of  $-1 \text{ mA/V}$ . Gain means the ratio of output to input; in this case the gain has units of current/voltage, or  $1/\text{resistance}$ . The inverse of resistance is called *conductance*.<sup>22</sup> An amplifier whose gain has units of conductance is called a *transconductance* amplifier; the ratio of changes  $\Delta I_{\text{out}}/\Delta V_{\text{in}}$  (usually written with lowercase:  $i_{\text{out}}/v_{\text{in}}$ ) is called the transconductance,  $g_m$ :

$$g_m = \frac{\Delta I_{\text{out}}}{\Delta V_{\text{in}}} = \frac{i_{\text{out}}}{v_{\text{in}}}. \quad (2.7)$$

Think of the first part of the circuit as a transconductance amplifier, i.e., a voltage-to-current amplifier with transconductance  $g_m$  (gain) of  $-1 \text{ mA/V}$  ( $1000 \mu\text{S}$ , or  $1 \text{ mS}$ , which is just  $1/R_E$ ). The second part of the circuit is the load resistor, an “amplifier” that converts current to voltage. This resistor could be called a *transresistance* converter, and its gain ( $r_m$ ) has units of voltage/current, or resistance. In this case its quiescent voltage is  $V_{\text{CC}}$ , and its gain (transresistance) is  $10 \text{ V/mA}$  ( $10\text{k}\Omega$ ), which is just  $R_C$ . Connecting the two parts together gives you a voltage amplifier. You get the overall gain by multiplying the two gains. In this case the voltage gain  $G_V = g_m R_C = -R_C/R_E$ , or  $-10$ , a unitless number equal to the ratio (output voltage change)/(input voltage change).

This is a useful way to think about an amplifier, because you can analyze performance of the sections independently. For example, you can analyze the transconductance part of the amplifier by evaluating  $g_m$  for different circuit configurations or even different devices, such as field-effect transistors FETs. Then you can analyze the transresistance (or load) part by considering gain versus voltage swing tradeoffs. If you are interested in the overall voltage gain, it is given by  $G_V = g_m r_m$ , where  $r_m$  is the transresistance of the load. Ultimately the substitution of an active load (current source), with its extremely high transresistance, can yield single-stage voltage gains of 10,000 or more. The *cascode* configuration, which we will discuss later, is another example easily understood with this approach.

In Chapter 4, which deals with operational amplifiers, you will see further examples of amplifiers with voltages or currents as inputs or outputs: voltage amplifiers (voltage to voltage), current amplifiers (current to current), and transresistance amplifiers (current to voltage).

<sup>22</sup> The inverse of reactance is *susceptance* (and the inverse of impedance is *admittance*), and has a special unit, the *siemens* (“S,” not to be confused with lowercase “s,” which means seconds), which used to be called the *mho* (ohm spelled backward, symbol “Ω”).

## A. Turning up the gain: limitations of the simple model

The voltage gain of the emitter-degenerated amplifier is  $-R_C/R_E$ , according to our model. What happens as  $R_E$  is reduced toward zero? The equation predicts that the gain will rise without limit. But if we made actual measurements of the preceding circuit, keeping the quiescent current constant at 1 mA, we would find that the gain would level off at about 400 when  $R_E$  is zero, i.e., with the emitter grounded. We would also find that the amplifier would become significantly nonlinear (the output would not be a faithful replica of the input), the input impedance would become small and nonlinear, and the biasing would become critical and unstable with temperature. Clearly our transistor model is incomplete and needs to be modified to handle this circuit situation, as well as others we will talk about presently. Our fixed-up model, which we will call the transconductance model, will be accurate enough for the remainder of the book.

## B. Recap: the “four topologies”

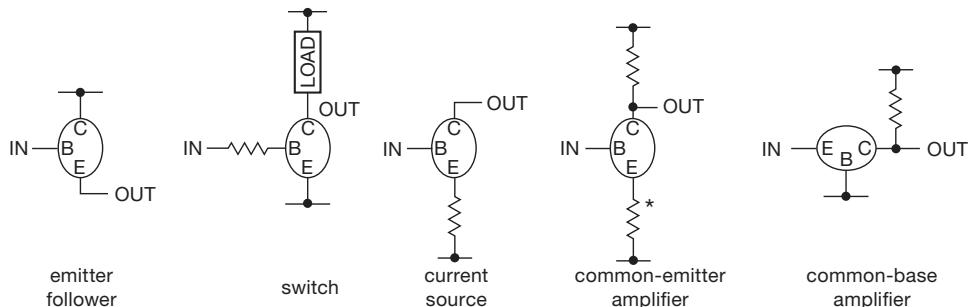
Before jumping into the complexity just ahead, let’s remind ourselves of the four transistor circuits we’ve seen, namely the switch, emitter follower, current source, and common-emitter amplifier. We’ve drawn these very schematically in Figure 2.40, omitting details like biasing, and even the polarity of transistor (i.e., *npn* or *pnp*). For completeness we’ve included also a fifth circuit, the *common-base amplifier*, which we’ll meet soon enough (§2.4.5B).

## 2.3 Ebers–Moll model applied to basic transistor circuits

We’ve enjoyed seeing some nice feats that can be accomplished with the simplest BJT model – switch, follower, current source, amplifier – but we’ve run up against some serious limitations (hey, would you believe, *infinite gain*?!). Now it’s time to go a level deeper, to address these limitations. The material that follows will suffice for our purposes. And – good news – for many BJT applications the simple model you’ve already seen is completely adequate.

### 2.3.1 Improved transistor model: transconductance amplifier

The important change is in rule 4 (§2.1.1), where we said earlier that  $I_C = \beta I_B$ . We thought of the transistor as a current amplifier whose input circuit behaved like a diode. That’s roughly correct, and for some applications it’s good enough. But to understand differential amplifiers,



**Figure 2.40.** Five basic transistor circuits. Fixed voltages (power supplies or ground) are indicated by connections to horizontal line segments. For the switch, the load may be a resistor, to produce a full-swing voltage output; for the common-emitter amplifier, the emitter resistor may be bypassed or omitted altogether.

logarithmic converters, temperature compensation, and other important applications, you must think of the transistor as a *transconductance* device – collector current is determined by base-to-emitter voltage.

Here's the modified rule 4.

**4. Transconductance amplifier** When rules 1–3 (§2.1.1) are obeyed,  $I_C$  is related to  $V_{BE}$  by<sup>23</sup>

$$I_C = I_S(T) \left( e^{V_{BE}/V_T} - 1 \right), \quad (2.8)$$

or, equivalently,

$$V_{BE} = \frac{kT}{q} \log_e \left( \frac{I_C}{I_S(T)} + 1 \right), \quad (2.9)$$

where

$$V_T = kT/q = 25.3 \text{ mV} \quad (2.10)$$

at room temperature (68°F, 20°C),  $q$  is the electron charge ( $1.60 \times 10^{-19}$  coulombs),  $k$  is Boltzmann's constant ( $1.38 \times 10^{-23}$  joules/K, sometimes written  $k_B$ ),  $T$  is the absolute temperature in degrees Kelvin ( $K = ^\circ C + 273.16$ ), and  $I_S(T)$  is the *saturation current* of the particular transistor (which depends strongly on temperature,  $T$ , as we'll see shortly). Then the base current, which also depends on  $V_{BE}$ , can be approximated by

$$I_B = I_C/\beta,$$

where the “constant”  $\beta$  is typically in the range 20 to 1000, but depends on transistor type,  $I_C$ ,  $V_{CE}$ , and temperature.  $I_S(T)$  approximates the reverse leakage current (roughly  $10^{-15}$  A for a small-signal transistor like the 2N3904). In the active region  $I_C \gg I_S$ , and therefore

<sup>23</sup> We indicate the important temperature dependence of  $I_S$  by explicitly showing it in functional form – “ $I_S(T)$ ”.

the  $-1$  term can be neglected in comparison with the exponential:

$$I_C \approx I_S(T) e^{V_{BE}/V_T}. \quad (2.11)$$

The equation for  $I_C$  is known as the Ebers–Moll equation.<sup>24</sup> It also describes approximately the current versus voltage for a diode, if  $V_T$  is multiplied by a correction factor  $m$  between 1 and 2. For transistors it is important to realize that the collector current is accurately determined by the base–emitter voltage, rather than by the base current (the base current is then roughly determined by  $\beta$ ), and that this exponential law is accurate over an enormous range of currents, typically from nanoamps to millamps. Figure 2.41 makes the point graphically.<sup>25</sup> If you measure the base current at various collector currents, you will get a graph of  $\beta$  versus  $I_C$  like that in Figure 2.42. Transistor beta versus collector current is discussed further in Chapter 2x.

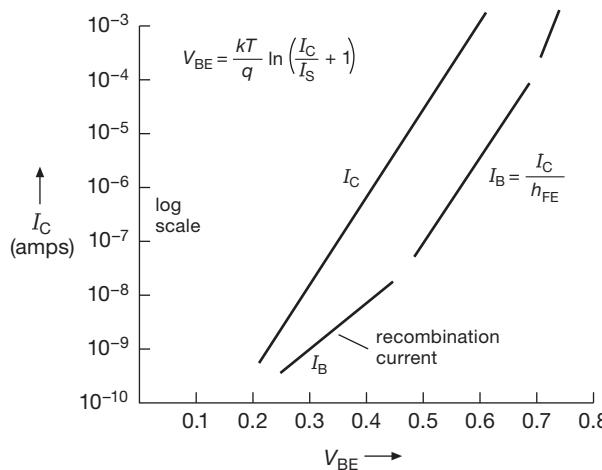
Although the Ebers–Moll equation tells us that the base–emitter voltage “programs” the collector current, this property is not easy to use in practice (biasing a transistor by applying a base voltage) because of the large temperature coefficient of base–emitter voltage. You will see later how the Ebers–Moll equation provides insight and solutions to this problem.

### 2.3.2 Consequences of the Ebers–Moll model: rules of thumb for transistor design

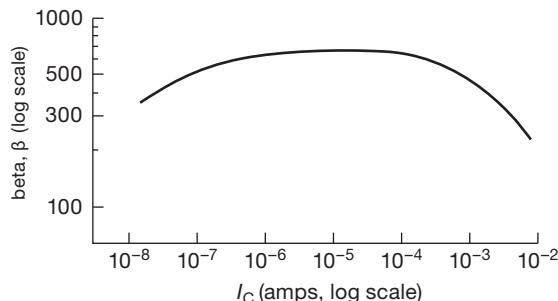
From the Ebers–Moll equation (2.8) we get these simple (but handy) “ratio rules” for collector current:  $I_{C2}/I_{C1} = \exp(\Delta V_{BE}/V_T)$  and  $\Delta V_{BE} = V_T \log_e(I_{C2}/I_{C1})$ . We

<sup>24</sup> J. J. Ebers & J. L. Moll, “Large-signal behavior of junction transistors,” *Proc. IRE* **42**, 1761 (1954).

<sup>25</sup> This is sometimes called a Gummel plot.



**Figure 2.41.** Transistor base and collector currents as functions of base-to-emitter voltage  $V_{BE}$ .



**Figure 2.42.** Typical transistor current gain ( $\beta$ ) versus collector current.

also get the following important quantities we will be using often in circuit design.

#### A. The steepness of the diode curve.

How much do we need to increase  $V_{BE}$  to increase  $I_C$  by a factor of 10? From the Ebers–Moll equation, that's just  $V_T \log_e 10$ , or 58.2 mV at room temperature. We like to remember this as *base–emitter voltage increases approximately 60 mV per decade of collector current*. (Two other formulations: collector current doubles for each 18 mV increase in base–emitter voltage; collector current increases 4% per millivolt increase in base–emitter voltage.) Equivalently,  $I_C = I_{C0} e^{\Delta V / 25}$ , where  $\Delta V$  is in millivolts.<sup>26</sup>

#### B. The small-signal impedance looking into the emitter, $r_e$ , for the base held at a fixed voltage.

Taking the derivative of  $V_{BE}$  with respect to  $I_C$ , you get

$$r_e = V_T / I_C = 25 / I_C \text{ ohms,} \quad (2.12)$$

where  $I_C$  is in millamps.<sup>27</sup> The numerical value  $25/I_C$  is for room temperature. This *intrinsic* emitter resistance,  $r_e$ , acts as if it is in series with the emitter in all transistor circuits. It limits the gain of a grounded-emitter amplifier, causes an emitter follower to have a voltage gain of slightly less than unity, and prevents the output impedance of an emitter follower from reaching zero. Note that the transconductance<sup>28</sup> of a grounded emitter amplifier is just

$$g_m = I_C / V_T = 1 / r_e \quad (= 40 I_C \text{ at room temp}). \quad (2.13)$$

#### C. The temperature dependence of $V_{BE}$ .

A glance at the Ebers–Moll equation suggests that  $V_{BE}$  (at constant  $I_C$ ) has a positive temperature coefficient because of the multiplying factor of  $T$  in  $V_T$ . However, the strong temperature dependence of  $I_S(T)$  more than compensates for that term, such that  $V_{BE}$  (at constant  $I_C$ ) *decreases* about 2.1 mV/ $^\circ\text{C}$ . It is roughly proportional to  $1/T_{abs}$ , where  $T_{abs}$  is the absolute temperature. Sometimes it's useful to cast this instead in terms of the temperature dependence of  $I_C$  (at constant  $V_{BE}$ ):  $I_C$  *increases* about 9%/ $^\circ\text{C}$ ; it doubles for an 8 $^\circ\text{C}$  rise.

There is one additional quantity we will need on occasion, although it is not derivable from the Ebers–Moll equation. It is known as the Early effect,<sup>29</sup> and it sets important limits on current-source and amplifier performance.

#### D. Early effect.

$V_{BE}$  (at constant  $I_C$ ) varies slightly with changing  $V_{CE}$ . This effect is caused by the variation of effective base width as  $V_{CE}$  changes, and it is given, approximately, by

$$\Delta V_{BE} = -\eta \Delta V_{CE}, \quad (2.14)$$

where  $\eta \approx 10^{-4}$ – $10^{-5}$ . (As an example, the *npn* 2N5088 has  $\eta = 1.3 \times 10^{-4}$ , thus a 1.3 mV change of  $V_{BE}$  to maintain constant collector current when  $V_{CE}$  changes by 10 V.)

opportunity to make a “silicon thermometer.” We’ll see more of this in Chapter 2x, and again in Chapter 9.

<sup>27</sup> We like to remember the fact that  $r_e = 25 \Omega$  at a collector current of 1 mA. Then we just scale inversely for other currents; thus  $r_e = 2.5 \Omega$  at  $I_C = 10 \text{ mA}$ , etc.

<sup>28</sup> At the next level of sophistication we’ll see that, since the quantity  $r_e$  is proportional to absolute temperature, a grounded emitter amplifier whose collector current is PTAT has transconductance (and gain) independent of temperature. More in Chapter 2x.

<sup>29</sup> J. M. Early, “Effects of space-charge layer widening in junction transistors,” *Proc. IRE* **40**, 1401 (1952). James Early died in 2004.

<sup>26</sup> The “25” in this and the following discussion is more precisely 25.3 mV, the value of  $k_B T/q$  at room temperature. It’s proportional to absolute temperature – engineers like to say “PTAT,” pronounced *pee-tat*. This has interesting (and useful) consequences, for example the

This is often described instead as a linear increase of collector current with increasing collector voltage when  $V_{BE}$  is held constant; you see it expressed as

$$I_C = I_{C0} \left( 1 + \frac{V_{CE}}{V_A} \right), \quad (2.15)$$

where  $V_A$  (typically 50–500 V) is known as the Early voltage.<sup>30</sup> This is shown graphically in Figure 2.59 in §2.3.7A. A low Early voltage indicates a low collector output resistance; *pnp* transistors tend to have low  $V_A$ , see measured values in Table 8.1. We treat the Early effect in more detail in Chapter 2x.<sup>31</sup>

These are the essential quantities we need. With them we will be able to handle most problems of transistor circuit design, and we will have little need to refer to the Ebers–Moll equation itself.<sup>32</sup>

### 2.3.3 The emitter follower revisited

Before looking again at the common-emitter amplifier with the benefit of our new transistor model, let's take a quick look at the humble emitter follower. The Ebers–Moll model predicts that an emitter follower should have nonzero output impedance, even when driven by a voltage source, because of finite  $r_e$  (item B in the above list). The same effect also produces a voltage gain slightly less than unity, because  $r_e$  forms a voltage divider with the load resistor.

These effects are easy to calculate. With fixed base voltage, the impedance looking back into the emitter is just  $R_{out} = dV_{BE}/dI_E$ ; but  $I_E \approx I_C$ , so  $R_{out} \approx r_e$ , the intrinsic emitter resistance [recall  $r_e = 25/I_C(\text{mA})$ ]. For example, in Figure 2.43A, the load sees a driving impedance of  $r_e = 25 \Omega$ , because  $I_C = 1 \text{ mA}$ . (This is paralleled by the emitter resistor  $R_E$ , if used; but in practice  $R_E$  will always be much larger than  $r_e$ .) Figure 2.43B shows a more typical situation, with finite source resistance  $R_s$  (for simplicity we've omitted the obligatory biasing components –

<sup>30</sup> The connection between Early voltage and  $\eta$  is  $\eta = V_T/(V_A + V_{CE}) \approx V_T/V_A$ ; see Chapter 2x.

<sup>31</sup> Previewing some of the results there, the Early effect (a) determines a transistor's collector output resistance  $r_o = V_A/I_C$ ; (b) sets a limit on single-stage voltage gain; and (c) limits the output resistance of a current source. Other things being equal, *pnp* transistors tend to have low Early voltages, as do transistors with high beta; high-voltage transistors usually have high Early voltages, along with low beta. These trends can be seen in the measured Early voltages listed in Table 8.1.

<sup>32</sup> The computer circuit-analysis program SPICE includes accurate transistor simulation with the Ebers–Moll formulas and Gummel–Poon charge models. It's a lot of fun to "wire up" circuits on your computer screen and set them running with SPICE. For more detail see the application of SPICE to BJT amplifier distortion in Chapter 2x.

base divider and blocking capacitor – which are shown in Figure 2.43C). In this case the emitter follower's output impedance is just  $r_e$  in series with  $R_s/(\beta + 1)$  (again paralleled by an unimportant  $R_E$ , if present). For example, if  $R_s = 1\text{k}$  and  $I_C = 1 \text{ mA}$ ,  $R_{out} = 35 \Omega$  (assuming  $\beta = 100$ ). It is easy to show that the intrinsic emitter  $r_e$  also figures into an emitter follower's *input* impedance, just as if it were in series with the load (actually, parallel combination of load resistor and emitter resistor). In other words, for the emitter follower circuit the effect of the Ebers–Moll model is simply to add a series emitter resistance  $r_e$  to our earlier results.<sup>33</sup>

The voltage gain of an emitter follower is slightly less than unity, owing to the voltage divider produced by  $r_e$  and the load. It is simple to calculate, because the output is at the junction of  $r_e$  and  $R_{load}$ :  $G_V = v_{out}/v_{in} = R_L/(r_e + R_L)$ . Thus, for example, a follower running at 1 mA quiescent current, with 1k load, has a voltage gain of 0.976. Engineers sometimes like to write the gain in terms of the transconductance, to put it in a form that holds for FETs also (see §3.2.3A); in that case (using  $g_m = 1/r_e$ ) you get  $G_V = R_L g_m / (1 + R_L g_m)$ .

### 2.3.4 The common-emitter amplifier revisited

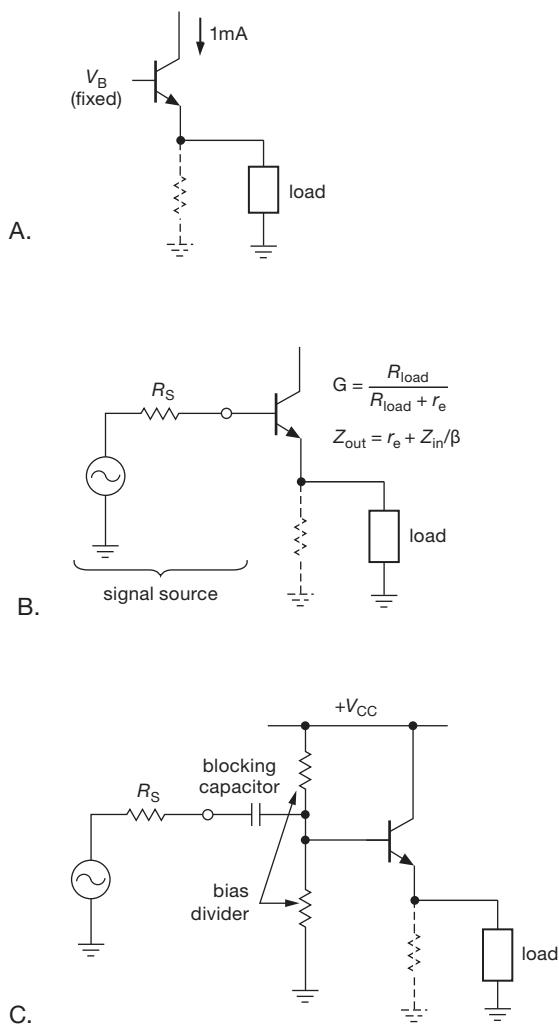
Previously we got wrong answers for the voltage gain of the common-emitter amplifier with emitter resistor (sometimes called emitter degeneration) when we set the emitter resistor equal to zero; recall that our wrong answer was  $G_V = -R_C/R_E = \infty$ !

The problem is that the transistor has  $25/I_C(\text{mA})$  ohms of built-in (intrinsic) emitter resistance  $r_e$  that must be added to the actual external emitter resistor. This resistance is significant only when small emitter resistors (or none at all) are used.<sup>34</sup> So, for instance, the amplifier we considered previously will have a voltage gain of  $-10k/r_e$ , or  $-400$ , when the external emitter resistor is zero. The input impedance is not zero, as we would have predicted earlier ( $\beta R_E$ ); it is approximately  $\beta r_e$ , or in this case (1 mA quiescent current) about 2.5k.<sup>35</sup>

<sup>33</sup> There's more, if you look deeper: at high frequencies (above  $f_T/\beta$ ) the effective current gain drops inversely with frequency; so you get a linearly rising output impedance from an emitter follower that is driven with low  $R_s$ . That is, it looks like an inductance, and a capacitive load can cause ringing or even oscillation; these effects are treated in Chapter 2x.

<sup>34</sup> Or, equivalently, when the emitter resistor is bypassed with a capacitor whose impedance at signal frequencies is comparable with, or less than,  $r_e$ .

<sup>35</sup> These estimates of gain and input impedance are reasonably good, as long as we stay away from operation at very high frequencies or



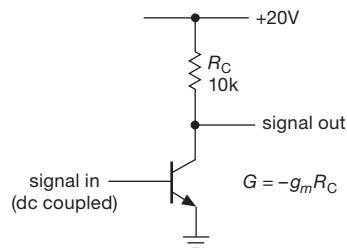
**Figure 2.43.** Output impedance of emitter followers (see text).

The terms “grounded emitter” and “common emitter” are sometimes used interchangeably, and they can be confusing. We will use the phrase “grounded-emitter amplifier” to mean a common-emitter amplifier with  $R_E = 0$  (or equivalent bypassing). A common-emitter amplifier stage may have an emitter resistor; what matters is that the emitter circuit is common to the input circuit and the output circuit.

from circuits in which the collector load resistor is replaced with a current source “active load” ( $R_C \rightarrow \infty$ ). The ultimate voltage gain of a grounded-emitter amplifier, in the latter situation, is limited by the Early effect; this is discussed in more detail Chapter 2x.

### A. Shortcomings of the single-stage grounded emitter amplifier

The extra voltage gain you get by using  $R_E = 0$  comes at the expense of other properties of the amplifier. In fact, the grounded-emitter amplifier, in spite of its popularity in textbooks, should be avoided except in circuits with overall negative feedback. In order to see why, consider Figure 2.44.



**Figure 2.44.** Common-emitter amplifier without emitter degeneration.

**1. Nonlinearity.** The voltage gain is  $G = -g_m R_C = -R_C/r_e = -R_C I_C(\text{mA})/25$ , so for a quiescent current of 1 mA, the gain is -400. But  $I_C$  varies as the output signal varies. For this example, the gain will vary from -800 ( $V_{\text{out}} = 0$ ,  $I_C = 2 \text{ mA}$ ) down to zero ( $V_{\text{out}} = V_{\text{CC}}$ ,  $I_C = 0$ ). For a triangle-wave input, the output will look as in Figure 2.45. The amplifier has lots of distortion, or poor linearity. The grounded-emitter amplifier without feedback is useful only for small-signal swings about the quiescent point. By contrast, the emitter-degenerated amplifier has a gain almost entirely independent of collector current, as long as  $R_E \gg r_e$ , and can be used for undistorted amplification even with large-signal swings.

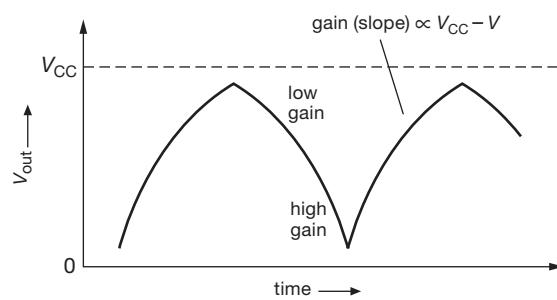
It's easy to estimate the distortion, both with and without an external emitter resistor. With a *grounded* emitter, the incremental (small-signal) gain is  $G_V = -R_C/r_e = -I_C R_C/V_T = -V_{\text{drop}}/V_T$ , where  $V_{\text{drop}}$  is the instantaneous voltage drop across the collector resistor. Because the gain is proportional to the drop across the collector resistor, the nonlinearity (fractional change of gain with swing) equals the ratio of instantaneous swing to average quiescent drop across the collector resistor:  $\Delta G/G \approx \Delta V_{\text{out}}/V_{\text{drop}}$ , where  $V_{\text{drop}}$  is the average, or quiescent, voltage drop across the collector resistor  $R_C$ . Because this represents the extreme variation of gain (i.e., at the peaks of the swing), the overall waveform “distortion” (usually stated as the amplitude of the residual waveform after subtraction of the strictly linear component) will be smaller by roughly a factor of 3. Note that the distortion depends on only the ratio of swing to quiescent drop, and not directly on the operating current, etc.

As an example, in a grounded emitter amplifier powered from +10 V, biased to half the supply (i.e.,  $V_{\text{drop}} = 5 \text{ V}$ ), we measured a distortion of 0.7% at 0.1 V output sinewave amplitude and 6.6% at 1 V amplitude; these values are in good agreement with the predicted values. Compare this with the situation with an added external emitter resistor  $R_E$ , in which the voltage gain becomes  $G_V = -R_C/(r_e + R_E) = -I_C R_C / (V_T + I_C R_E)$ . Only the first term in the denominator contributes distortion, so the distortion is reduced by the ratio of  $r_e$  to the total effective emitter resistance: the nonlinearity becomes  $\Delta G/G \approx (\Delta V_{\text{out}}/V_{\text{drop}})[r_e/(r_e + R_E)] = (\Delta V_{\text{out}}/V_{\text{drop}})[V_T/(V_T + I_C R_E)]$ ; the second term is the factor by which the distortion is reduced. When we added an emitter resistor, chosen to drop 0.25 V at the quiescent current – which by this estimate should reduce the nonlinearity by a factor of 10 – the measured distortion of the previous amplifier dropped to 0.08% and 0.74% for 0.1 V and 1 V output amplitudes, respectively. Once again, these measurement agree well with our prediction.

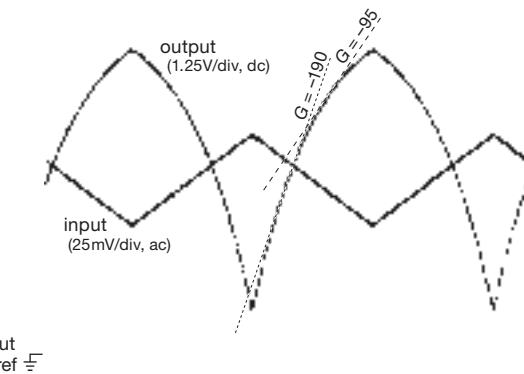
**Exercise 2.12.** Calculate the predicted distortion for these two amplifiers at the two output levels that were measured.

As we remarked, the nonlinearity of a common-emitter amplifier, when driven by a triangle wave, takes the form of the asymmetric “barn roof” distortion sketched in Figure 2.45.<sup>36</sup> For comparison we took a real-life ‘scope (oscilloscope) trace of a grounded emitter amplifier (Figure 2.46); we used a 2N3904 with a 5k collector resistor to a +10 V supply, biased (carefully!) to half the supply. With a ruler we estimated the incremental gain at output voltages of +5 V (halfway to  $V_+$ ) and at +7.5 V, as shown, where the collector current is 1 mA and 0.5 mA, respectively. The gain values are in pretty good agreement with the predictions ( $G = R_C/r_e = I_C(\text{mA})R_C/25\Omega$ ) of  $G = -200$  and  $G = -100$ , respectively. By comparison, Figure 2.47 shows what happened when we added a 225  $\Omega$  emitter resistor: the gain is reduced by a factor of 10 at the quiescent point ( $G = R_C/(R_E + r_e) \approx R_C/250\Omega$ ), but with much improved linearity (because changes in  $r_e$  contribute little to the overall resistance in the denominator, which is now dominated by the fixed 225  $\Omega$  external emitter resistor).

For sinusoidal input, the output contains all harmonics of the fundamental wave. Later in the chapter we’ll see how to make differential amplifiers with a pair of transistors; for these the residual distortion is symmetric, and contains only the odd harmonics. And in Chapter 2x we’ll see some very clever methods for cancelling distortion in differential



**Figure 2.45.** Nonlinear output waveform from grounded-emitter amplifier.



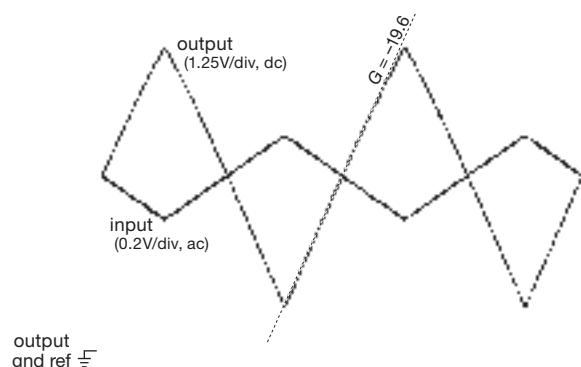
**Figure 2.46.** *Real life!* The grounded-emitter amplifier of Figure 2.44, with  $R_C = 5\text{k}$ ,  $V_+ = +10 \text{ V}$ , and a 1 kHz triangle wave input. Top and bottom of the screen are +10 V and ground for the dc-coupled output trace (note sensitive scale for the ac-coupled input signal). Gain estimates (tangent lines) are at  $V_{\text{out}}$  values of  $0.5V_+$  and  $0.75V_+$ . Horizontal: 0.2 ms/div.

amplifiers, along with the use of SPICE simulation software for rapid analysis and circuit iteration. Finally, to set things in perspective, we should add that any amplifier’s residual distortion can be reduced dramatically by use of *negative feedback*. We’ll introduce feedback later in this chapter (§2.5), after you’ve gained familiarity with common transistor circuits. Feedback will finally take center stage when we get to *operational amplifiers* in Chapter 4.

**2. Input impedance.** The input impedance is roughly  $Z_{\text{in}} = \beta r_e = 25\beta/I_C(\text{mA})$  ohms. Once again,  $I_C$  varies over the signal swing, giving a varying input impedance. Unless the signal source driving the base has low impedance, you will wind up with nonlinearity because of the nonlinear (variable) voltage divider formed from the signal source and the amplifier’s input impedance. By contrast, the input impedance of an emitter-degenerated amplifier is nearly constant, and high.

**3. Biasing.** The grounded emitter amplifier is difficult to

<sup>36</sup> Because the gain (i.e., the slope of  $V_{\text{out}}$  versus  $V_{\text{in}}$ ) is proportional to the distance from the  $V_{\text{CC}}$  line, the shape of the curve is in fact an exponential.



**Figure 2.47.** Adding a  $225\Omega$  emitter resistor improves the linearity dramatically at the expense of gain (which drops by a factor of 10 at the quiescent point). Horizontal: 0.2 ms/div.

bias. It might be tempting just to apply a voltage (from a voltage divider) that gives the right quiescent current according to the Ebers–Moll equation. That won't work, because of the temperature dependence of  $V_{BE}$  (at fixed  $I_C$ ), which varies about  $2.1 \text{ mV}/^\circ\text{C}$  [it actually decreases with increasing  $T$  because of the variation of  $I_S(T)$  with temperature; as a result,  $V_{BE}$  is roughly proportional to  $1/T$ , the absolute temperature]. This means that the collector current (for fixed  $V_{BE}$ ) will increase by a factor of 10 for a  $30^\circ\text{C}$  rise in temperature (which corresponds to a  $60 \text{ mV}$  change in  $V_{BE}$ ), or about  $9\%/\text{C}$ . Such unstable biasing is useless, because even rather small changes in temperature will cause the amplifier to saturate. For example, a grounded emitter stage biased with the collector at half the supply voltage will go into saturation if the temperature rises by  $8^\circ\text{C}$ .

**Exercise 2.13.** Verify that an  $8^\circ\text{C}$  rise in ambient temperature will cause a base-voltage-biased grounded emitter stage to saturate, assuming that it was initially biased for  $V_C = 0.5V_{CC}$ .

Some solutions to the biasing problem are discussed in the following sections. By contrast, the emitter-degenerated amplifier achieves stable biasing by applying a voltage to the base, most of which appears across the emitter resistor, thus determining the quiescent current.

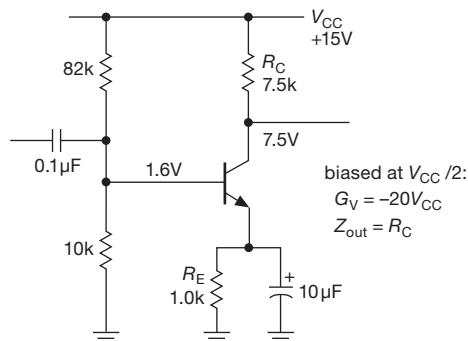
### B. Emitter resistor as feedback

Adding an external series resistor to the intrinsic emitter resistance  $r_e$  (emitter degeneration) improves many properties of the common-emitter amplifier, but at the expense of gain. You will see the same thing happening in Chapters 4 and 5, when we discuss *negative feedback*, an important technique for improving amplifier characteristics by feeding back some of the output signal to reduce the effective input signal. The similarity here is no coincidence –

the emitter-degenerated amplifier itself uses a form of negative feedback. Think of the transistor as a transconductance device, determining collector current (and therefore output voltage) according to the voltage applied between the base and emitter; but the input to the amplifier is the voltage from base to ground. So the voltage from base to emitter is the input voltage, minus a sample of the output (namely  $I_E R_E$ ). That's negative feedback, and that's why emitter degeneration improves most properties of the amplifier (here improved linearity and stability and increased input impedance.<sup>37</sup>) Later in the chapter, in §2.5, we'll make these statements quantitative when we first look at feedback. And there are great things to look forward to, with the full flowering of feedback in Chapters 4 and 5!

### 2.3.5 Biasing the common-emitter amplifier

If you must have the highest possible gain (or if the amplifier stage is inside a feedback loop), it is possible to arrange successful biasing of a common-emitter amplifier. There are three solutions that can be applied, separately or in combination: bypassed emitter resistor, matched biasing transistor, and dc feedback.



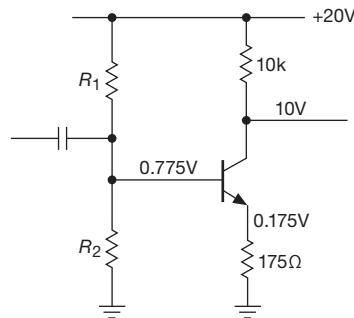
**Figure 2.48.** A bypassed emitter resistor can be used to improve the bias stability of a grounded-emitter amplifier.

#### A. Bypassed emitter resistor

You can use a bypassed emitter resistor, biasing as for the degenerated amplifier, as shown in Figure 2.48. In this case  $R_E$  has been chosen about  $0.1R_C$ , for ease of biasing; if  $R_E$  is too small, the emitter voltage will be much smaller than the base–emitter drop, leading to temperature instability of the quiescent point as  $V_{BE}$  varies with temperature. The emitter bypass capacitor is chosen to make its impedance small compared with  $r_e$  (not  $R_E$  – why?) at the

<sup>37</sup> And, as we'll learn, the output impedance would be reduced – a desirable feature in a voltage amplifier – if the feedback were taken directly from the collector.

lowest frequency of interest. In this case its impedance is  $25\Omega$  at 650 Hz. At signal frequencies the input coupling capacitor sees an impedance of  $10k$  in parallel with the base impedance, in this case  $\beta \times 25\Omega$ , or roughly  $2.5k$ . At dc, the impedance looking into the base is much larger ( $\beta$  times the emitter resistor, or about  $100k$ ).



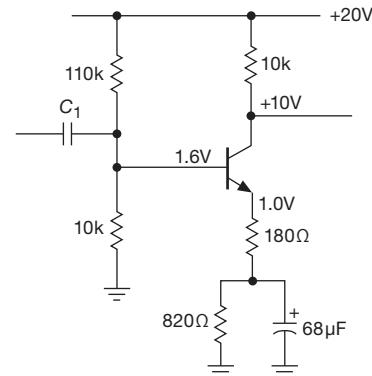
**Figure 2.49.** Gain-of-50 stage presents bias stability problem.

A variation on this circuit consists of using two emitter resistors in series, one of them bypassed. For instance, suppose you want an amplifier with a voltage gain of 50, quiescent current of 1 mA, and  $V_{CC}$  of +20 volts, for signals from 20 Hz to 20 kHz. If you try to use the emitter-degenerated circuit, you will have the circuit shown in Figure 2.49. The collector resistor is chosen to put the quiescent collector voltage at  $0.5V_{CC}$ . Then the emitter resistor is chosen for the required gain, including the effects of the  $r_e$  of  $25/I_C$ (mA). The problem is that the emitter voltage of only 0.175 V will vary significantly as the  $\sim 0.6$  V of base-emitter drop varies with temperature ( $-2.1$  mV/ $^{\circ}\text{C}$ , approximately), since the base is held at constant voltage by  $R_1$  and  $R_2$ ; for instance, you can verify that an increase of  $20^{\circ}\text{C}$  will cause the collector current to increase by nearly 25%.

**Exercise 2.14.** Show that this statement is correct.

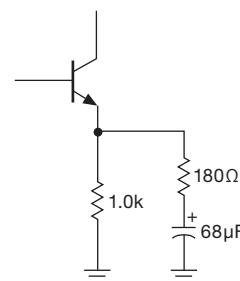
The solution here is to add some bypassed emitter resistance for stable biasing, with no change in gain at signal frequencies (Figure 2.50). As before, the collector resistor is chosen to put the collector at 10 volts ( $0.5V_{CC}$ ). Then the unbypassed emitter resistor is chosen to give a gain of 50, including the intrinsic emitter resistance  $r_e = 25/I_C$ (mA). Enough bypassed emitter resistance is added to make stable biasing possible (one-tenth of the collector resistance is a good guideline). The base voltage is chosen to give 1 mA of emitter current, with impedance about one-tenth the dc impedance looking into the base (in this case about  $100k$ ). The emitter bypass capacitor is chosen to have low impedance compared with  $180 + 25\Omega$  at the lowest signal frequencies. Finally, the input coupling ca-

pacitor is chosen to have low impedance compared with the *signal-frequency* input impedance of the amplifier, which is equal to the voltage-divider impedance in parallel with  $\beta \times (180 + 25)\Omega$  (the  $820\Omega$  is bypassed and looks like a short at signal frequencies).



**Figure 2.50.** A common-emitter amplifier combining bias stability, linearity, and large voltage gain.

An alternative circuit splits the signal and dc paths (Figure 2.51). This lets you vary the gain (by changing the  $180\Omega$  resistor) without bias change.



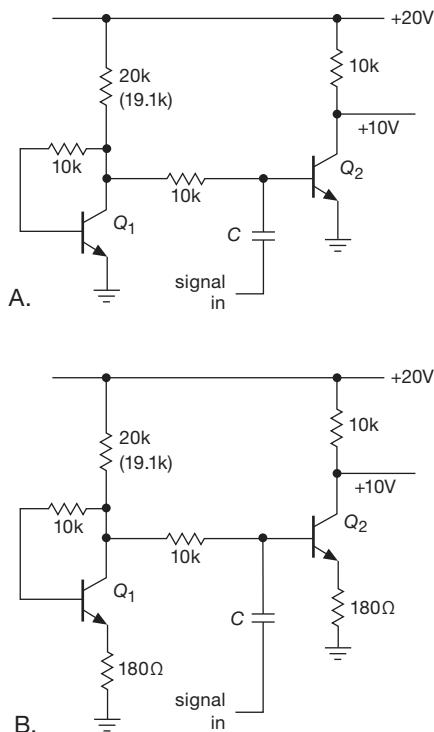
**Figure 2.51.** Equivalent emitter circuit for Figure 2.50.

## B. Matched biasing transistor

You can use a matched transistor to generate the correct base voltage for the required collector current; this ensures automatic temperature compensation (Figure 2.52).<sup>38</sup>  $Q_1$ 's collector is drawing 1 mA, since it is guaranteed to be near ground (about one  $V_{BE}$  drop above ground, to be exact); if  $Q_1$  and  $Q_2$  are a matched pair (available as a single device, with the two transistors on one piece of silicon), then  $Q_2$  will also be biased to draw 1 mA, putting its collector at

<sup>38</sup> R. Widlar, "Some circuit design techniques for linear integrated circuits," *IEEE Trans. Circuit Theory* **CT-12**, 586 (1965). See also US Patent 3,364,434.

+10 volts and allowing a full  $\pm 10$  V symmetrical swing on its collector. Changes in temperature are of no importance as long as both transistors are at the same temperature. This is a good reason for using a “monolithic” dual transistor.



**Figure 2.52.** Biasing scheme with compensated  $V_{BE}$  drop, for both grounded emitter (A) and degenerated emitter (B) stages. With the values shown,  $V_C$  would be approximately 10.5 V; reducing the 20k resistor to 19.1k (a standard value) would take into account the effects of  $V_{BE}$  and finite  $\beta$  and put  $V_C$  at 10 V.

### C. Feedback at dc

You can use dc feedback to stabilize the quiescent point. Figure 2.53A shows one method. By taking the bias voltage from the collector, rather than from  $V_{CC}$ , you get some measure of bias stability. The base sits one diode drop above ground – and because its bias comes from a 10:1 divider, the collector must be at 11 diode drops above ground, or about 7 volts. Any tendency for the transistor to saturate (e.g., if it happens to have unusually high beta) is stabilized, since the dropping collector voltage will reduce the base bias. This scheme is acceptable if great stability is not required. The quiescent point is liable to drift a volt or so as the ambient (surrounding) temperature changes, because the base–emitter voltage has a significant temperature coefficient (Ebers–Moll, again). Better stability is possible if

several stages of amplification are included within the feedback loop. You will see examples later in connection with feedback.

A better understanding of feedback is really necessary to understand this circuit. For instance, feedback acts to reduce the input and output impedances. The input signal sees  $R_1$ ’s resistance effectively reduced by the voltage gain of the stage. In this case it is equivalent to a resistor of about  $200\Omega$  to ground (not pleasant at all!). Later in this chapter (and again in Chapter 4) we treat feedback in enough detail so that you will be able to figure the voltage gain and terminal impedance of this circuit.

Figures 2.53B–D illustrate some variations on the basic dc-feedback bias scheme: circuit B adds some emitter degeneration to improve linearity and predictability of gain; circuit C adds to that an input follower to increase the input impedance (with appropriately increased  $R_1R_2$  divider values and changed ratio to accommodate the additional  $V_{BE}$  drop); and circuit D combines the methods of Figure 2.51 with circuit B to achieve greater bias stability.

Note that the base bias resistor values in these circuits could be increased to raise the input impedance, but you should then take into account the non-negligible base current. Suitable values might be  $R_1 = 220\text{k}$  and  $R_2 = 33\text{k}$ . An alternative approach might be to bypass the feedback resistance in order to eliminate feedback (and therefore lowered input impedance) at signal frequencies (Figure 2.54).<sup>39</sup>

### D. Comments on biasing and gain

One important point about grounded-emitter amplifier stages: you might think that the voltage gain can be raised by increasing the quiescent current, since the intrinsic emitter resistance  $r_e$  drops with rising current. Although  $r_e$  does decrease with increasing collector current, the smaller collector resistor you need to obtain the same quiescent collector voltage just cancels the advantage. In fact, you can show that the small-signal voltage gain of a grounded-emitter amplifier biased to  $0.5V_{CC}$  is given by  $G = 20V_{CC}$  (in volts), independent of quiescent current.

**Exercise 2.15.** Show that the preceding statement is true.

If you need more voltage gain in one stage, one approach is to use a current source as an *active load*. Because its impedance is very high, single-stage voltage gains of

<sup>39</sup> But *caution*: the cascaded  $RC$  sections ( $33\text{k}$  into  $10\mu\text{F}$ ,  $33\text{k}$  into the input capacitor) can cause peaking or instability, unless care is taken (for example by avoiding similar  $RC$  products).

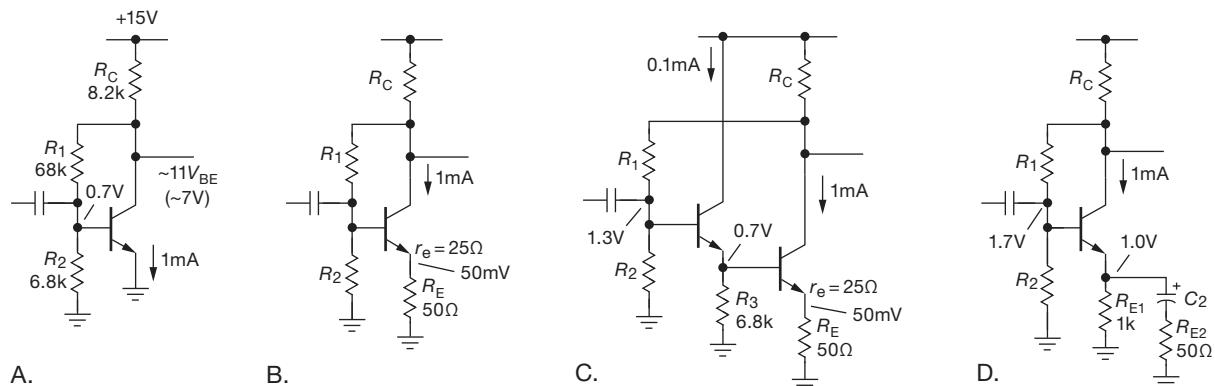


Figure 2.53. Bias stability is improved by feedback.

1000 or more are possible.<sup>40</sup> Such an arrangement cannot be used with the biasing schemes we have discussed, but must be part of an overall dc feedback loop, a subject we will discuss in Chapter 4. You should be sure such an amplifier looks into a high-impedance load; otherwise the gain obtained by high collector load impedance will be lost. Something like an emitter follower, an FET, or an op-amp presents a good load.

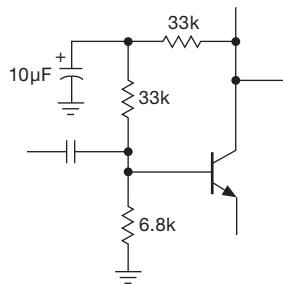
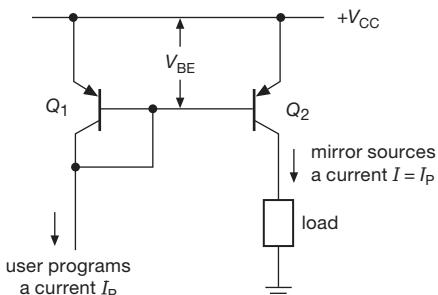


Figure 2.54. Eliminating impedance-lowering feedback at signal frequencies.

In RF amplifiers intended for use only over a narrow frequency range, it is common to use a parallel  $LC$  circuit as a collector load. In that case a very high voltage gain is possible since the  $LC$  circuit has high impedance (like a current source) at the signal frequency, with low impedance at dc. Because the  $LC$  is “tuned,” out-of-band interfering signals (and distortion) are effectively rejected. Additional bonuses are the possibility of peak-to-peak (pp) output swings of  $2V_{CC}$ , and the use of transformer coupling from the inductor.

<sup>40</sup> Ultimately limited by the transistor's finite collector output resistance (a consequence of the Early effect); see the Early effect discussion in Chapter 2x.

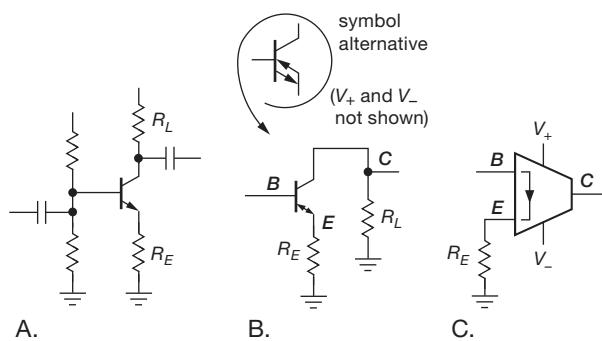
**Exercise 2.16.** Design a tuned common-emitter amplifier stage to operate at 100 kHz. Use a bypassed emitter resistor, and set the quiescent current at 1.0 mA. Assume  $V_{CC} = +15$  V and  $L = 1.0$  mH, and put a 6.2k resistor across the  $LC$  to set  $Q = 10$  (to get a 10% bandpass; see §1.7.14). Use capacitive input coupling.

Figure 2.55. Classic bipolar-transistor matched-pair current mirror. Note the common convention of referring to the positive supply as  $V_{CC}$ , even when  $pnp$  transistors are used.

### 2.3.6 An aside: the perfect transistor

Looking at BJT transistor properties like the non-zero (and temperature-dependent)  $V_{BE}$ , the finite (and current-dependent) emitter impedance  $r_e$  and transconductance  $g_m$ , the collector current that varies with collector voltage (Early effect) etc., one is tempted to ask which transistor is better? Is there a “best” transistor, or perhaps even a *perfect* transistor? If you go through our transistor tables, e.g., Tables 2.1 and 2.2, and especially Table 8.1 for small-signal transistors, you'll see there is no best transistor candidate. That's because all physical bipolar transistors are subject to the same device physics, and their parameters tend to scale with die size and current, etc.

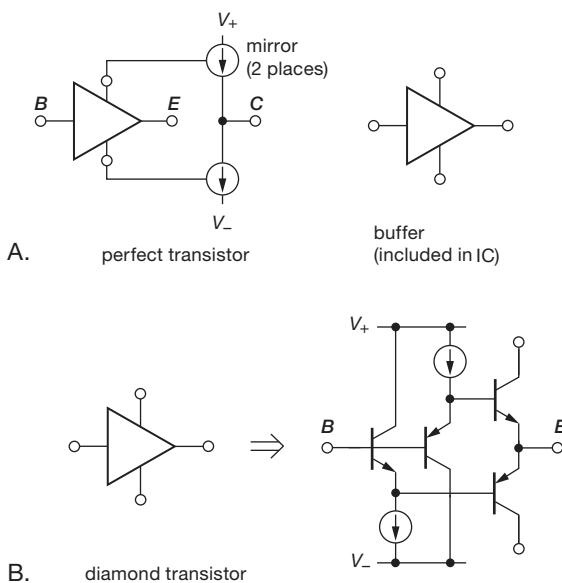
However, it turns out that there *is* a candidate for a “perfect transistor,” if you don’t limit yourself to a single *npn* or *pnp* structure; see Figure 2.56. This device has nearly ideal properties:  $V_{BE}=0$  V (!), along with very high  $g_m$  (thus low  $r_e$ ), and very high beta. And just to top it off, current can flow in either direction – it’s ambidextrous, or “bipolarity” (saying it’s bipolarity is better than saying it’s a bipolar bipolar transistor). Like a regular BJT, it’s a transconductance device: when driven with a positive  $V_{BE}$  input signal, it sources an output current that is  $g_m$  times greater, and vice versa (with a negative  $V_{BE}$  it sinks a current). Unlike a BJT, though, it’s noninverting. All signals are referenced to ground. Very nice.



**Figure 2.56.** A. An ordinary common-emitter BJT amplifier stage, with an emitter-degeneration resistor  $R_E$  and a load resistor  $R_L$ . B. In a common-emitter amplifier built with the “perfect” transistor, all signals are referenced to ground, to which the load  $R_L$  is also returned (the power supplies are not shown). C. The OTA symbol for the perfect transistor, implemented as an Operational Transconductance Amplifier device. The truncated apex symbol means the device has a current output.

How does the perfect transistor work? Figure 2.57 shows a four-transistor circuit known as a *diamond transistor stage*. This circuit is a variation of the cascaded *pnpn-pnppn* emitter follower in Figure 2.29: a complementary *nppn-pnppn* input follower is wired in parallel, and biased with current sources; the emitter outputs (2 $V_{BE}$  apart) drive a matched push-pull output follower, which therefore runs at the same quiescent current. The common node is the effective emitter,  $E$ . Finally, a pair of current mirrors brings the two individual collector currents to a common output, the effective collector,  $C$ , where the output current is zero if the input voltage (between terminals  $B$  and  $E$ ) is zero. As with an ordinary BJT, any current into (or out of) the emitter has to appear at the collector. The part does require two power supply connections. We’ll have more to say about this interesting component in Chapters 2x and 4x.

Texas Instruments calls their perfect transistor (its part



**Figure 2.57.** A. The OPA860 perfect transistor includes a diamond transistor (the triangle) and a pair of current mirrors. A second diamond transistor acts as an output buffer. B. The diamond transistor consists of a complementary pair of matched offset-cancelled emitter followers.

number is OPA860<sup>41</sup>) an Operational Transconductance Amplifier (OTA). Other names they use are “Voltage-Controlled Current source,” “Transconductor,” “Macro Transistor,” and “positive second-generation current conveyor” (CCII+). We fear it has a branding identity crisis, so, with characteristic understatement, we’re calling it a “perfect transistor.”

How close to perfection do these parts come? The OPA860 and OPA861 perfect transistors have these specs:  $V_{os}=3$  mV typ (12 mV max),  $g_m=95$  mS,  $r_e=10.5\Omega$ ,  $Z_{out}=54k\Omega||2\text{ pF}$ ,  $Z_{in}=455k\Omega||2\text{ pF}$ ,  $I_{out(max)}=\pm 15$  mA. Its maximum gain is 5100. Hardly perfect, but, hey, not half bad. You can create many nice circuits with these puppies (e.g., active filter, wideband current summing circuit, or integrator for nanosecond-scale pulses); see the OPA860 datasheet for details.

<sup>41</sup> TI’s OP861 version omits the output buffer and is available in a small SOT-23 package. That’s one of our favorite surface-mount package styles, available for many of the other transistors mentioned in our tables. Knowledgeable readers will recognize the circuitry from inside a current-feedback, or CFB opamp. Some of these devices (for example the AD844) make the internal node available.

### 2.3.7 Current mirrors

The technique of matched base-emitter biasing can be used to make what is called a *current mirror*, an interesting current-source circuit that simply reverses the sign of a “programming” current. (Figure 2.55). You program the mirror by sinking a current from  $Q_1$ ’s collector. That causes a  $V_{BE}$  for  $Q_1$  appropriate to that current at the circuit temperature and for that transistor type.  $Q_2$ , matched to  $Q_1$ ,<sup>42</sup> is thereby programmed to source the same current to the load. The small base currents are unimportant.<sup>43</sup>

One nice feature of this circuit is voltage compliance of the output transistor current source to within a few tenths of a volt of  $V_{CC}$ , as there is no emitter resistor drop to contend with. Also, in many applications it is handy to be able to program a current with a current. An easy way to generate the control current  $I_P$  is with a resistor (Figure 2.58). Because the bases are a diode drop below  $V_{CC}$ , the 14.4k resistor produces a control current, and therefore an output current, of 1 mA. Current mirrors can be used in transistor circuits whenever a current source is needed. They’re very popular in integrated circuits, where (a) matched transistors abound and (b) the designer tries to make circuits that will work over a large range of supply voltages. There are even resistorless IC op-amps in which the operating current of the whole amplifier is set by one external resistor, with all the quiescent currents of the individual amplifier stages inside being determined by current mirrors.

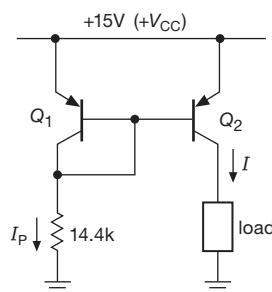


Figure 2.58. Programming current-mirror current.

#### A. Current-mirror limitations due to the Early effect

One problem with the simple current mirror is that the output current varies a bit with changes in output voltage, i.e.,

<sup>42</sup> A monolithic dual transistor is ideal; Table 8.1b on page 502 lists most available matched transistors. Some, such as the DMMT3904 and 3906, are matched to 1 mV and are quite affordable, \$0.36 in small quantities.

<sup>43</sup> This circuit is often called the Widlar current mirror; see the reference on page 97 and US Patent 3,320,439.

the output impedance is not infinite. This is because of the slight variation of  $V_{BE}$  with collector voltage at a given current in  $Q_2$  (which is due to the Early effect); said a different way, the curve of collector current versus collector-emitter voltage at a fixed base-emitter voltage is not flat (Figure 2.59). In practice, the current might vary 25% or so over the output compliance range – much poorer performance than that of the current source with an emitter resistor discussed earlier.

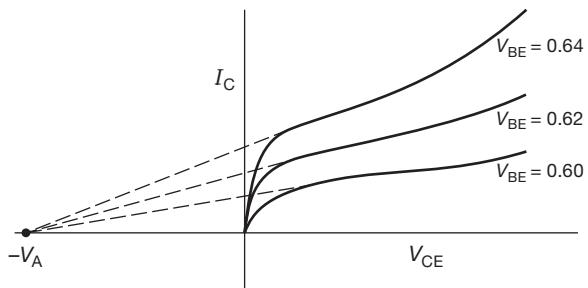


Figure 2.59. Early effect: collector current varies with  $V_{CE}$ . (Interestingly, you get a very similar curve, with comparable  $V_A$ , if you apply instead a family of constant base currents.)

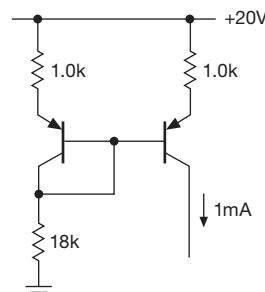
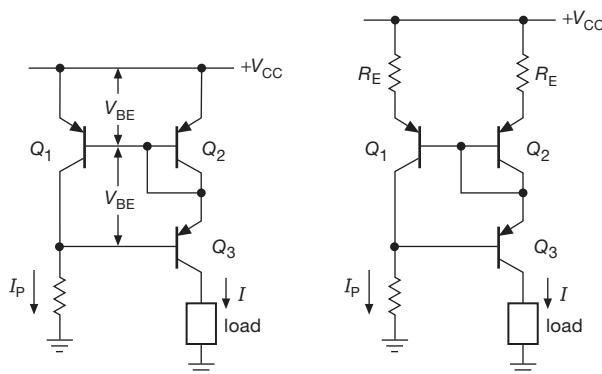


Figure 2.60. Improved current mirror with emitter resistors.

One solution, if a better current source is needed (it often isn’t), is the circuit shown in Figure 2.60. The emitter resistors are chosen to have at least a few tenths of a volt drop; this makes the circuit a far better current source, since the small variations of  $V_{BE}$  with  $V_{CE}$  are now negligible in determining the output current. Again, matched transistors should be used. Note that this circuit loses its effectiveness if operation over a wide range of programming current is intended (figure out why).<sup>44</sup>

<sup>44</sup> Current sources and current mirrors are discussed in more detail in Chapter 2x.



**Figure 2.61.** Wilson current mirror. Good stability with load variations is achieved through cascode transistor  $Q_3$ , which reduces voltage variations across  $Q_1$ . Adding a pair of emitter resistors  $R_E$ , as shown, reduces output current error caused by  $V_{BE}$  mismatch, when chosen such that  $I_P R_E$  is of order 100 mV or more.

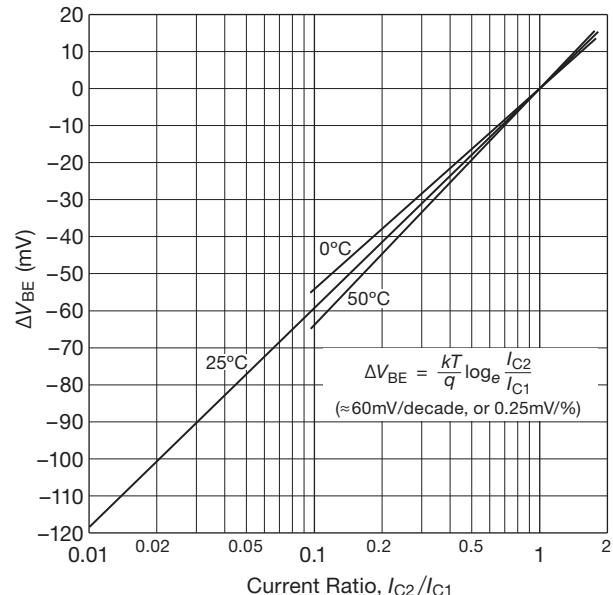
### B. Wilson mirror

Another current mirror with improved consistency of current is shown in the clever circuit of Figure 2.61.  $Q_1$  and  $Q_2$  are in the usual mirror configuration, but  $Q_3$  now keeps  $Q_1$ 's collector fixed at two diode drops below  $V_{CC}$ . That circumvents the Early effect in  $Q_1$ , whose collector is now the programming terminal, with  $Q_2$  now sourcing the output current. The result is that both current-determining transistors ( $Q_1$  and  $Q_2$ ) have fixed collector-emitter drops; you can think of  $Q_3$  as simply passing the output current through to a variable-voltage load (a similar trick is used in the cascode connection, which you will see later in the chapter). Transistor  $Q_3$ , by the way, does not have to be matched to  $Q_1$  and  $Q_2$ ; but if it has the same beta, then you get an exact cancellation of the (small) base current error that afflicts the simple mirror of Figure 2.55 (or the beta-enhanced mirror in Chapter 2x).

**Exercise 2.17.** Show that this statement is true.

There are additional nice tricks you can do with current mirrors, such as generating multiple independent outputs, or an output that is a fixed multiple of the programming current. One trick (invented by the legendary Widlar) is to unbalance the  $R_E$ 's in Figure 2.61; as a rough estimate, the output current ratio is approximately the ratio of resistor values (because the base-emitter drops are roughly equal). But to get it right you need to take into account the difference of  $V_{BE}$ 's (because the transistors are running at different currents), for which the graph in Figure 2.62 is helpful. This graph is also useful for estimating the current unbalance in a current mirror built with discrete (i.e.,

not matched) transistors. We treat current mirrors further in Chapter 2x (§§2x.3 and 2x.11).



**Figure 2.62.** Collector current ratios for matched transistors, as determined by the difference in applied base-emitter voltages. See Table 8.1b for low-noise matched BJTs.

### 2.3.8 Differential amplifiers

The differential amplifier is a very common configuration used to amplify the difference voltage between two input signals. In the ideal case the output is entirely independent of the individual signal levels – only the difference matters.

Differential amplifiers are important in applications in which weak signals are contaminated by “pickup” and other miscellaneous noise. Examples include digital and RF signals transferred over twisted-pair cables, audio signals (the term “balanced” means differential, usually  $600\Omega$  impedance, in the audio business), local-area-network signals (such as 100BASE-TX and 1000BASE-T Ethernet), electrocardiogram voltages, magnetic disk head amplifiers, and numerous other applications. A differential amplifier at the receiving end restores the original signal if the interfering “common-mode” signals (see below) are not too large. Differential amplifiers are universally used in operational amplifiers, an essential building block that is the subject of Chapter 4. They’re very important in dc amplifier design (amplifiers that amplify clear down to dc, i.e., have no coupling capacitors) because their symmetrical design is inherently compensated against thermal drifts.

*Some nomenclature:* when both inputs change levels together, that's a *common-mode* input change. A differential change is called *normal mode*, or sometimes *differential mode*. A good differential amplifier has a high *common-mode rejection ratio* (CMRR), the ratio of response for a normal-mode signal to the response for a common-mode signal of the same amplitude. CMRR is usually specified in decibels. The common-mode input range is the voltage level over which the inputs may vary. The differential amplifier is sometimes called a “long-tailed pair.”

Figure 2.63 shows the basic circuit. The output is taken off one collector with respect to ground; that is called a *single-ended output* and is the most common configuration. You can think of this amplifier as a device that amplifies a difference signal and converts it to a single-ended signal so that ordinary subcircuits (followers, current sources, etc.) can make use of the output. (If, instead, a differential output is desired, it is taken between the collectors.)

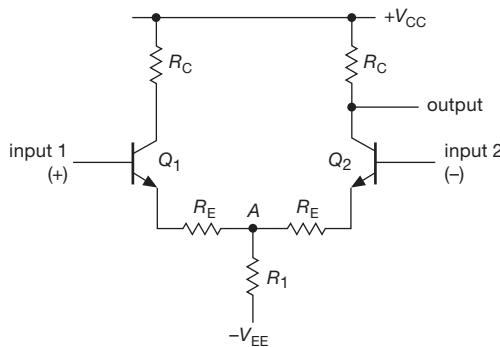


Figure 2.63. Classic transistor differential amplifier.

What is the gain? That's easy enough to calculate: imagine a symmetrical input signal wiggle, in which input 1 rises by  $v_{in}$  (a small-signal variation) and input 2 drops by the same amount. As long as both transistors stay in the active region, point A remains fixed. You then determine the gain as with the single transistor amplifier, remembering that the input change is actually twice the wiggle on either base:  $G_{diff} = R_C / 2(r_e + R_E)$ . Typically  $R_E$  is small,  $100\ \Omega$  or less, or it may be omitted entirely. Differential voltage gains of a few hundred are possible.

You can determine the common-mode gain by putting identical signals  $v_{in}$  on both inputs. If you think about it correctly<sup>45</sup> (remembering that  $R_1$  carries both emitter currents), you'll find  $G_{CM} = -R_C / (2R_1 + R_E)$ . Here we've ig-

nored the small  $r_e$ , because  $R_1$  is typically large, at least a few thousand ohms. We really could have ignored  $R_E$  as well. The CMRR is thus roughly  $R_1 / (r_e + R_E)$ . Let's look at a typical example (Figure 2.64) to get some familiarity with differential amplifiers.

Collector resistor  $R_C$  is chosen for a quiescent current of  $100\ \mu A$ . As usual, we put the collector at  $0.5V_{CC}$  for large dynamic range.  $Q_1$ 's collector resistor can be omitted, since no output is taken there.<sup>46</sup>  $R_1$  is chosen to give total emitter current of  $200\ \mu A$ , split equally between the two sides when the (differential) input is zero. From the formulas just derived, this amplifier has a differential gain of 10 and a common-mode gain of 0.55. Omitting the  $1.0k$  resistors raises the differential gain to 50, but drops the (differential) input impedance from about  $250k$  to about  $50k$  (you can substitute Darlington transistors<sup>47</sup> in the input stage to raise the impedance into the megohm range, if necessary).

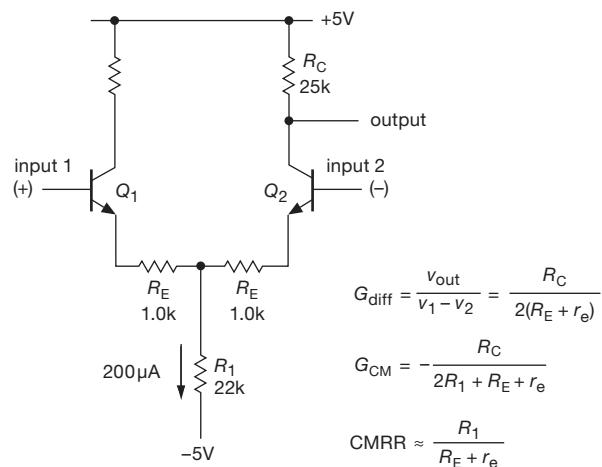


Figure 2.64. Calculating differential amplifier performance.

Remember that the maximum gain of a single-ended grounded emitter amplifier biased to  $0.5V_{CC}$  is  $20V_{CC}$  (in volts). In the case of a differential amplifier the maximum differential gain ( $R_E = 0$ ) is half that figure, or (for arbitrary quiescent point) 20 times the voltage (in volts) across the collector resistor. The corresponding maximum CMRR (again with  $R_E = 0$ ) is equal to 20 times the voltage (in volts) across  $R_1$ . As with the single-ended common-emitter amplifier, the emitter resistors  $R_E$  reduce distortion, at the

<sup>45</sup> Hint: replace  $R_1$  with a parallel pair, each of resistance  $2R_1$ ; then notice that you can cut the wire connecting them together at point A (because no current flows); take it from there.

<sup>46</sup> See §2.4.2.

expense of gain. See the extensive discussion of BJT amplifier distortion in Chapter 2x.

**Exercise 2.18.** Verify that these expressions are correct. Then design a differential amplifier to run from  $\pm 5$  V supply rails, with  $G_{\text{diff}} = 25$  and  $R_{\text{out}} = 10\text{k}$ . As usual, put the collector's quiescent point at half of  $V_{\text{CC}}$ .

### A. Biasing with a current source

The common-mode gain of the differential amplifier can be reduced enormously by the substitution of a current source for  $R_1$ . Then  $R_1$  effectively becomes very large and the common-mode gain is nearly zero. If you prefer, just imagine a common-mode input swing; the emitter current source maintains a constant total emitter current, shared equally by the two collector circuits, by symmetry. The output is therefore unchanged. Figure 2.65 shows an example. The CMRR of this circuit, using an LM394 monolithic transistor pair for  $Q_1$  and  $Q_2$ , will be around 100,000:1 (100 dB) at dc. The common-mode input range for this circuit goes from  $-3.5$  V to  $+3$  V; it is limited at the low end by the compliance of the emitter current source and at the high end by the collector's quiescent voltage.<sup>48</sup>

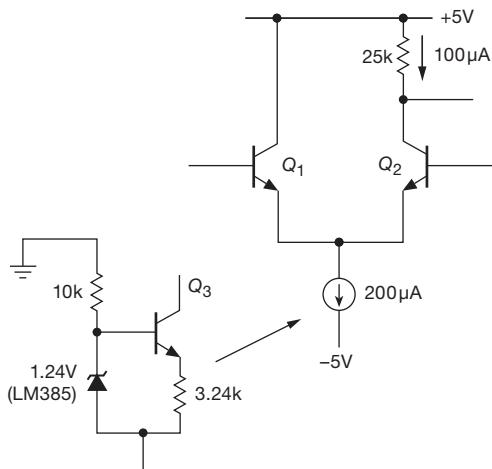


Figure 2.65. Improving CMRR of the differential amplifier with a current source.

Be sure to remember that this amplifier, like all transistor amplifiers, must have a dc bias path to the bases. If the input is capacitively coupled, for instance, you must have

base resistors to ground. An additional caution for differential amplifiers, particularly those without inter-emitter resistors: bipolar transistors can tolerate only 6 volts of base–emitter reverse bias before breakdown; thus, applying a differential input voltage larger than this will destroy the input stage (if there is no inter-emitter resistor). An inter-emitter resistor limits the breakdown current and prevents destruction, but the transistors may be degraded nonetheless (in beta, noise, etc.). In either case the input impedance drops drastically during reverse conduction.

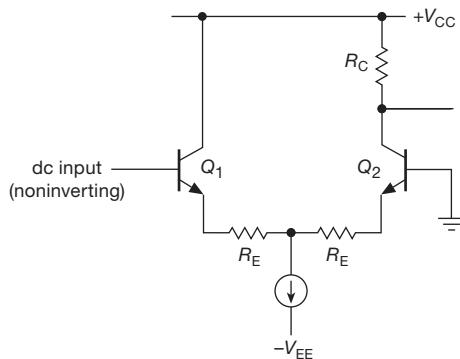
An interesting aside: the emitter current sink shown in Figure 2.65 has some variation with temperature, because  $V_{\text{BE}}$  decreases with increasing temperature (to the tune of approximately  $-2.1 \text{ mV}/^{\circ}\text{C}$ , §2.3.2), causing the current to increase. More explicitly, if we call the 1.24 V zener-like reference “ $V_{\text{ref}}$ ,” then the drop across the emitter resistor equals  $V_{\text{ref}} - V_{\text{BE}}$ ; the current is proportional, thus increasing with temperature. As it happens, this is in fact *beneficial*: it can be shown from basic transistor theory that the quantity  $V_{g0} - V_{\text{BE}}$  is approximately proportional to absolute temperature (PTAT), where  $V_{g0}$  is the silicon bandgap voltage (extrapolated to absolute zero), approximately 1.23 V. So, by choosing our  $V_{\text{ref}}$  voltage equal to the bandgap voltage, we have an emitter current that increases PTAT; and this cancels the temperature dependence of differential-pair voltage gain ( $g_m \propto 1/T_{\text{abs}}$ , §2.3.2). We’ll explore this sort of cleverness a bit more in §9.10.2. And in Chapter 9 there’s extensive discussion of the differential amplifier and the closely related “instrumentation amplifier.”

### B. Use in single-ended dc amplifiers

A differential amplifier makes an excellent dc amplifier, even for single-ended inputs. You just ground one of the inputs and connect the signal to the other (Figure 2.66). You might think that the “unused” transistor could be eliminated. Not so! The differential configuration is inherently compensated for temperature drifts, and even when one input is at ground that transistor is still doing something: a temperature change causes both  $V_{\text{BE}}$ ’s to change the same amount, with no change in balance or output. That is, changes in  $V_{\text{BE}}$  are not amplified by  $G_{\text{diff}}$  (only by  $G_{\text{CM}}$ , which can be made essentially zero). Furthermore, the cancellation of  $V_{\text{BE}}$ ’s means that there are no 0.6 V drops at the input to worry about. The quality of a dc amplifier constructed this way is limited only by mismatching of input  $V_{\text{BE}}$ ’s or their temperature coefficients. Commercial monolithic transistor pairs and commercial differential amplifier ICs are available with extremely good matching (e.g., the MAT12 *npn* monolithic matched pair has a typical drift of

<sup>48</sup> You can make good current sinks also with JFETs (see the discussion in §3.2.2C), but BJTs are better for this task in many ways. See for example Figure 3.26, where we show four configurations of BJT current sinks that improve upon the JFET alternative.

$V_{BE}$  between the two transistors of  $0.15 \mu\text{V}/^\circ\text{C}$ ). See Table 8.1b on page 502 for a listing of matched BJTs.



**Figure 2.66.** A differential amplifier can be used as a precision single-ended dc amplifier.

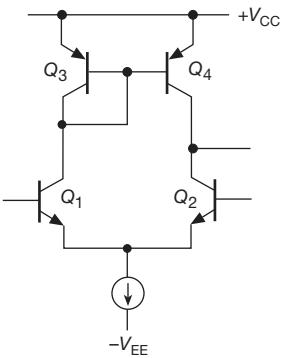
Either input could have been grounded in the preceding circuit example. The choice depends on whether or not the amplifier is supposed to invert the signal. (The configuration shown is preferable at high frequencies, however, because of the *Miller effect*; see §2.4.5.) The connection shown is noninverting, and so the inverting input has been grounded. This terminology carries over to op-amps, which are versatile high-gain differential amplifiers.

### C. Current-mirror active load

As with the simple grounded-emitter amplifier, it is sometimes desirable to have a single-stage differential amplifier with very high gain. An elegant solution is a current-mirror active load (Figure 2.67).  $Q_1Q_2$  is the differential pair with emitter current source.  $Q_3$  and  $Q_4$ , a current mirror, form the collector load. The high effective collector load impedance provided by the mirror yields voltage gains of 5000 or more, assuming no load at the amplifier's output.<sup>49</sup> Such an amplifier is very common as the input stage in a larger circuit, and is usually used only within a feedback loop, or as a comparator (discussed in the next section). Be sure to keep the load impedance of such an amplifier very high, or the gain will drop enormously.

### D. Differential amplifiers as phase splitters

The collectors of a symmetrical differential amplifier generate equal signal swings of opposite phase. By taking outputs from both collectors, you've got a phase splitter. Of course, you could also use a differential amplifier with both



**Figure 2.67.** Differential amplifier with active current mirror load.

differential inputs and differential outputs. This differential output signal could then be used to drive an additional differential amplifier stage, with greatly improved overall common-mode rejection.

### E. Differential amplifiers as comparators

Because of its high gain and stable characteristics, the differential amplifier is the main building block of the *comparator* (which we saw in §1.4.2E), a circuit that tells which of two inputs is larger. They are used for all sorts of applications: switching on lights and heaters, generating square waves from triangles, detecting when a level in a circuit exceeds some particular threshold, class-D amplifiers and pulse-code modulation, switching power supplies, etc. The basic idea is to connect a differential amplifier so that it turns a transistor switch on or off, depending on the relative levels of the input signals. The linear region of amplification is ignored, with one or the other of the two input transistors cut off at any time. A typical hookup is illustrated in §2.6.2 by a temperature-controlling circuit that uses a resistive temperature sensor (thermistor).

## 2.4 Some amplifier building blocks

We've now seen most of the basic – and important – transistor circuit configurations: switch, follower, current source (and mirror), and common-emitter amplifier (both single-ended and differential). For the remainder of the chapter we look at some circuit elaborations and their consequences: push-pull, Darlington and Sziklai, bootstrapping, Miller effect, and the cascode configuration. We'll finish with an introduction to the wonderful (and essential) technique of *negative feedback*. Chapter 2x deals with follow-on transistor circuits and techniques at a greater level of sophistication.

<sup>49</sup> The dc gain is limited primarily by the Early effect; see §2.3.2 and the discussion in Chapter 2x.

**Table 2.2 Bipolar Power Transistors<sup>a</sup>**

NPN	PNP	Case	$V_{CEO}$ max (V)	$I_c$ max <sup>b</sup> (A)	$P_{diss}$ max <sup>b,h</sup> (W)	$R_{\theta JC}^c$ (°C/W)	$h_{FE}$ min	$h_{FE}$ typ	at $I_c$ (A)	$f_T$ min (MHz)	multiple manf?
<b><i>standard BJT</i></b>											
BD139	BD140	TO-126	80	1.5	12.5	10	40 <sup>e</sup>	100	0.15	50	•
2N3055	2N2955	TO-3	60	15	115	1.5	20	--	4	2.5	•
2N6292	2N6107	TO-220	70	7	40	3.1	30	--	2	4	-
TIP31C	TIP32C	TO-220	100	3	40	3.1	25	100	1	3	•
TIP33C	TIP34C	TO-218 <sup>d</sup>	100	10	80	1.6	40	100	1	3	•
TIP35C	TIP36C	TO-218 <sup>d</sup>	100	25	125	1.0	25	150	1.5	3	•
MJ15015	MJ15016	TO-3	120	15	180	1.0	20	35	4	0.8 <sup>g</sup>	-
MJE15030	MJE15031	TO-220	150	8	50	2.5	40	80	3	30	•,z
MJE15032	MJE15033	TO-220	250	8	50	2.5	50	100	1	30	•
2SC5200	2SA1943	TO-264	230	17	150	0.8	55	80	1	30	•
2SC5242 <sup>k</sup>	2SA1962 <sup>k</sup>	TO-3P	250	s	s	s	s	s	s	s	•
MJE340	MJE350	TO-126	300	0.5	20	6	30	--	0.05	--	•
TIP47	MJE5730	TO-220	250	1	40	3.1	30	--	0.3	10	•
TIP50 <sup>u</sup>	MJE5731A <sup>u</sup>	TO-220	400	s	s	s	--	s	s	s	•
MJE13007	MJE5852	TO-220	400 <sup>f</sup>	8	80	1.6	8 <sup>g</sup>	20 <sup>g</sup>	2	14 <sup>t</sup>	-
<b><i>Darlington</i></b>											
MJD112	MJD117	DPak	100	2	20	6.3	1000	2000	2	25	•
TIP122	TIP127	TO-220	100	5	65	1.9	1000	--	3	--	•
TIP142	TIP147	TO-218	100	10	125	1.0	1000	--	5	--	•
MJ11015	MJ11016	TO-3	120	30	200	0.9	1000	--	20	4	•
MJ11032	MJ11033	TO-3	120	50	300	0.6	1000	--	25	--	•
MJH11019	MJH11020	TO-218	200 <sup>v</sup>	15	150	0.8	400	--	10	3	•

Notes: (a) sorted more or less by voltage, current and families; see also additional tables in Chapter 2x. (b) with case at 25°C. (c)  $P_{diss}(\text{reality}) = (T_J[\text{your-max-value}] - T_{\text{amb}}) / (R_{\theta JC} + R_{\text{ECS}} + R_{\text{ESA}})$ ; this is a much lower number than the "spec," especially if you're careful with  $T_J$  max, say 100°C. (d) similar to TO-247. (e) higher gain grades are available. (f) much higher  $V_{CES}$  "blocking" capability (compared with  $V_{CEO}$ ), e.g. 700V for MJE13007. (g) higher for the PNP device. (h)  $P_{diss}(\text{max}) = (150^\circ\text{C} - 25^\circ\text{C}) / R_{\theta JC}$ ; this is a classic datasheet specsmanship value. (k) larger pkg version of above. (s) same as above. (t) typical. (u) higher voltage version of above. (v) there are also 150V and 250V versions. (z) if these are hard to get, try the '028 and '029 versions (120V rather than 150V).

### 2.4.1 Push-pull output stages

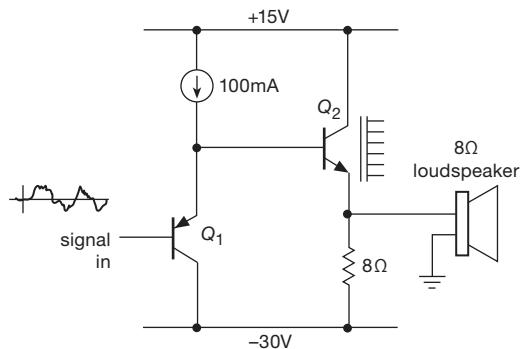
As we mentioned earlier in the chapter, an *npn* emitter follower cannot sink current and a *pnp* follower cannot source current. The result is that a single-ended follower operating between split supplies can drive a ground-return load only if a high quiescent current is used.<sup>50</sup> The quiescent current must be at least as large as the maximum output current during peaks of the waveform, resulting in high quiescent power dissipation. For example, Figure 2.68 shows a follower circuit to drive an  $8\Omega$  loudspeaker load with up to 10 watts of audio.

An explanation of the driver stage: the *pnp* follower  $Q_1$  is included to reduce drive requirements and to cancel  $Q_2$ 's  $V_{BE}$  offset (0 V input gives approximately 0 V output).

$Q_1$  could, of course, be omitted for simplicity. The hefty current source in  $Q_1$ 's emitter load is used to ensure that there is sufficient base drive to  $Q_2$  at the top of the signal swing. A resistor as emitter load would be inferior because it would have to be a rather low value ( $50\Omega$  or less) in order to guarantee at least 50 mA of base drive to  $Q_2$  at the peak of the swing, when load current would be maximum and the drop across the resistor would be minimum; the resultant quiescent current in  $Q_1$  would be excessive.

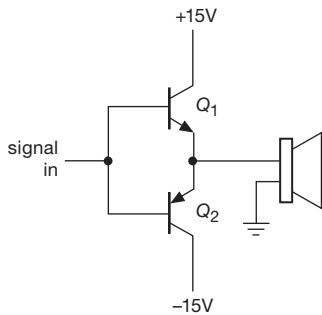
The output of this example circuit can swing to nearly  $\pm 15$  volts (peak) in both directions, giving the desired output power (9 V rms across  $8\Omega$ ). However, the output transistor dissipates 55 watts with no signal (hence the heatsink symbol), and the emitter resistor dissipates another 110 watts. Quiescent power dissipation many times greater than the maximum output power is characteristic of this kind of class-A circuit (transistor always in conduction);

<sup>50</sup> An amplifier in which current flows in the output transistor over the full waveform swing is sometimes called a "class-A" amplifier.



**Figure 2.68.** A 10W loudspeaker amplifier, built with a single-ended emitter follower, dissipates 165W of quiescent power!

this obviously leaves a lot to be desired in applications in which any significant amount of power is involved.



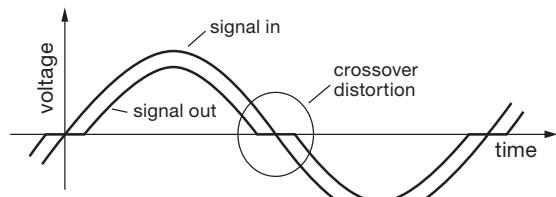
**Figure 2.69.** Push-pull emitter follower.

Figure 2.69 shows a *push-pull* follower doing the same job.  $Q_1$  conducts on positive swings,  $Q_2$  on negative swings. With zero input voltage, there is no collector current and no power dissipation. At 10 watts of output power there is less than 10 watts of dissipation in each transistor.<sup>51</sup>

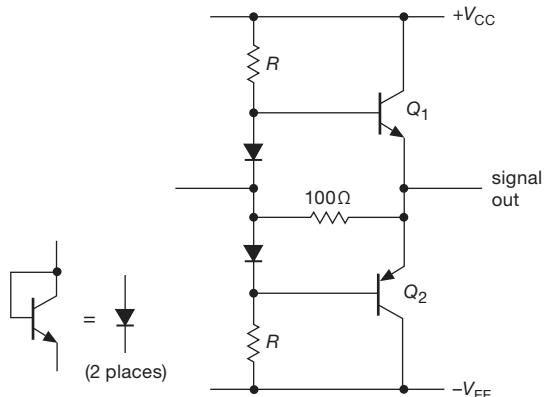
#### A. Crossover distortion in push-pull stages

There is a problem with the preceding circuit as drawn. The output trails the input by a  $V_{BE}$  drop; on positive swings the output is about 0.6 V less positive than the input, and the reverse for negative swings. For an input sine wave, the output would look as shown in Figure 2.70. In the language of the audio business, this is called *crossover distortion*. The best cure (feedback offers another method, although by itself it is not entirely satisfactory; see §4.3.1E) is to bias the push-pull stage into slight conduction, as in Figure 2.71.

The bias resistors  $R$  bring the diodes into forward conduction, holding  $Q_1$ 's base a diode drop above the input



**Figure 2.70.** Crossover distortion in the push-pull follower.



**Figure 2.71.** Biasing the push-pull follower to eliminate crossover distortion.

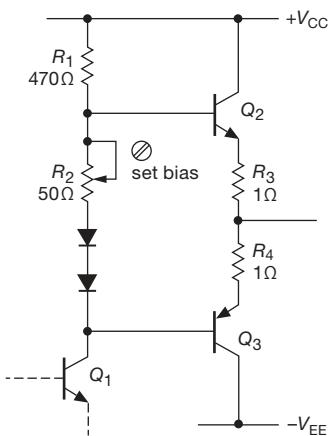
signal and  $Q_2$ 's base a diode drop below the input signal. Now, as the input signal crosses through zero, conduction passes from  $Q_2$  to  $Q_1$ ; one of the output transistors is always on. The value  $R$  of the base resistors is chosen to provide enough base current for the output transistors at the peak output swing. For instance, with  $\pm 20$  V supplies and an  $8\Omega$  load running up to 10 watts sinewave power, the peak base voltage is about 13.5 volts and the peak load current is about 1.6 amps. Assuming a transistor beta of 50 (power transistors generally have lower current gain than small-signal transistors), the 32 mA of necessary base current will require base resistors of about  $220\Omega$  (6.5 V from  $V_{CC}$  to base at peak swing).

In this circuit we've added a resistor from input to output (this could have been done in Figure 2.69 as well). This serves to eliminate the "dead zone" as conduction passes from one transistor to the other (particularly in the first circuit), which is desirable especially when this circuit is included within a larger feedback circuit. However, it does not substitute for the better procedure of linearizing by biasing, as in Figure 2.71, to achieve transistor conduction over the full output waveform. We have more to say about this in Chapter 2x.

<sup>51</sup> An amplifier like this, with half-cycle conduction in each of the output transistors, is sometimes called a "class-B" amplifier.

### B. Thermal stability in class-B push-pull amplifiers

The preceding amplifier has one unfortunate feature: it is not thermally stable. As the output transistors warm up (and they will get hot, because they are dissipating power when signal is applied), their  $V_{BE}$  drops, causing quiescent current to flow. The added heat this produces causes the situation to get worse, with the strong possibility of what is called *thermal runaway* (whether it runs away or not depends on a number of factors, including how large a “heatsink” is used, how well the diode’s temperature tracks the transistor’s temperature, etc.). Even without runaway, better control over the circuit is needed, usually with the sort of arrangement shown in Figure 2.72.



**Figure 2.72.** Adding (small) emitter resistors improve thermal stability in the push-pull follower.

For variety, the input is shown coming from the collector of the previous stage;  $R_1$  now serves the dual purpose of being  $Q_1$ ’s collector resistor, and also providing current to bias the diodes and bias-setting resistor in the push-pull base circuit. Here  $R_3$  and  $R_4$ , typically a few ohms or less, provide a “cushion” for the critical quiescent current biasing: the voltage between the bases of the output transistors must now be a bit greater than two diode drops, and you provide the extra with adjustable biasing resistor  $R_2$  (often replaced with a third series diode, or, better, with the more elegant biasing circuit of Figure 2.78 on page 111). With a few tenths of a volt across  $R_3$  and  $R_4$ , the temperature variation of  $V_{BE}$  doesn’t cause the current to rise very rapidly (the larger the drop across  $R_3$  and  $R_4$ , the less sensitive it is), and the circuit will be stable. Stability is improved by mounting the diodes<sup>52</sup>

in physical contact with the output transistors (or their heatsinks).

You can estimate the thermal stability of such a circuit by remembering that the base-emitter drop decreases by about 2.1 mV for each  $1^{\circ}\text{C}$  rise, and that the collector current increases by a factor of 10 for every 60 mV increase in base-emitter voltage (or 4% per mV). For example, if  $R_2$  were replaced with a diode, you would have three diode drops between the bases of  $Q_2$  and  $Q_3$ , leaving about one diode drop across the series combination of  $R_3$  and  $R_4$ . (The latter would then be chosen to give an appropriate quiescent current, perhaps 100 mA for an audio power amplifier.) The worst case for thermal stability occurs if the biasing diodes are not thermally coupled to the output transistors.

Let us assume the worst and calculate the increase in output-stage quiescent current corresponding to a  $30^{\circ}\text{C}$  temperature rise in output transistor temperature. That’s not a lot for a power amplifier, by the way. For that temperature rise, the  $V_{BE}$  of each output transistor will decrease by about 63 mV at constant current, raising the voltage across  $R_3$  and  $R_4$  by about 50% (i.e., the quiescent current will rise by about 50%). The corresponding figure for the preceding amplifier circuit without emitter resistors (Figure 2.71) will be a factor of 10 rise in quiescent current (recall that  $I_C$  increases a decade per 60 mV increase in  $V_{BE}$ ), i.e., 1000%. The improved thermal stability of this biasing arrangement (even without having the diodes thermally coupled to the output transistors) is evident. And you’ll do significantly better when the diodes (or diode-connected transistors, or, best of all,  $V_{BE}$ -referenced biasing as shown in Figure 2.78) ride along on the heatsink.

This circuit has the additional advantage that, by adjusting the quiescent current, you have some control over the amount of residual crossover distortion. A push-pull amplifier biased in this way to obtain substantial quiescent current at the crossover point is sometimes referred to as a “class-AB” amplifier, meaning that both transistors conduct simultaneously during a significant portion of the cycle. In practice, you choose a quiescent current that is a good compromise between low distortion and excessive quiescent dissipation. Feedback, introduced later in this chapter (and exploited shamelessly, and with joy, in Chapter 4), is almost always used to reduce distortion still further.

We will see a further evolution of this circuit in §2.4.2, where we supplement it with the intriguingly named techniques of  $V_{BE}$ -referenced biasing, collector bootstrapping, and  $\beta$ -boosting complementary Darlington output stage.

<sup>52</sup> Or, better, diode-connected transistors: connect base and collector together as “anode” with emitter as “cathode.”

### C. “Class-D” amplifiers

An interesting solution to this whole business of power dissipation (and distortion) in class-AB linear power amplifiers is to abandon the idea of a linear stage entirely and use instead a *switching* scheme: imagine that the push-pull follower transistors  $Q_2$  and  $Q_3$  in Figure 2.72 are replaced with a pair of transistor *switches*, with one ON and the other OFF at any time, so that the output is switched completely to  $+V_{CC}$  or to  $-V_{CC}$  at any instant. Imagine also that these switches are operated at a high frequency (say at least 10 times the highest audio frequency) and that their relative timing is controlled (by techniques we’ll see later, in Chapters 10–13) such that the *average* output voltage is equal to the desired analog output. Finally, we add an *LC* lowpass filter to kill the high switching signal, leaving the desired (lower-frequency) analog output intact.

This is a *Class-D*, or *switching* amplifier. It has the advantage of very high efficiency, because the switching transistors are either off (no current) or in saturation (near-zero voltage); that is, the power dissipated in the switching transistors (the product  $V_{CE} \times I_C$ ) is always small. There’s also no worry about thermal runaway. The downsides are the problems of emission of high-frequency noise, switching feedthrough to the output, and the difficulty of achieving excellent linearity.

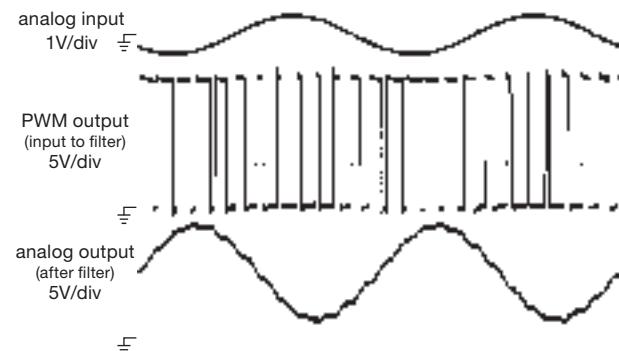
Class-D amplifiers are nearly universal in inexpensive audio equipment, and they are increasingly finding their way into high-end audio equipment. Figure 2.73 shows measured waveforms of an inexpensive (and tiny!) class-D amplifier IC driving a  $5\Omega$  load with a sinewave at the high end of the audio range (20 kHz). This particular IC uses a 250 kHz switching frequency, and can drive 20 watts each into a pair of stereo speakers; pretty much everything you need (except for the output *LC* filters) is on the chip, which costs about \$3 in small quantities. Pretty neat.

### 2.4.2 Darlington connection

If you hook two transistors together as in Figure 2.74, the result – called a *Darlington connection*<sup>53</sup> (or *Darlington pair*) – behaves like a single transistor with beta equal to the product of the two transistor betas.<sup>54</sup> This can be very handy where high currents are involved (e.g., voltage regulators or power amplifier output stages), or for input stages

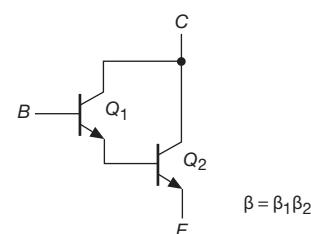
<sup>53</sup> Sidney Darlington, US Patent 2,663,806: “Semiconductor Signal Translating Device.” Darlington wanted the patent to cover any number of transistors in one package, but the lawyers at Bell Laboratories overruled him, thus forgoing a patent that would have covered every IC.

<sup>54</sup> At the operating current of each transistor, of course.



**Figure 2.73.** Class-D amplifier waveforms: a 20 kHz input sinewave controls the “duty cycle” (fraction of time the output is HIGH) of a push-pull switched output. These waveforms are from a TPA3123 stereo amplifier chip running from  $+15\text{ V}$ , and show the prefiltered PWM (pulse-width modulated) output, and the final smoothed output after the *LC* lowpass output filter. Horizontal: 10  $\mu\text{s}/\text{div}$ .

of amplifiers where very high input impedance is necessary.



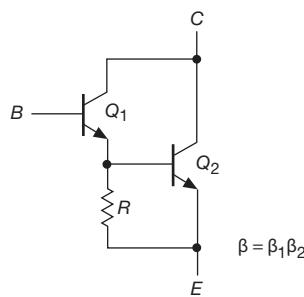
**Figure 2.74.** Darlington transistor configuration.

For a Darlington transistor the base-emitter drop is twice normal and the saturation voltage is at least one diode drop (since  $Q_1$ ’s emitter must be a diode drop above  $Q_2$ ’s emitter). Also, the combination tends to act like a rather slow transistor because  $Q_1$  cannot turn off  $Q_2$  quickly. This problem is usually taken care of by including a resistor from base to emitter of  $Q_2$  (Figure 2.75). Resistor  $R$  also prevents leakage current through  $Q_1$  from biasing  $Q_2$  into conduction;<sup>55</sup> its value is chosen so that  $Q_1$ ’s leakage current (nanoamps for small-signal transistors, as much as hundreds of microamps for power transistors) produces less than a diode drop across  $R$ , and so that  $R$  doesn’t sink a large proportion of  $Q_2$ ’s base current when it has a diode drop across it. Typically  $R$  might be a few hundred ohms in

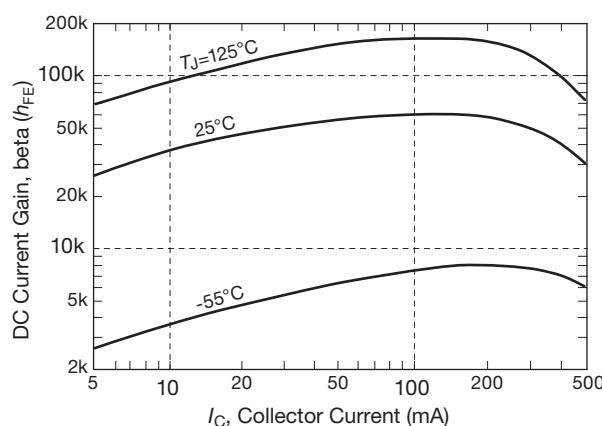
<sup>55</sup> And, by stabilizing  $Q_1$ ’s collector current, it improves the predictability of the Darlington’s total  $V_{BE}$ .

a power transistor Darlington, or a few thousand ohms for a small-signal Darlington.

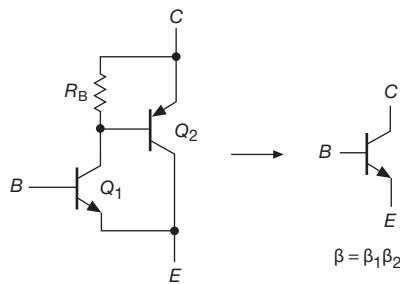
Darlington transistors are available as single packages, usually with the base-emitter resistor included. A typical example is the *npn* power Darlington MJH6284 (and *pnp* cousin MJH6287), with a current gain of 1000 (typical) at a collector current of 10 amps. Another popular power Darlington is the inexpensive *npn* TIP142 (and *pnp* cousin TIP147); these cost \$1 in small quantities and have typical  $\beta=4000$  at  $I_C=5\text{ A}$ . And for small-signal applications we like the widely available MPSA14 or MMBTA14 (in TO-92 and SOT23 packages, respectively), with a minimum beta of 10,000 at 10 mA and 20,000 at 100 mA. These 30 volt parts have no internal base-emitter resistor (so you can use them at very low currents); they cost less than \$0.10 in small quantities. Figure 2.76 shows beta versus collector current for these parts; note the pleasantly high values of beta, but with substantial dependence both on temperature and on collector current.



**Figure 2.75.** Improving turn-off speed in a Darlington pair. (The beta formula is valid as long as  $R$  does not rob significant base current from  $Q_2$ .)



**Figure 2.76.** Typical beta versus collector current for the popular MPSA14 *npn* Darlington (adapted from the datasheet).



**Figure 2.77.** Sziklai connection ("complementary Darlington").

### A. Sziklai connection

A similar beta-boosting configuration is the Sziklai connection,<sup>56</sup> sometimes referred to as a complementary Darlington (Figure 2.77). This combination behaves like an *npn* transistor, again with large beta. It has only a single base-emitter drop, but (like the Darlington) it also cannot saturate to less than a diode drop. A resistor from base to emitter of  $Q_2$  is advisable, for the same reasons as with the Darlington (leakage current; speed; stability of  $V_{BE}$ ). This connection is common in push-pull power output stages in which the designer may wish to use one polarity of high-current output transistor only. However, even when used as complementary polarity pairs, it is generally to be preferred over the Darlington for amplifiers and other linear applications; that is because it has the advantage of a single  $V_{BE}$  drop (versus two), and that voltage drop is stabilized by the base-emitter resistor of the output transistor. For example, if  $R_B$  is chosen such that its current (with a nominal  $V_{BE}$  drop across it) is 25% of the output transistor's base current at peak output, then the driver transistor sees a collector current that ranges over only a factor of 5; so its  $V_{BE}$  (which is the Sziklai's  $V_{BE}$ ) varies only 40 mV ( $V_T \ln 5$ ) over the full output current swing. The Sziklai configuration is discussed in more detail in Chapter 2x (see §2x.10); and you'll find nice examples of circuits that rely on the

<sup>56</sup> George C. Sziklai, "Symmetrical properties of transistors and their applications," Proc. IRE 41, 717–24 (1953), and US patents 2,762,870 and 2,791,644. His new complementary configuration is buried as Figure 8, where he remarks that "The complementary symmetry of transistors finds an interesting application when it is applied to the cascading of push-pull amplifier stages." The circuit evidently was devised by Sziklai, Lohman, and Herzog, for a transistorized TV demonstration at RCA; the common wisdom was that transistors weren't good enough for the task. In early ICs, where only poor *pnp* transistors were available, an additional *npn* was added, in Sziklai fashion, to boost the current capability of the *pnp*; the combination was called a "composite lateral *pnp*."

Sziklai's unique properties in that chapter's section "BJT amplifier distortion: a SPICE exploration."

Figure 2.78 shows a nice example of a push-pull Sziklai output stage. This has an important advantage compared with the Darlington alternative, namely that the biasing of the  $Q_3Q_5$  pair into class-AB conduction (to minimize crossover distortion) has just two base-emitter drops, rather than four; and, more importantly,  $Q_3$  and  $Q_5$  are running cool compared with the output transistors ( $Q_4$  and  $Q_6$ ), so they can be relied upon to have a stable base-emitter drop. This allows higher quiescent currents than with the conventional Darlington, where you have to leave a larger safety margin; bottom line, lower distortion.<sup>57</sup>

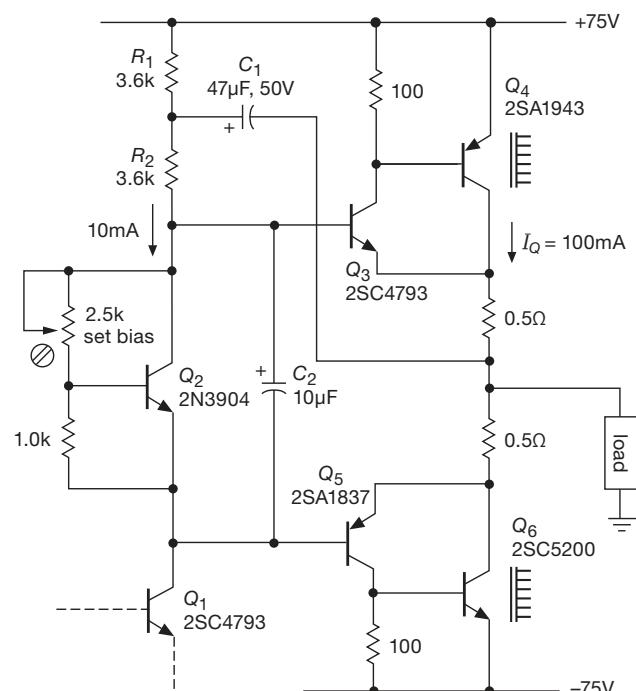
In this circuit  $Q_2$  functions as an "adjustable  $V_{BE}$  multiplier" for biasing, here settable from 1 to 3.5  $V_{BE}$ 's; it is bypassed at signal frequencies. Another circuit trick is the "bootstrapping" of  $Q_1$ 's collector resistor by  $C_1$  (see §2.4.3), raising its effective resistance at signal frequencies and increasing the amplifier's loop gain to produce lower distortion.

### B. Superbeta transistor

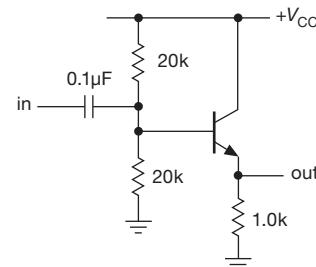
The Darlington connection and its near relatives should not be confused with the so-called superbeta transistor, a device with very high current gain achieved through the manufacturing process. A typical superbeta transistor is the 2N5962, with a guaranteed minimum current gain of 450 at collector currents from  $10\ \mu\text{A}$  to  $10\ \text{mA}$  (see, for example, Table 8.1a on page 501). Superbeta matched pairs are available for use in low-level amplifiers that require matched characteristics, for example the differential amplifier of §2.3.8. Legendary examples are the LM394 and MAT-01 series; these provide high-gain *npn* transistor pairs whose  $V_{BE}$ 's are matched to a fraction of a millivolt (as little as  $50\ \mu\text{V}$  in the best versions) and whose betas are matched to about 1%. The MAT-03 is a *pnp* matched pair (see Table 8.1b on page 502). Some commercial op-amps use superbeta differential input stages to achieve input (i.e., base bias) currents as low as  $50\ \text{picoamps}$  this way; examples are the LT1008 and LT1012.

### 2.4.3 Bootstrapping

When biasing an emitter follower, for instance, you choose the base voltage-divider resistors so that the divider presents a stiff voltage source to the base, i.e., their par-



**Figure 2.78.** Push-pull power stage with Sziklai-pair output transistors, capable of output swings to  $\pm 70\text{ V}$  and output currents of  $\pm 2\text{ A}$  peak.



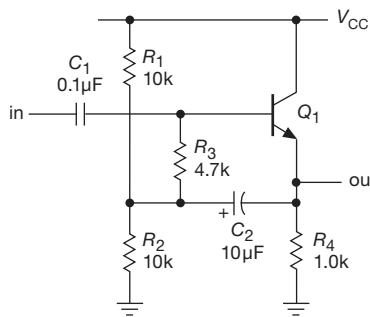
**Figure 2.79.** Bias network lowers input impedance.

allel impedance is much less than the impedance looking into the base. For this reason the resulting circuit has an input impedance dominated by the voltage divider – the driving signal sees a much lower impedance than would otherwise be necessary. Figure 2.79 shows an example. The input resistance of about  $9.1\text{k}$  is mostly due to the voltage-divider impedance of  $10\text{k}$ . It is always desirable to keep input impedances high, and anyway it's a shame to load the input with the divider, which, after all, is only there to bias the transistor.

"Bootstrapping" is the colorful name given to a technique that circumvents this problem (Figure 2.80). The

<sup>57</sup> To handle higher power, a common practice is to connect in parallel several identical  $Q_3Q_4$  stages (each with its  $0.5\Omega$  emitter resistor), and similarly for  $Q_5Q_6$ . See §2.4.4.

transistor is biased by the divider  $R_1R_2$  through series resistor  $R_3$ . Capacitor  $C_2$  is chosen to have low impedance at signal frequencies compared with the bias resistors. As always, bias is stable if the dc impedance seen from the base (in this case 9.7k) is much less than the dc impedance looking into the base (in this case approximately 100k). But now the signal-frequency input impedance is no longer the same as the dc impedance. Look at it this way: an input wiggle  $v_{in}$  results in an emitter wiggle  $v_E \approx v_{in}$ . So the change in current through bias resistor  $R_3$  is  $i = (v_{in} - v_E)/R_3 \approx 0$ , i.e.,  $Z_{in}$  (of the bias string) =  $v_{in}/i_{in} \approx \infty$ . We've made the loading (shunt) impedance of the bias network very large *at signal frequencies*.



**Figure 2.80.** Raising the input impedance of an emitter follower at signal frequencies by bootstrapping the base bias divider.

Another way of seeing this is to notice that  $R_3$  always has the same voltage across it at signal frequencies (since both ends of the resistor have the same voltage changes), i.e., it's a current source. But a current source has infinite impedance. In reality the effective impedance is less than infinity because the gain of a follower is slightly less than unity. That is so because the base-emitter drop depends on collector current, which changes with the signal level. You could have predicted the same result from the voltage-dividing effect of the impedance looking into the emitter [ $r_e = 25/I_C$  (mA) ohms] combined with the emitter resistor. If the follower has voltage gain  $A$  (slightly less than unity), the effective value of  $R_3$  at signal frequencies is

$$R_3/(1-A).$$

The voltage gain of a follower can be written  $A=R_L/(R_L+r_e)$ , where  $R_L$  is the total load seen at the emitter (here  $R_1 \parallel R_2 \parallel R_4$ ), so the effective value of bias resistor  $R_3$  at signal frequencies can be written as  $R_3 \rightarrow R_3(1+R_L/r_e)$ . In practice the value of  $R_3$  is effectively increased by a hundred or so, and the input impedance is then dominated by the transistor's base impedance. The emitter-degenerated amplifier can be bootstrapped in the same way,

since the signal on the emitter follows the base. The bias divider circuit is driven by the low-impedance emitter output at signal frequencies, which is what isolates the input signal from this usual task, and makes possible the beneficial increase of input impedance.

### A. Bootstrapping collector load resistors

The bootstrap principle can be used to increase the effective value of a transistor's collector load resistor, if that stage drives a follower. That can increase the voltage gain of the stage substantially – recall that  $G_V = -g_m R_C$ , with  $g_m = 1/(R_E + r_e)$ . This technique is used in Figure 2.78, where we bootstrapped  $Q_1$ 's collector load resistor ( $R_2$ ), forming an approximate current-source load. This serves two useful functions: (a) it raises the voltage gain of  $Q_1$ , and (b) it provides base drive current to  $Q_3Q_4$  that does not drop off toward the top of the swing (as would a resistive load, just when you need it most).

### 2.4.4 Current sharing in paralleled BJTs

It's not unusual in power electronics design to find that the power transistor you've chosen is not able to handle the required power dissipation, and needs to share the job with additional transistors. This is a fine idea, but you need a way to ensure that each transistor handles an equal portion of the power dissipation. In §9.13.5 we illustrate the use of transistors *in series*. This can simplify the problem, because we know they'll all be running at the same current. But it's often more attractive to divide up the current by connecting the transistors in parallel, as in Figure 2.81A.

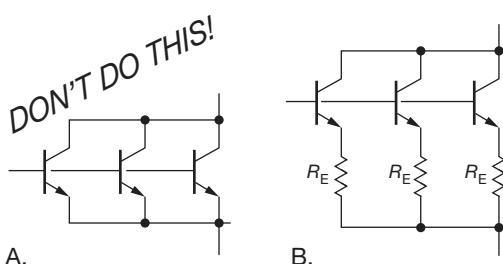
There are two problems with this approach. First, we know the bipolar transistor is a transconductance device, with its collector current determined in a precise way by its base-to-emitter voltage  $V_{BE}$ , as given by the Ebers–Moll equations 2.8 and 2.9. As we saw in §2.3.2, the temperature coefficient of  $V_{BE}$  (at constant collector current) is about  $-2.1\text{ mV}^{\circ}\text{C}$ ; or, equivalently,  $I_C$  increases with temperature for a fixed  $V_{BE}$ .<sup>58</sup> This is unfortunate, because if the junction of one of the transistors becomes hotter than the rest, it takes more of the total current, thereby heating up even more. It's in danger of the dreaded thermal runaway.

The second problem is that transistors of the same part number are not identical. They come off the shelf with

<sup>58</sup> This result comes directly from  $\partial I_C / \partial T = -g_m \partial V_{BE} / \partial T$ , which, after substituting  $g_m = I_C / V_T$  tells us that the fractional change of collector current is just  $(\partial I_C / \partial T) / I_C = -(\partial V_{BE} / \partial T) / V_T$ . So the collector current increases fractionally by about  $2.1\text{mV}/25\text{mV}$  (or 8.4%) per  $^{\circ}\text{C}$  – a rather large amount!

differing values of  $V_{BE}$  for a given  $I_C$ . This is true even for parts made at the same time on the same fabrication line, and from the same silicon wafer. To see how large a variation you are likely to get, we measured 100 adjacent ZTX851 transistors on a reel, with an observed spread of about 17 mV, shown in Figure 8.44. This really represents a “best case,” because you cannot be certain that a batch of incoming transistors derive from a single lot, much less a single wafer. When you first build something, the  $V_{BE}$ ’s of “identical” transistors may be within 20–50 mV of each other, but that matching is lost when one of them has to be replaced someday. It’s always safer to assume a possible 100 mV or so spread of base-emitter voltages. Recalling that  $\Delta V_{BE}=60$  mV corresponds to a factor of ten current ratio, it’s clear that you cannot get away with a direct parallel connection like Figure 2.81A.

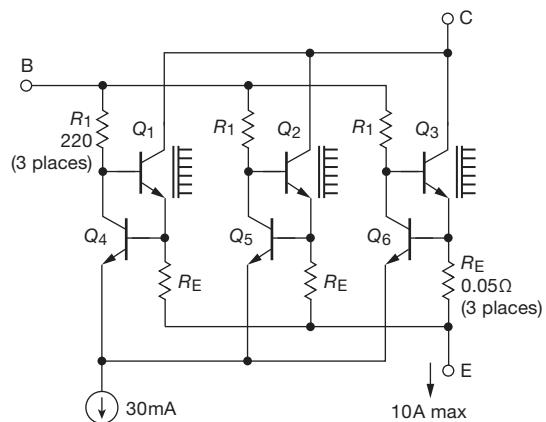
The usual solution to this problem is the use of small resistors in the emitters, as shown in Figure 2.81B. These are called *emitter-ballasting* resistors, and their value is chosen to drop at least a few tenths of a volt at the higher end of the anticipated operating current range. That voltage drop must be adequate to swamp the  $V_{BE}$  spread of the individual transistors, and is ordinarily chosen somewhere in the range of 300–500 mV.



**Figure 2.81.** To equalize the currents of parallel transistors, use emitter ballasting resistors  $R_E$ , as in circuit B.

At high currents the resistors may suffer from an inconveniently-high power dissipation, so you may want to use the current-sharing trick shown in Figure 2.82. Here the current-sensing transistors  $Q_4$ – $Q_6$  adjust the base drive to the “paralleled” power transistors  $Q_1$ – $Q_3$  to maintain equal emitter currents (you can think of  $Q_4$ – $Q_6$  as a high-gain differential amplifier with three inputs). This “active ballast” technique works well with power Darlington BJTs, and it works particularly well with MOSFETs (see Figure 3.117), thanks to their negligible input (gate) current, thus making

MOSFETs a good choice for circuits with lots of power dissipation.<sup>59</sup>



**Figure 2.82.** Active ballasting of parallel transistors  $Q_1$ – $Q_3$  via feedback from current sensing transistors  $Q_4$ – $Q_6$  lets you configure a parallel power transistors with very low drops across the emitter resistors.

## 2.4.5 Capacitance and Miller effect

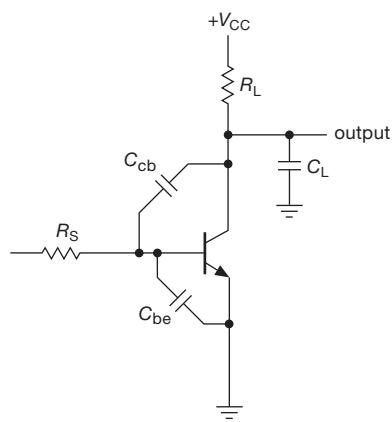
In our discussion so far we have used what amounts to a dc, or low-frequency, model of the transistor. Our simple current amplifier model and the more sophisticated Ebers–Moll transconductance model both deal with voltages, currents, and resistances seen at the various terminals. With these models alone we have managed to go quite far, and in fact these simple models contain nearly everything you will ever need to know to design transistor circuits. However, one important aspect that has serious impact on high-speed and high-frequency circuits has been neglected: the existence of capacitance in the external circuit and in the transistor junctions themselves. Indeed, at high frequencies the effects of capacitance often dominate circuit behavior; at 100 MHz a typical junction capacitance of 5 pF has an impedance of just 320  $\Omega$ !

In this brief subsection we introduce the problem, illustrate some of its circuit incarnations, and suggest some methods of circumventing its effects. It would be a mistake to leave this chapter without realizing the nature of this problem. In the course of this brief discussion we will encounter the infamous *Miller effect*, and the use of configurations such as the cascode to overcome it.

<sup>59</sup> Another nice feature of MOSFETs is their lack of second breakdown, thus a wider safe-operating area; see §3.6.4C.

### A. Junction and circuit capacitance

Capacitance limits the speed at which the voltages within a circuit can swing (“slew rate”), owing to finite driving impedance or current. When a capacitance is driven by a finite source resistance, you see  $RC$  exponential charging behavior, whereas a capacitance driven by a current source leads to slew-rate-limited waveforms (ramps). As general guidance, reducing the source impedances and load capacitances and increasing the drive currents within a circuit will speed things up. However, there are some subtleties connected with feedback capacitance and input capacitance. Let’s take a brief look.



**Figure 2.83.** Junction and load capacitances in a transistor amplifier.

The circuit in Figure 2.83 illustrates most of the problems of junction capacitance. The output capacitance forms a time constant with the output resistance  $R_L$  ( $R_L$  includes both the collector and load resistances, and  $C_L$  includes both junction and load capacitances), giving a rolloff starting at some frequency  $f = 1/2\pi R_L C_L$ .

The same is true for the input capacitance,  $C_{be}$ , in combination with the source impedance  $R_S$ . Of greater significance, at high frequencies the input capacitance robs base current, effectively decreasing the transistor’s beta. In fact, transistor datasheets specify a cutoff frequency,  $f_T$ , at which the beta has decreased to unity – not much of an amplifier anymore! We discuss this further in Chapter 2x.

### B. Miller effect

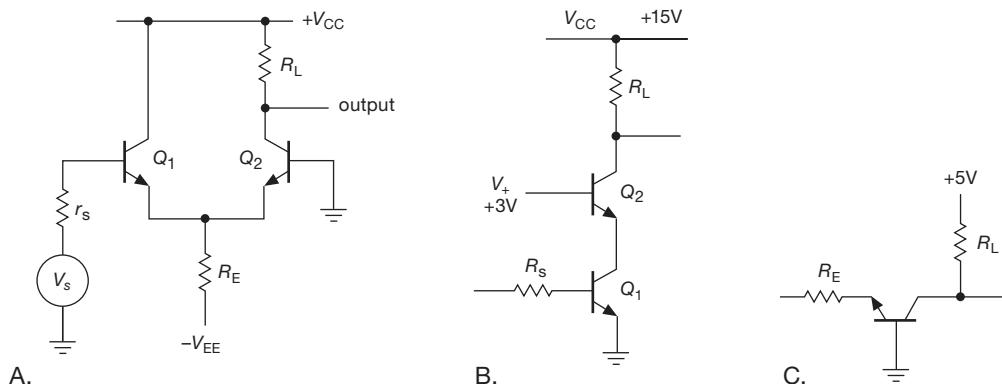
The feedback impedance  $C_{cb}$  is another matter. The amplifier has some overall voltage gain  $G_V$ , so a small voltage wiggle at the input results in a wiggle  $G_V$  times larger (and inverted) at the collector. This means that the signal source sees a current through  $C_{cb}$  that is  $G_V + 1$  times as large as if  $C_{cb}$  were connected from base to ground; i.e., for the pur-

pose of input rolloff frequency calculations, the feedback capacitance behaves like a capacitor of value  $C_{cb}(G_V + 1)$  from input to ground. This effective increase of  $C_{cb}$  is known as the Miller effect. It often dominates the rolloff characteristics of amplifiers, because a typical feedback capacitance of 4 pF can look like several hundred picofarads to ground.

There are several methods available for beating the Miller effect: (a) you can decrease the source impedance driving a grounded-emitter stage by using an emitter follower. Figure 2.84 shows three other possibilities; (b) the differential amplifier circuit with no collector resistor in  $Q_1$  (Figure 2.84A) has no Miller effect; you can think of it as an emitter follower driving a grounded-base amplifier (see below); (c) the famous cascode configuration (Figure 2.84B) elegantly defeats the Miller effect. Here  $Q_1$  is a grounded-emitter amplifier with  $R_L$  as its collector resistor;  $Q_2$  is interposed in the collector path to prevent  $Q_1$ ’s collector from swinging (thereby eliminating the Miller effect) while passing the collector current through to the load resistor unchanged. The input labeled  $V_+$  is a fixed-bias voltage, usually set a few volts above  $Q_1$ ’s emitter voltage to pin  $Q_1$ ’s collector and keep it in the active region. This circuit fragment is incomplete, because biasing is not shown; you could either include a bypassed emitter resistor and base divider for biasing  $Q_1$  (as we did earlier in the chapter) or include it within an overall loop with feedback at dc.  $V_+$  might be provided from a divider or zener, with bypassing to keep it stiff at signal frequencies. (d) Finally, the grounded-base amplifier can be used by itself, as shown in Figure 2.84C. It has no Miller effect because the base is driven by zero source impedance (ground), and the amplifier is noninverting from input to output.

**Exercise 2.19.** Explain in detail why there is no Miller effect in either transistor in the preceding differential amplifier and cascode circuits.

Capacitive effects can be somewhat more complicated than this brief introduction might indicate. In particular: (a) the rolloffs that are due to feedback and output capacitances are not entirely independent; in the terminology of the trade there is *pole splitting*; (b) the transistor’s input capacitance still has an effect, even with a stiff input signal source. In particular, current that flows through  $C_{be}$  is not amplified by the transistor. This base current “robbing” by the input capacitance causes the transistor’s small-signal current gain  $h_{fe}$  to drop at high frequencies, eventually reaching unity at a frequency known as  $f_T$ . (c) To complicate matters, the junction capacitances depend on voltage: a dominant portion of  $C_{be}$  changes proportionally with



**Figure 2.84.** Three circuit configurations that avoid the Miller effect. A. Differential amplifier with inverting input grounded. B. Cascode connection. C. Grounded base amplifier.

operating current, so  $f_T$  is given instead.<sup>60</sup> (d) When a transistor is operated as a switch, effects associated with charge stored in the base region of a saturated transistor cause an additional loss of speed.

The Miller effect looms large in high-speed and wide-band circuits, and we'll be seeing it again and again in subsequent chapters.

#### 2.4.6 Field-effect transistors

In this chapter we have dealt exclusively with BJTs, characterized by the Ebers–Moll equation. BJTs were the original transistors, and they are widely used in analog circuit design. However, it would be a mistake to continue without a few words of explanation about the other kind of transistor, the FET, which we will take up in detail in Chapter 3.

The FET behaves in many ways like an ordinary bipolar transistor. It is a three-terminal amplifying device, available in both polarities, with a terminal (the *gate*) that controls the current flow between the other two terminals (*source* and *drain*). It has a unique property, though: the gate draws no dc current, except for leakage. This means that extremely high input impedances are possible, limited only by capacitance and leakage effects. With FETs you don't have to worry about providing substantial base current, as was necessary with the BJT circuit design of this chapter. Input currents measured in *picoamperes* are commonplace. Yet the FET is a rugged and capable device, with voltage and current ratings comparable to those of bipolar transistors.

Most of the available devices fabricated with BJTs

(matched pairs, differential and operational amplifiers, comparators, high-current switches and amplifiers, and RF amplifiers) are also available with FET construction, often with superior performance. Furthermore, digital logic, microprocessors, memory, and all manner of complex and wonderful large-scale digital chips are built almost exclusively with FETs. Finally, the area of micropower design is dominated by FET circuits. It is not exaggeration to say that, demographically, almost all transistors are FETs.<sup>61</sup>

FETs are so important in electronic design that we devote the next chapter to them before treating operational amplifiers and feedback in Chapter 4. We urge the reader to be patient with us as we lay the groundwork in these first three difficult chapters; that patience will be rewarded many times over in the succeeding chapters, as we explore the enjoyable topics of circuit design with operational amplifiers and digital integrated circuits.

#### 2.5 Negative feedback

We've hinted earlier in the chapter that feedback offers a cure to some vexing problems: biasing the grounded-emitter amplifier (§2.3.4 and 2.3.5), biasing the differential amplifier with current-mirror active load (§2.3.8C), and minimizing crossover distortion in push–pull followers (§2.4.1A). It's even better than that – *negative feedback*

<sup>60</sup> See values of  $f_T$  versus collector current for 25 transistors, plotted and tabulated in Chapter 2x's section titled “BJT Bandwidth and  $f_T$ .”

<sup>61</sup> Lest this outpouring of enthusiasm leave the wrong impression, we hasten to point out that BJTs are alive and well, largely because they are unbeatable when it comes to characteristics like accuracy and noise (the subjects of Chapters 5 and 8). They excel also in transconductance (i.e., gain). Those muscular power FETs suffer from rather high input capacitance; and, as discrete parts, you cannot get small-signal MOSFETs, only *power* MOSFETs.

is a wonderful technique that can cure all manner of ills: distortion and nonlinearities, frequency dependence of amplifier gain, departure from ideal performance of voltage sources, current sources, or pretty much anything else.

We'll be enjoying the benefits of negative feedback fully in Chapter 4, where we introduce the universal analog component called an *operational amplifier* ("op-amp"), a creature that thrives on negative feedback. But this is a good place to introduce feedback, both because it is widely used in discrete transistor circuits and also because it is present already in our common emitter amplifier, whose improved linearity (compared with that of the grounded-emitter amplifier) is due to negative feedback.

### 2.5.1 Introduction to feedback

Feedback has become such a well-known concept that the word has entered the general vocabulary. In control systems, feedback consists of comparing the actual output of the system with the desired output and making a correction accordingly. The "system" can be almost anything: for instance, the process of driving a car down the road, in which the output (the position and velocity of the car) is sensed by the driver, who compares it with expectations and makes corrections to the input (steering wheel, throttle, brake). In amplifier circuits the output should be a multiple of the input, so in a feedback amplifier the input is compared with an attenuated version of the output.

As used in amplifiers, negative feedback is implemented simply by coupling the output back in such a way as to cancel some of the input. You might think that this would only have the effect of reducing the amplifier's gain and would be a pretty stupid thing to do. Harold S. Black, who attempted to patent negative feedback in 1928, was greeted with the same response. In his words, "Our patent application was treated in the same manner as one for a perpetual-motion machine."<sup>62</sup> True, it does lower the gain, but in exchange it also improves other characteristics, most notably freedom from distortion and nonlinearity, flatness of response (or conformity to some desired frequency response), and predictability. In fact, as more negative feedback is used, the resultant amplifier characteristics become less dependent on the characteristics of the open-loop (no-feedback) amplifier and finally depend on the properties only of the feedback network itself. Operational amplifiers

(the very high-gain differential amplifier building blocks of Chapter 4) are typically used in this high-loop-gain limit, with *open-loop* voltage gain (no feedback) of a million or so.

A feedback network can be frequency dependent, to produce an equalization amplifier (with specific gain-versus-frequency characteristics), or it can be amplitude dependent, producing a nonlinear amplifier (an example is a logarithmic amplifier, built with feedback that exploits the logarithmic  $V_{BE}$  versus  $I_C$  of a diode or transistor). It can be arranged to produce a current source (near-infinite output impedance) or a voltage source (near-zero output impedance), and it can be connected to generate very high or very low input impedance. Speaking in general terms, the property that is sampled to produce feedback is the property that is improved. Thus, if you feed back a signal proportional to the output current, you will generate a good current source.<sup>63</sup>

Let's look at how feedback works, and how it affects what an amplifier does. We will find simple expressions for the input impedance, output impedance, and gain of an amplifier with negative feedback.

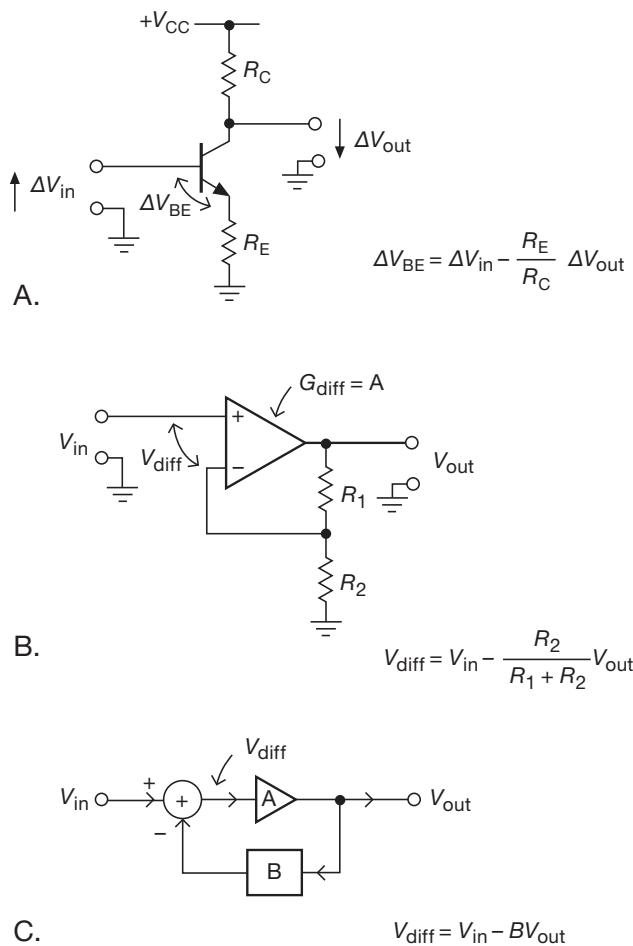
### 2.5.2 Gain equation

Look at Figure 2.85. To get started we've drawn the familiar common-emitter amplifier with emitter degeneration. Thinking of the transistor in the Ebers–Moll sense, the small-signal voltage from base to emitter ( $\Delta V_{BE}$ ) programs the collector current. But  $\Delta V_{BE}$  is less than the input voltage  $V_{in}$ , because of the drop across  $R_E$ . If the output is unloaded, it's easy to get the equation in the figure. In other words, the common-emitter amplifier with emitter degeneration is a grounded-emitter amplifier with negative feedback, as we hinted earlier.

This circuit has some subtleties, which we'd like to sidestep for now by looking instead at the more straightforward configuration shown in Figure 2.85B. Here we've drawn a differential amplifier (with differential gain  $A$ ), with a fraction of its output signal subtracted from the circuit input  $v_{in}$ . That fraction, of course, is given simply by

<sup>62</sup> See the fascinating article in *IEEE Spectrum*, December 1977. His patent for negative feedback (No. 2,102,671, modestly titled "Wave translation system") was granted in 1937, nine years after his initial filing.

<sup>63</sup> Feedback can also be *positive*; that's how you make an oscillator, for instance. As much fun as that may sound, it simply isn't as important as negative feedback. More often it's a nuisance, because a negative-feedback circuit may have large enough phase shifts at some high frequency to produce positive feedback and oscillations. It is surprisingly easy to have this happen, and the prevention of unwanted oscillations is the object of what is called *compensation*, a subject we treat briefly at the end of Chapter 4.



**Figure 2.85.** Negative feedback subtracts a fraction of the output from the input: A. Common-emitter amplifier. B. Differential amplifier configured as a noninverting voltage amplifier. C. Conventional block diagram.

the voltage divider equation, as shown. This is a very common configuration, widely used with op-amps (Chapter 4), and known simply as a “noninverting amplifier.”

When talking about negative feedback, it’s conventional to draw a diagram like Figure 2.85C, in which the feedback fraction is simply labeled  $B$ . This is useful because it allows more generality than a voltage divider (feedback can include frequency-dependent components like capacitors, and nonlinear components like diodes), and it keeps the equations simple. For a voltage divider, of course,  $B$  would simply be equal to  $R_2/(R_1 + R_2)$ .

Let’s figure out the gain. The amplifier has open-loop voltage gain  $A$ , and the feedback network subtracts a fraction  $B$  of the output voltage from the input. (Later we will generalize things so that inputs and outputs can be currents

or voltages.) The input to the gain block is then  $V_{in} - BV_{out}$ . But the output is just the input times  $A$ :

$$A(V_{in} - BV_{out}) = V_{out}.$$

In other words,

$$V_{out} = \frac{A}{1+AB} V_{in},$$

and so the closed-loop voltage gain,  $V_{out}/V_{in}$ , is just

$$G = \frac{A}{1+AB}. \quad (2.16)$$

Some terminology: the standard designations for these quantities are as follows:  $G$  = closed-loop gain,  $A$  = open-loop gain,  $AB$  = loop gain,  $1+AB$  = return difference, or desensitivity. The feedback network is sometimes called the beta network (no relation to transistor beta,  $h_{fe}$ ).<sup>64</sup>

### 2.5.3 Effects of feedback on amplifier circuits

Let’s look at the important effects of feedback. The most significant are predictability of gain (and reduction of distortion), changed input impedance, and changed output impedance.

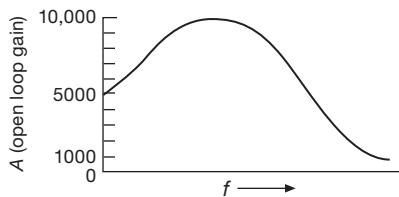
#### A. Predictability of gain

The voltage gain is  $G = A/(1+AB)$ . In the limit of infinite<sup>65</sup> open-loop gain  $A$ ,  $G = 1/B$ . For finite gain  $A$ , feedback acts to reduce the effects of variations of  $A$  (with frequency, temperature, amplitude, etc.). For instance, suppose  $A$  depends on frequency as in Figure 2.86. This will surely satisfy anyone’s definition of a poor amplifier (the gain varies over a factor of 10 with frequency). Now imagine we introduce feedback, with  $B = 0.1$  (a simple voltage divider will do). The closed-loop voltage gain now varies from  $1000/[1+(1000 \times 0.1)]$ , or 9.90, to  $10,000/[1+(10,000 \times 0.1)]$ , or 9.99, a variation of just 1% over the same range of frequency! To put it in audio terms, the original amplifier is flat to  $\pm 10$  dB, whereas the feedback amplifier is flat to  $\pm 0.04$  dB. We can now recover the original gain of 1000 with nearly this linearity simply by cascading three such stages.

It was for just this reason (namely, the need for extremely flat-response telephone repeater amplifiers) that

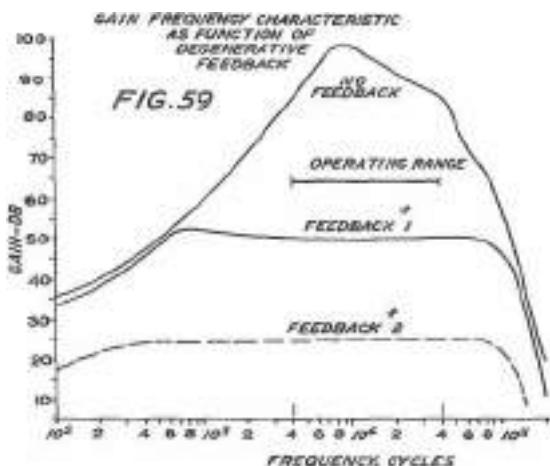
<sup>64</sup> We’ll see later that amplifiers used with feedback commonly have significant lagging phase shifts from input to output. So the open-loop voltage gain  $A$  should properly be represented as a complex number. We’ll treat this in §2.5.4; for now we’ll adopt the simplification that the amplifier’s output voltage is proportional to its input voltage.

<sup>65</sup> Which is not a bad approximation for an op-amp, whose typical open-loop gain is in the neighborhood of  $A_{OL} \approx 10^6$ .



**Figure 2.86.** Amplifier with open-loop gain  $A$  that varies widely with frequency  $f$ .

negative feedback in electronics was invented. As the inventor, Harold Black, described it in his first open publication on the invention [Elec. Eng., 53, 114, (1934)], “by building an amplifier whose gain is made deliberately, say 40 decibels higher than necessary (10,000-fold excess on energy basis) and then feeding the output back to the input in such a way as to throw away the excess gain, it has been found possible to effect extraordinary improvement in constancy of amplification and freedom from nonlinearity.” Black’s patent is spectacular, with dozens of elegant figures; we reproduce one of them here (Figure 2.87), which makes the point eloquently.



**Figure 2.87.** Harold Black explains it in his historic 1937 patent, with the unassuming title “Wave translation system.”

It is easy to show, by taking the partial derivative of  $G$  with respect to  $A$  (i.e.,  $\partial G / \partial A$ ), that relative variations in the open-loop gain are reduced by the desensitivity:

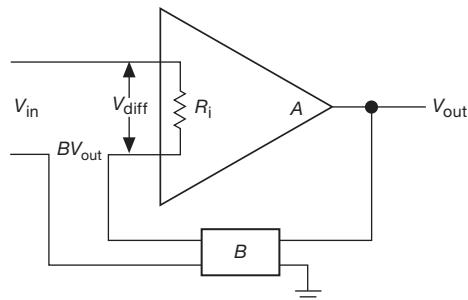
$$\frac{\Delta G}{G} = \frac{1}{1+AB} \frac{\Delta A}{A}. \quad (2.17)$$

Thus, for good performance the loop gain  $AB$  should be much larger than 1. That’s equivalent to saying that the open-loop gain should be much larger than the closed-loop gain.

A very important consequence of this is that nonlinearities, which are simply gain variations that depend on signal level, are reduced in exactly the same way.

### B. Input impedance

Feedback can be arranged to subtract a voltage or a current from the input (these are sometimes called *series feedback* and *shunt feedback*, respectively). The noninverting amplifier configuration we’ve been considering, for instance, subtracts a sample of the output voltage from the differential *voltage* appearing at the input, whereas the feedback scheme in Figure 2.89B subtracts a *current* from the input. The effects on input impedance are opposite in the two cases: voltage feedback multiplies the open-loop input impedance by  $1+AB$ , whereas current feedback reduces it by the same factor. In the limit of infinite loop gain the input impedance (at the amplifier’s input terminal) goes to infinity or zero, respectively. This is easy to understand, since voltage feedback tends to subtract signal from the input, resulting in a smaller change (by the factor  $AB$ ) across the amplifier’s input resistance; it’s a form of bootstrapping. Current feedback reduces the input signal by bucking it with an equal current.



**Figure 2.88.** Series-feedback input impedance.

### Series (voltage) feedback

Let’s see explicitly how the effective input impedance is changed by feedback. We illustrate the case of voltage feedback only, since the derivations are similar for the two cases. We begin with a differential amplifier model with (finite) input resistance as shown in Figure 2.88. An input  $V_{in}$  is reduced by  $BV_{out}$ , putting a voltage  $V_{diff} = V_{in} - BV_{out}$  across the inputs of the amplifier. The input current is therefore

$$I_{in} = \frac{V_{in} - BV_{out}}{R_i} = \frac{V_{in} \left( 1 - B \frac{A}{1+AB} \right)}{R_i} = \frac{V_{in}}{(1+AB)R_i},$$

giving an effective input impedance

$$Z_{in} = V_{in}/I_{in} = (1 + AB)R_i.$$

In other words, the input impedance is boosted by a factor of the loop gain plus one. If you were to use the circuit of Figure 2.85B to close the feedback loop around a differential amplifier whose native input impedance is  $100\text{ k}\Omega$  and whose differential gain is  $10^4$ , choosing the resistor ratio (99:1) for a target gain of 100 (in the limit of infinite amplifier gain), the input impedance seen by the signal source would be approximately  $10\text{ M}\Omega$ , and the closed-loop gain would be 99.<sup>66</sup>

### Shunt (current) feedback

Look at Figure 2.89A. The impedance seen looking into the input of a voltage amplifier with current feedback is reduced by the feedback current, which opposes voltage changes at the input.<sup>67</sup> By considering the current change produced by a voltage change at the input, you find that the input signal sees a parallel combination of (a) the amplifier's native input impedance  $R_i$  and (b) the feedback resistor  $R_f$  divided by  $1 + A$ . That is,

$$Z_{in} = R_i \parallel \frac{R_f}{1 + A}$$

(see if you can prove this). In cases of very high loop gain (e.g., an op-amp), the input impedance is reduced to a fraction of an ohm, which might seem bad. But in fact this configuration is used to convert an input current into an output voltage (a “transresistance amplifier”), for which a low input impedance is a good characteristic. We'll see examples in Chapters 4 and 4x.

By the addition of an input resistor (Figure 2.89B) the circuit becomes an “inverting amplifier,” with input resistance as shown. You can think of this (particularly in the high-loop-gain limit) as a resistor feeding a current-to-voltage amplifier. In that limit  $R_{in}$  approximately equals  $R_1$  (and the closed-loop gain approximately equals  $-R_2/R_1$ ).

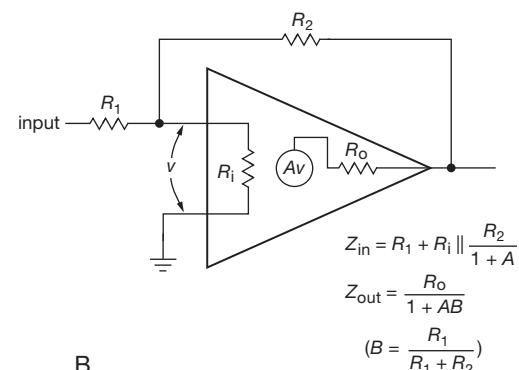
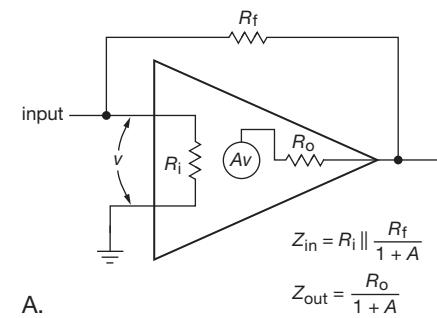
It is a straightforward exercise to derive an expression for the closed-loop voltage gain of the inverting amplifier with finite loop gain. The answer is

$$G = -A(1 - B)/(1 + AB)$$

where  $B$  is defined as before,  $B = R_1/(R_1 + R_2)$ . In the limit of large open-loop gain  $A$ ,  $G = 1 - 1/B$  (i.e.,  $G = -R_2/R_1$ ).

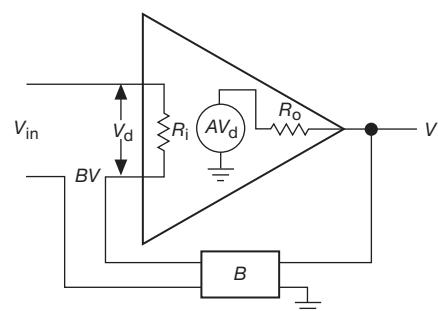
<sup>66</sup> Of course, knowing that the open-loop gain is approximately  $10^4$ , you might bump the resistor ratio up to 100:1 to compensate. With an op-amp there's no need: with a typical open-loop gain of  $\sim 10^6$ , the closed-loop gain would be  $G_{CL} = 99.99$ .

<sup>67</sup> As in the circuit of Figure 2.53 in §2.3.5C.



**Figure 2.89.** Input and output impedances for (A) transresistance amplifier and (B) inverting amplifier.

**Exercise 2.20.** Derive the foregoing expressions for input impedance and gain of the inverting amplifier.



**Figure 2.90.** Output impedance.

### C. Output impedance

Again, feedback can extract a sample of the output voltage or the output current. In the first case the open-loop output impedance will be reduced by the factor  $1 + AB$ , whereas in the second case it will be increased by the same factor. We illustrate this effect for the case of voltage sampling.

We begin with the model shown in Figure 2.90. This time we have shown the output impedance explicitly. The calculation is simplified by a trick: short the input and apply a voltage  $V$  to the output; by calculating the output current  $I$ , we get the output impedance  $R'_o = V/I$ . Voltage  $V$  at the output puts a voltage  $-BV$  across the amplifier's input, producing a voltage  $-ABV$  in the amplifier's internal generator. The output current is therefore

$$I = \frac{V - (-ABV)}{R_o} = \frac{V(1+AB)}{R_o}$$

giving an effective output impedance<sup>68</sup> of

$$Z_{\text{out}} = V/I = R_o/(1+AB).$$

#### D. Sensing output current

Feedback can be connected instead to sample the output current. Then the expression for output impedance becomes

$$Z_{\text{out}} = R_o(1+AB).$$

In fact, it is possible to have multiple feedback paths, sampling both voltage and current. In the general case the output impedance is given by Blackman's impedance relation<sup>69</sup>

$$Z_{\text{out}} = R_o \frac{1 + (AB)_{\text{SC}}}{1 + (AB)_{\text{OC}}},$$

where  $(AB)_{\text{SC}}$  is the loop gain with the output shorted to ground and  $(AB)_{\text{OC}}$  is the loop gain with no load attached. Thus feedback can be used to generate a desired output impedance. This equation reduces to the previous results for the usual situation in which feedback is derived from either the output voltage or the output current. See additional discussion in Chapter 2x.

#### 2.5.4 Two important details

Feedback is a rich subject, which we've simplified shamelessly in this brief introduction. Here are two details that should not be overlooked, however, even at this somewhat superficial level of understanding.

<sup>68</sup> If the open-loop gain  $A$  is real (i.e. no phase shift), then the output impedance  $Z_{\text{out}}$  will be real (i.e., resistive:  $R_{\text{out}}$ ). As we'll see in Chapter 4, however,  $A$  can be (and often is) complex, representing a lagging phase shift. For op-amps the phase shift is  $90^\circ$  over most of the amplifier's bandwidth. The result is an *inductive* closed-loop output impedance. See for example Figure 4.53 in Chapter 4.

<sup>69</sup> R. B. Blackman, "Effect of feedback on impedance," *Bell. Sys. Tech. J.* 22, 269 (1943).

#### A. Loading by the feedback network

In feedback computations, you usually assume that the beta network doesn't load the amplifier's output. If it does, that must be taken into account in computing the open-loop gain. Likewise, if the connection of the beta network at the amplifier's input affects the open-loop gain (feedback removed, but network still connected), you must use the modified open-loop gain. Finally, the preceding expressions assume that the beta network is unidirectional, i.e., it does not couple any signal from the input to the output.

#### B. Phase shifts, stability, and "compensation"

The open-loop amplifier gain  $A$  is central in the expressions we've found for closed-loop gain and the corresponding input and output impedances. By default one might reasonably assume that  $A$  is a real number – that is, that the output is in phase with the input. In real life things are more complex,<sup>70</sup> because of the effects of circuit capacitances (and Miller effect, §2.4.5), and also the limited bandwidth ( $f_T$ ) of the active components themselves. The result is that the open-loop amplifier will exhibit lagging phase shifts that increase with frequency. This has several consequences for the closed-loop amplifier.

#### Stability

If the open-loop amplifier's lagging phase shift reaches  $180^\circ$ , then negative feedback becomes *positive* feedback, with the possibility of oscillation. This is not what you want! (The actual criterion for oscillation is that the phase shift be  $180^\circ$  at a frequency at which the loop gain  $AB$  equals 1.) This is a serious concern, particularly in amplifiers with plenty of gain (such as op-amps). The problem is only exacerbated if the feedback network contributes additional lagging phase shift (as it often will). The subject of *frequency compensation* in feedback amplifiers deals directly with this essential issue; you can read about it in §4.9.

#### Gain and phase shift

The expressions we found for closed-loop gain and for the input and output impedances contain the open-loop gain  $A$ . For example, the voltage amplifier with series feedback (Figures 2.85B&C, 2.88, and 2.90) has closed-loop gain  $G_{\text{CL}} = A/(1+AB)$ , where  $A = G_{\text{OL}}$ , the amplifier's open-loop gain. Let's imagine that the open-loop gain  $A$  is 100, and that we've chosen  $B = 0.1$  for a target closed-loop gain of  $G_{\text{CL}} \approx 10$ . Now, if the open-loop amplifier had no phase shifts, then  $G_{\text{CL}} \approx 9.09$ , also without phase shift. If instead the amplifier has a  $90^\circ$  lagging phase shift, then  $A$  is pure

<sup>70</sup> That's a pun, get it?

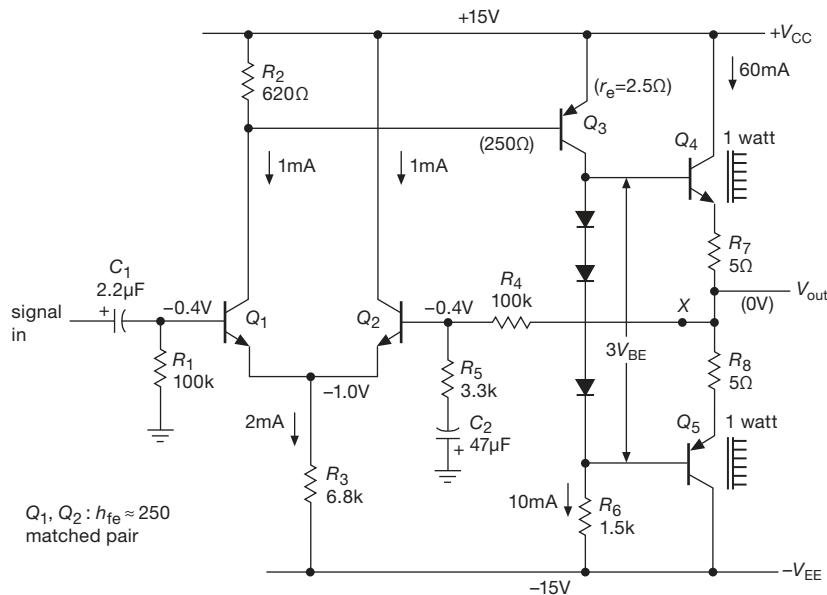


Figure 2.91. Transistor power amplifier with negative feedback.

imaginary ( $A = -100j$ ), and the closed-loop gain becomes  $G_{CL} = 9.90 - 0.99j$ . That's a magnitude  $|G_{CL}| = 9.95$ , with a lagging phase shift of approximately  $6^\circ$ . In other words, the effect of a pretty significant (halfway to oscillation!) open-loop phase shift turns out, in fact, to be favorable: the closed-loop gain is only 0.5% less than the target, compared with 9% for the case of the same amplifier without phase shift. The price you pay is some residual phase shift and, of course, an approach to instability.

As artificial as this example may seem, it in fact reflects a reality of op-amps, which usually have an  $\sim 90^\circ$  lagging phase shift over almost their entire bandwidth (typically from  $\sim 10$  Hz to 1 MHz or more). Because of their much higher open-loop gain, the amplifier with feedback exhibits very little phase shift, and an accurate gain set almost entirely by the feedback network. Much more on this in Chapter 4, and in §4.9.

**Exercise 2.21.** Verify that the above expressions for  $G_{CL}$  are correct.

## 2.5.5 Two examples of transistor amplifiers with feedback

Let's look at two transistor amplifier designs to see how the performance is affected by negative feedback. There's a bit of complexity in this analysis . . . don't be discouraged!<sup>71</sup>

<sup>71</sup> Those fearful of discouragement may wish to skip over this section in a first reading.

Figure 2.91 shows a complete transistor amplifier with negative feedback. Let's see how it goes.

### A. Circuit description

It may look complicated, but it is extremely straightforward in design and is relatively easy to analyze.  $Q_1$  and  $Q_2$  form a differential pair, with common-emitter amplifier  $Q_3$  amplifying its output.  $R_6$  is  $Q_3$ 's collector load resistor, and push-pull pair  $Q_4$  and  $Q_5$  form the output emitter follower. The output voltage is sampled by the feedback network consisting of voltage divider  $R_4$  and  $R_5$ , with  $C_2$  included to reduce the gain to unity at dc for stable biasing.  $R_3$  sets the quiescent current in the differential pair, and since overall feedback guarantees that the quiescent output voltage is at ground,  $Q_3$ 's quiescent current is easily seen to be 10 mA ( $V_{EE}$  across  $R_6$ , approximately). As we discussed earlier (§2.4.1B), the diodes bias the push-pull pair into conduction, leaving one diode drop across the series pair  $R_7$  and  $R_8$ , i.e., 60 mA quiescent current. That's class-AB operation, good for minimizing crossover distortion, at the cost of 1 watt standby dissipation in each output transistor.

From the point of view of our earlier circuits, the only unusual feature is  $Q_1$ 's quiescent collector voltage, one diode drop below  $V_{CC}$ . That is where it must sit in order to hold  $Q_3$  in conduction, and the feedback path ensures that it will. (For instance, if  $Q_1$  were to pull its collector closer to ground,  $Q_3$  would conduct heavily, raising the output

voltage, which in turn would force  $Q_2$  to conduct more heavily, reducing  $Q_1$ 's collector current and hence restoring the status quo.)  $R_2$  was chosen to give a diode drop at  $Q_1$ 's quiescent current in order to keep the collector currents in the differential pair approximately equal at the quiescent point. In this transistor circuit the input bias current is not negligible ( $4\ \mu\text{A}$ ), resulting in a  $0.4\ \text{V}$  drop across the  $100\text{k}$  input resistors. In transistor amplifier circuits like this, in which the input currents are considerably larger than in op-amps, it is particularly important to make sure that the dc resistances seen from the inputs are equal, as shown (a Darlington input stage would probably be better here).

## B. Analysis

Let's analyze this circuit in detail, determining the gain, input and output impedances, and distortion. To illustrate the utility of feedback, we will find these parameters for both the open-loop and closed-loop situations (recognizing that biasing would be hopeless in the open-loop case). To get a feeling for the linearizing effect of the feedback, the gain will be calculated at  $+10$  volts and  $-10$  volts output, as well as at the quiescent point ( $0\ \text{V}$ ).

### Open loop

**Input impedance** We cut the feedback at point  $X$  and ground the right-hand side of  $R_4$ . The input signal sees  $100\text{k}$  in parallel with the impedance looking into the base. The latter is  $h_{\text{fe}}$  times twice the intrinsic emitter resistance plus the impedance seen at  $Q_2$ 's emitter caused by the finite impedance of the feedback network at  $Q_2$ 's base. For  $h_{\text{fe}} \approx 250$ ,  $Z_{\text{in}} \approx 250 \times [(2 \times 25) + (3.3\text{k}/250)]$ ; i.e.,  $Z_{\text{in}} \approx 16\text{k}$ .

**Output impedance** Since the impedance looking back into  $Q_3$ 's collector is high, the output transistors are driven by a  $1.5\text{k}$  source ( $R_6$ ). The output impedance is about  $15\Omega$  ( $\beta \approx 100$ ) plus the  $5\ \Omega$  emitter resistance, or  $20\ \Omega$ . The intrinsic emitter resistance of  $0.4\ \Omega$  is negligible.

**Gain** The differential input stage sees a load of  $R_2$  paralleled by  $Q_3$ 's base resistance. Since  $Q_3$  is running  $10\ \text{mA}$  quiescent current, its intrinsic emitter resistance is  $2.5\ \Omega$ , giving a base impedance of about  $250\ \Omega$  (again,  $\beta \approx 100$ ). The differential pair thus has a gain of

$$\frac{250 \parallel 620}{2 \times 25\ \Omega} \quad \text{or} \quad 3.5.$$

The second stage,  $Q_3$ , has a voltage gain of  $1.5\text{k}\ \Omega/2.5\ \Omega$ , or 600. The overall voltage gain at the quiescent point is  $3.5 \times 600$ , or 2100. Since  $Q_3$ 's gain depends on its collector current, there is substantial change of gain with signal

swing, i.e., nonlinearity. The gain is tabulated in the following section for three values of output voltage.

### Closed loop

**Input impedance** This circuit uses series feedback, so the input impedance is raised by  $(1 + \text{loop gain})$ . The feedback network is a voltage divider with  $B=1/30$  at signal frequencies, so the loop gain  $AB$  is 70. The input impedance is therefore  $70 \times 16\text{k}$ , still paralleled by the  $100\text{k}$  bias resistor, i.e., about  $92\text{k}$ . The bias resistor now dominates the input impedance.

**Output impedance** Since the output voltage is sampled, the output impedance is reduced by  $(1 + \text{loop gain})$ . The output impedance is therefore  $0.3\ \Omega$ . Note that this is a small-signal impedance and does not mean that a  $1\ \Omega$  load could be driven to nearly full swing, for instance. The  $5\ \Omega$  emitter resistors in the output stage limit the large signal swing. For instance, a  $4\ \Omega$  load could be driven only to  $10\ \text{Vpp}$ , approximately.

**Gain** The gain is  $A/(1+AB)$ . At the quiescent point that equals 30.84, using the exact value for  $B$ . To illustrate the gain stability achieved with negative feedback, the overall voltage gain of the circuit with and without feedback is tabulated at three values of output level at the end of this paragraph. It should be obvious that negative feedback has brought about considerable improvement in the amplifier's characteristics, although in fairness it should be pointed out that the amplifier could have been designed for better open-loop performance, e.g., by using a current source for  $Q_3$ 's collector load and degenerating its emitter, by using a current source for the differential-pair emitter circuit, etc. Even so, feedback would still make a large improvement.

$V_{\text{out}}$	<i>Open loop</i>			<i>Closed loop</i>		
	-10	0	+10	-10	0	+10
$Z_{\text{in}}$	16k	16k	16k	92k	92k	92k
$Z_{\text{out}}$	$20\Omega$	$20\Omega$	$20\Omega$	$0.3\Omega$	$0.3\Omega$	$0.3\Omega$
Gain	1360	2100	2400	30.60	30.84	30.90

## C. Series-feedback pair

Figure 2.92 shows another transistor amplifier with feedback. Thinking of  $Q_1$  as an amplifier of its base-emitter voltage drop (thinking in the Ebers–Moll sense), the feedback samples the output voltage and subtracts a fraction of it from the input signal. This circuit is a bit tricky because  $Q_2$ 's collector resistor doubles as the feedback network. Applying the techniques we used earlier, one can show that  $G(\text{open loop}) \approx 200$ , loop gain  $\approx 20$ ,  $Z_{\text{out}}(\text{open}$

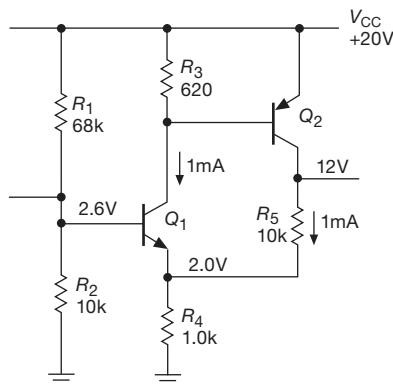


Figure 2.92. Series-feedback pair.

loop)  $\approx 10\text{k}$ ,  $Z_{\text{out}}(\text{closed loop}) \approx 500\Omega$ , and  $G(\text{closed loop}) \approx 9.5$ .

**Exercise 2.22.** Go for it!

## 2.6 Some typical transistor circuits

To illustrate some of the ideas of this chapter, let's look at a few examples of circuits with transistors. The range of circuits we can cover at this point is necessarily limited, because real-world circuits usually incorporate op-amps (the subject of Chapter 4) and other useful ICs – but we'll see plenty of transistors used alongside ICs in those later chapters.

### 2.6.1 Regulated power supply

Figure 2.93 shows a very common configuration.  $R_1$  normally holds  $Q_1$  on; when the output reaches 10 volts,  $Q_2$  goes into conduction (base at 5 V), preventing further rise of output voltage by shunting base current from  $Q_1$ 's base. The supply can be made adjustable by replacing  $R_2$  and  $R_3$  with a potentiometer. In this *voltage regulator* (or “regulated dc supply”) circuit, negative feedback acts to stabilize the output voltage:  $Q_2$  “looks at” the output and does something about it if the output isn't at the right voltage.

A few details: (a) Adding a biasing resistor  $R_4$  ensures a relatively constant zener current, so that the zener voltage does not change significantly with load current. It is tempting to provide that bias current from the input, but it is far better to use the regulated output. A warning is in order: whenever you use an output voltage to make something happen within a circuit, make sure that the circuit will start up correctly; here, however, there is no problem (why not?). (b) The capacitor  $C_1$  would probably be needed in

this circuit to ensure stability (i.e., to prevent oscillation), particularly if the output were capacitively bypassed (as it should be), for reasons we will see later in connection with feedback loop stability (§4.9).

We'll see much more of this subject in Chapter 9.

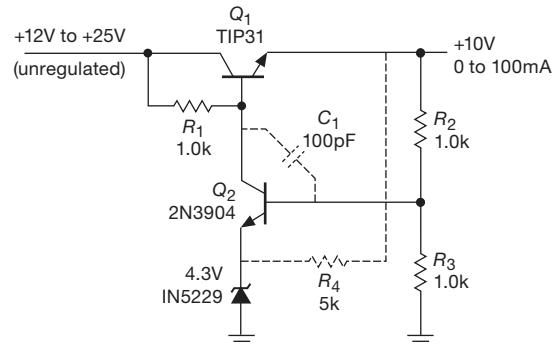


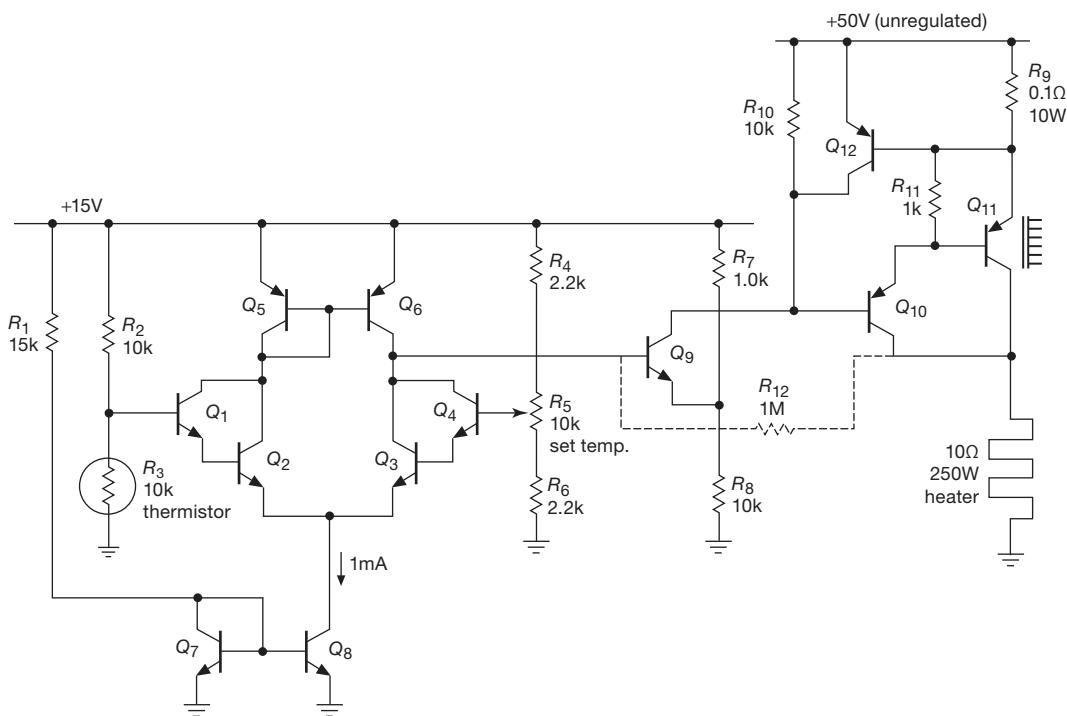
Figure 2.93. Feedback voltage regulator.

### 2.6.2 Temperature controller

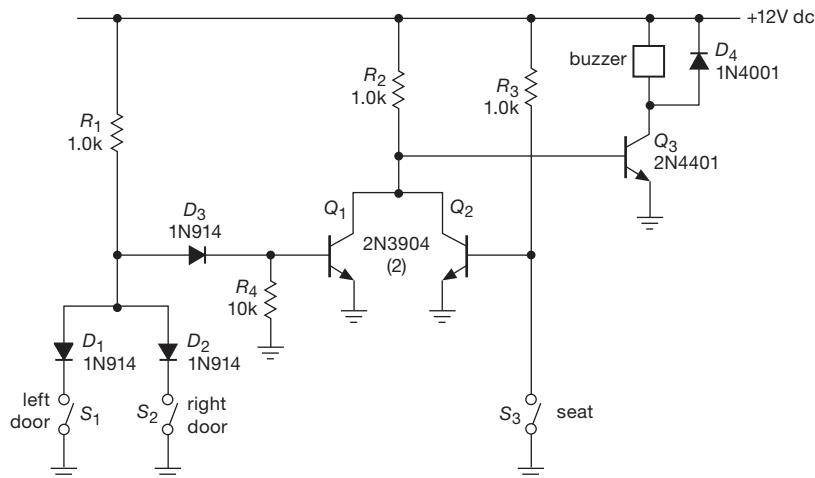
The schematic diagram in Figure 2.94 shows a temperature controller based on a *thermistor* sensing element, a device that changes resistance with temperature. Differential Darlington  $Q_1-Q_4$  compares the voltage of the adjustable reference divider  $R_4-R_6$  with the divider formed from the thermistor and  $R_2$ . (By comparing *ratios* from the same supply, the comparison becomes insensitive to supply variations; this particular configuration is called a Wheatstone bridge.) Current mirror  $Q_5Q_6$  provides an active load to raise the gain, and mirror  $Q_7Q_8$  provides emitter current.  $Q_9$  compares the differential amplifier output with a fixed voltage, saturating Darlington  $Q_{10}Q_{11}$  (which supplies power to the heater) if the thermistor is too cold.  $R_9$  is a current-sensing resistor that turns on protection transistor  $Q_{12}$  if the output current exceeds about 6 amps; that steals base drive from  $Q_{10}Q_{11}$ , preventing damage. And  $R_{12}$  adds a small amount of positive feedback, to cause the heater to snap on and off abruptly; this is the same trick (a “Schmitt trigger”) as in Figure 2.13.

### 2.6.3 Simple logic with transistors and diodes

Figure 2.95 shows a circuit that performs a task we illustrated in §1.9.1F: sounding a buzzer if either car door is open and the driver is seated. In this circuit the transistors all operate as switches (either OFF or saturated). Diodes  $D_1$  and  $D_2$  form what is called an OR gate, turning off  $Q_1$  if either door is open (switch closed). However, the collector of



**Figure 2.94.** Temperature controller for 250 W heater.



**Figure 2.95.** Both diodes and transistors are used to make digital logic “gates” in this seat-belt buzzer circuit.

$Q_1$  stays near ground, preventing the buzzer from sounding unless switch  $S_3$  is also closed (driver seated); in that case  $R_2$  turns on  $Q_3$ , putting 12 volts across the buzzer.  $D_3$  provides a diode drop so that  $Q_1$  is OFF with  $S_1$  or  $S_2$  closed, and  $D_4$  protects  $Q_3$  from the buzzer's inductive turn-off transient. In Chapters 10–15 we discuss logic circuitry in detail.

### **Additional Exercises for Chapter 2**

**Exercise 2.23.** Design a transistor switch circuit that allows you to switch two loads to ground by means of saturated *npn* transistors. Closing switch *A* should cause both loads to be powered, whereas closing switch *B* should power only one load. *Hint:* use diodes.

**Exercise 2.24.** Consider the current source in Figure 2.96. (a) What is  $I_{load}$ ? What is the output compliance? Assume  $V_{BE}$  is 0.6 V. (b) If  $\beta$  varies from 50 to 100 for collector voltages within the output compliance range, how much will the output current vary? (There are two effects here.) (c) If  $V_{BE}$  varies according to  $\Delta V_{BE} = -0.0001 \Delta V_{CE}$  (Early effect), how much will the load current vary over the compliance range? (d) What is the temperature coefficient of output current assuming that  $\beta$  does not vary with temperature? What is the temperature coefficient of output current assuming that  $\beta$  increases from its nominal value of 100 by 0.4%/°C?

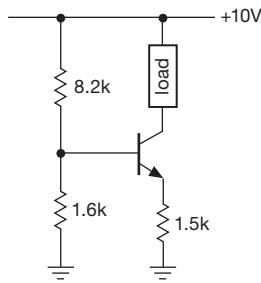


Figure 2.96. Current source exercise.

**Exercise 2.25.** Design a common-emitter *npn* amplifier with a voltage gain of 15,  $V_{CC}$  of +15 V, and  $I_C$  of 0.5 mA. Bias the collector at  $0.5V_{CC}$ , and put the low-frequency 3 dB point at 100 Hz.

**Exercise 2.26.** Bootstrap the circuit in the preceding problem to raise the input impedance. Choose the rolloff of the bootstrap appropriately.

**Exercise 2.27.** Design a dc-coupled differential amplifier with a voltage gain of 50 (to a single-ended output) for input signals near ground, supply voltages of ±15 volts, and quiescent currents of 0.1 mA in each transistor. Use a current source in the emitter and an emitter follower output stage.

**Exercise 2.28.** In this problem you will ultimately design an amplifier whose gain is controlled by an externally applied voltage (in Chapter 3 you will see how to do the same thing with FETs). (a) Begin by designing a long-tailed pair differential amplifier with emitter current source and no emitter resistors (undegenerated). Use ±15 V supplies. Set  $I_C$  (each transistor) at  $100 \mu\text{A}$ , and use  $R_C = 10\text{k}$ . Calculate the voltage gain from a single-ended input (other input grounded) to a single-ended output. (b) Now modify the circuit so that an externally applied voltage controls the emitter current source. Give an approximate formula for the gain as a function of controlling voltage. (In a real circuit you might arrange a second set of voltage-controlled current sources

to cancel the quiescent-point shift that gain changes produce in this circuit, or a differential-input second stage could be added to your circuit.)

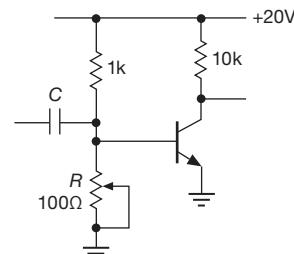


Figure 2.97. Bad biasing.

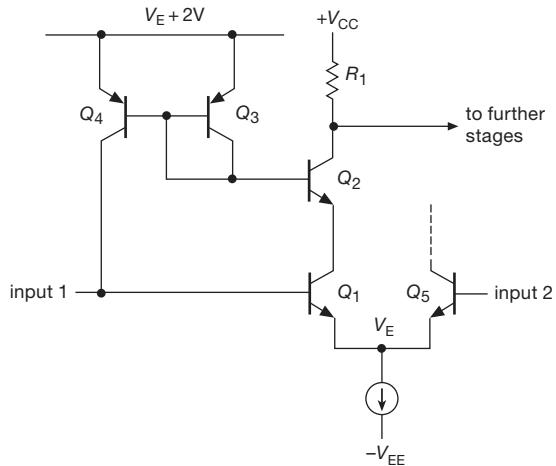


Figure 2.98. Base-current cancellation scheme used in precision operational amplifiers. Bias-current cancellation is discussed in detail in Chapter 4x.

**Exercise 2.29.** Disregarding the lessons of this chapter, a disgruntled student builds the amplifier shown in Figure 2.97. He adjusts  $R$  until the quiescent point is  $0.5V_{CC}$ . (a) What is  $Z_{in}$  (at high frequencies where  $Z_C \approx 0$ )? (b) What is the small-signal voltage gain? (c) What rise in ambient temperature (roughly) will cause the transistor to saturate?

**Exercise 2.30.** Several commercially available precision op-amps (e.g., the venerable OP-07) use the circuit in Figure 2.98 to cancel input bias current (only half of the symmetrical-input differential amplifier is shown in detail; the other half works the same way). Explain how the circuit works. Note:  $Q_1$  and  $Q_2$  are a beta-matched pair. Hint: it's all done with mirrors.

## Review of Chapter 2

An A-to-W review of Chapter 2. This review doesn't follow the exact topic order in the chapter: here we first cover transistor theory, then circle back to discuss some applications. In the chapter circuits have been interspersed with theory to provide motivation and illustrate how to use the theory.

### ¶A. Pin-Labeling Conventions.

The introduction (§2.1) describes some transistor and circuit-labeling conventions. For example,  $V_B$  (with a single subscript) indicates the voltage at the base terminal, and similarly  $I_B$  indicates current flowing into the base terminal.  $V_{BE}$  (two subscripts) indicates base-to-emitter voltage. Symbols like  $V_{CC}$  and  $V_{EE}$  (repeated subscripts) indicate the positive and negative supply voltages.

### ¶B. Transistor Types and Polarities.

Transistors are three-terminal devices capable of amplifying signals. They come in two broad classes, *bipolar junction transistors* (BJTs, the subject of this chapter), and *field-effect transistors* (FETs, the subject of Chapter 3). BJTs have a control terminal called the *base*, and a pair of output terminals, called the *collector* and the *emitter* (the corresponding terminals in a FET are *gate*, *drain*, and *source*). A signal applied to the base controls the current flowing from collector to emitter. There are two BJT polarities available, *npn* and *pnp*; for *npn* devices the collector is more positive than the emitter, and the opposite is true for *pnp*. Figure 2.2 illustrates this and identifies intrinsic diodes that are part of the transistor structure, see ¶D and ¶F below. The figure also illustrates that the collector current and the (much smaller) base current combine to form the emitter current.

**Operating modes** Transistors can operate as *switches* – turned ON or OFF – or they can be used as *linear* devices, for example as amplifiers, with an output current proportional to an input signal. Put another way, a transistor can be in one of three states: *cutoff* (non-zero  $V_{CE}$  but zero  $I_C$ ), *saturated* (non-zero  $I_C$  but near-zero  $V_{CE}$ ), or in the *linear region* (non-zero  $V_{CE}$  and  $I_C$ ). If you prefer prose (and using “voltage” as shorthand for collector-to-emitter voltage  $V_{CE}$ , and “current” as shorthand for collector current  $I_C$ ), the cutoff state has voltage but no current, the saturated state has current but near-zero voltage, and the linear region has both voltage and current.

### ¶C. Transistor Man and Current Gain.

In the simplest analysis, §2.1.1, the transistor is simply a current amplifier, with a *current gain* called *beta* (symbol  $\beta$ , or sometimes  $h_{FE}$ ). A current into the base causes

a current  $\beta$  times larger to flow from collector to emitter,  $I_C = \beta I_B$ , if the external circuit allows it. When currents are flowing, the base-emitter diode is conducting, so the base is  $\sim 0.65$  V more positive (for *npn*) than the emitter. The transistor doesn't *create* the collector current out of thin air; it simply throttles current from an available supply voltage. This important point is emphasized by our “transistor man” creation (Figure 2.7), a little homunculus whose job is to continuously examine the base current and attempt to adjust the collector's current to be a factor of  $\beta$  (or  $h_{FE}$ ) times larger. For a typical BJT the beta might be around 150, but beta is only loosely specified, and a particular transistor type may have a 3:1 spread (or more) in specified beta at some collector current (and further 3:1 spreads of  $\beta$  versus  $I_C$  and  $\beta$  versus temperature, see for example Figure 2.76).

### ¶D. Switches and Saturation.

When operated as a switch, §2.2.1, a current must be injected into the base to keep the transistor “ON.” This current must be substantially more than  $I_B = I_C / \beta$ . In practice a value of 1/10th of the maximum expected collector current is common, but you could use less, depending on the manufacturer's recommendations. Under this condition the transistor is in *saturation*, with 25–200 mV across the terminals. At such low collector-to-emitter voltages the base-to-collector diode in Figure 2.2 is conducting, and it robs some of the base-current drive. This creates an equilibrium at the saturation voltage. We'll return in ¶K to look at some circuit examples. See also the discussion of transistor saturation in Chapter 2x.

### ¶E. The BJT is a Transconductance Device.

As we point out in §2.1.1, “A circuit that depends on a particular value for beta is a bad circuit.” That's because  $\beta$  can vary by factors of 2 to 3 from the manufacturer's nominal datasheet value. A more reliable design approach is to use other highly-predictable BJT parameters that take into account that it is a *transconductance* device. In keeping with the definition of transconductance (an output current proportional to an input voltage), a BJT's collector current,  $I_C$ , is controlled by its base-to-emitter voltage,  $V_{BE}$ , see §2.3. (We can then rely on  $I_B = I_C / \beta$  to estimate the base current, the other way around from the simple approach in ¶C.) The transconductance view of BJTs is helpful in many circumstances (estimating gain, distortion, tempco), and it is essential in understanding and designing circuits such as differential amplifiers and current mirrors. However, in many situations you can circumvent the beta-uncertainty problem with circuit design tricks such as dc feedback or emitter degeneration, without explicitly invoking Ebers-Moll

(¶F). Note also that, just as it would be a bad idea to bias a BJT by applying a base current calculated from  $I_C/\beta$  (from an assumed  $\beta$ ), it would be even worse to attempt to bias a BJT by applying a calculated  $V_{BE}$  (from an assumed  $I_s$ , see ¶F); more on this in ¶Q, below. We might paraphrase this by saying “a circuit that depends on a particular value for  $I_s$ , or for operation at a precise ambient temperature, is a bad circuit.”

#### ¶F. Ebers–Moll.

Figure 2.41 shows a typical *Gummel plot*, with  $V_{BE}$  dictating  $I_C$ , and thus an approximate  $I_B$ . Equations (2.8) and (2.9) show the exponential (or logarithmic) nature of this relationship. A simple form of the equation,  $I_C=I_s \exp(V_{BE}/V_T)$  and its inverse,  $V_{BE}=V_T \log_e(I_C/I_s)$ , where the constant  $V_T=25\text{ mV}$  at  $25^\circ\text{C}$ , reveals that collector current is determined by  $V_{BE}$  and a parameter  $I_s$ , the latter related to the transistor die size and its current density.  $I_s$  is a very small current, typically some  $10^{11}$  times smaller than  $I_C$ . The Ebers–Moll formula accurately holds for the entire range of silicon BJT types, for example those listed in Table 8.1. The integrated-circuit (IC) industry relies on Ebers–Moll for the design of their highly-successful BJT linear circuits.

#### ¶G. Collector Current versus Base Voltage: Rules of Thumb.

See §2.3.2. It’s useful to remember a few rules of thumb, which we can derive from Ebers–Moll:  $I_C$  increases by a factor of ten for a  $\approx 60\text{ mV}$  increase in  $V_{BE}$ ; it doubles for an  $\approx 18\text{ mV}$   $V_{BE}$  increase, and it increases 4% for a  $1\text{ mV}$   $V_{BE}$  increase.

#### ¶H. Small Signals, Transconductance and $r_e$ .

See §2.3.2B. It’s convenient to assume operation at fixed  $I_C$ , and look for the effect of small changes (“small signals”). First, thinking about the rules of thumb above, we can calculate (eq’n 2.13,) the transconductance,  $g_m=\partial I_C/\partial V_{BE}=I_C/V_T$ . This evaluates to  $g_m=40\text{ mS}$  at  $1\text{ mA}$ , with  $g_m$  proportional to current. To put it another way, we can assign an effective internal resistance  $r_e$  in series with the emitter,  $r_e=1/g_m=V_T/I_C$ , see eq’n 2.12. (The small  $r$  indicates *small signal*.) A useful fact to memorize:  $r_e$  is about  $25\Omega$  at a collector current of  $1\text{ mA}$ , and it scales inversely with current.

#### ¶I. Dependence on Temperature.

See §2.3.2C. In ¶F we said  $V_T=25\text{ mV}$  at  $25^\circ\text{C}$ , which suggests it’s not exactly a constant, but changes with temperature. Because  $V_T=kT/q$  (§2.3.1), you might guess that  $V_{BE}$  is proportional to absolute temperature, thus a temper-

ature coefficient of about  $+2\text{mV}^\circ\text{C}$  (because  $V_{BE}\approx 600\text{ mV}$  at  $T=300\text{K}$ ). But the scaling parameter  $I_s$  has a large opposite tempco, producing an overall tempco of about  $-2.1\text{ mV}^\circ\text{C}$ . Memorize this fact also! Because  $V_T$  is proportional to absolute temperature, the tempco of transconductance at fixed collector current is inversely proportional to absolute temperature (recall  $g_m=I_C/V_T$ ), and thus drops by about  $0.34\%^\circ\text{C}$  at  $25^\circ\text{C}$ .

#### ¶J. Early Effect.

See §2.3.2D. In our simple understanding so far, base voltages (or currents) “program” a BJT’s collector current, independent of collector voltage. But in reality  $I_C$  increases slightly with increasing  $V_{CE}$ . This is called the *Early effect*, see eq’n 2.14 and Figure 2.59, which can be characterized by an *Early voltage*  $V_A$ , a parameter independent of operating current; see eq’n 2.15. If the Early voltage is low (a common drawback of *pnp* transistors) the effect can be quite large. For example, a *pnp* 2N5087 with  $V_A=55\text{ V}$  has  $\eta=4\times 10^{-4}$ , and would experience a  $4\text{ mV}$  shift of  $V_{BE}$  with a  $10\text{ V}$  change of  $V_{CE}$ ; if instead the base voltage were held constant, a  $10\text{ V}$  increase of collector voltage would cause a 17% increase of collector current. We hasten to point out there are circuit configurations, such as *degeneration*, or the *cascode*, that alleviate the Early effect. For more detail see the discussion in Chapter 2x.

### Circuit Examples

With this summary of basic BJT theory, we now circle back and review some circuit examples from Chapter 2. One way to review the circuits is to flip through the chapter looking at the pictures (and reading the captions), and refer to the associated text wherever you are uncertain of the underlying principles.

#### ¶K. Transistor Switches.

BJT switches are discussed in §2.2.1, and circuit examples appear in Figures 2.9 (driving an LED), 2.10 (high-side switching, including level shifting), and 2.16 (with an emitter-follower driver). Simply put, you arrange to drive a current into the base to put the transistor into solid saturation for the anticipated collector load current (i.e.,  $I_B \gg I_C/\beta$ ), bringing its collector within tens of millivolts of the emitter. More like this appears in Chapter 12 (Logic Interfacing). Looking forward, the use of *MOSFET* switches often provide a superior switching solution (§§3.4.4 and 3.5); their control terminal (the gate) conveniently requires *no* static gate current, though you may have to provide significant transient currents to charge its gate capacitance during rapid switching.

### ¶L. Transistor Pulsers.

Basic timer and pulse generator circuits are shown in Figures 2.11 (pulse from a step) and 2.12 (pulse from a pulse). These are simple, but not terribly accurate or stable; better to use a dedicated timer or pulse generator IC, see §7.2.

### ¶M. Schmitt Trigger.

A *Schmitt trigger* is a threshold level-detecting circuit (Figure 2.13) with hysteresis to prevent multiple transitions when noisy input signals go through the threshold(s). Although you can make a Schmitt trigger circuit with discrete transistors, good design practice favors the use of dedicated *comparator* ICs, see §§4.3.2 and 12.3.

### ¶N. Emitter Follower.

The emitter follower is a linear amplifier with an ideal voltage gain of unity, see §2.2.3. The beta of the transistor increases the follower's input impedance and reduces its output impedance, see §2.2.3B and eq'n 2.2. There's more detail in §2.3.3 and Figure 2.43, where the effect of the intrinsic emitter resistance  $r_e$  is taken into account. In simplified form  $R_{out} = r_e + R_s/\beta$ , where  $R_s$  is the signal source resistance seen at the base. The dc output voltage is offset from the dc input by  $V_{BE}$ , about 0.6 V to 0.7 V, unless a cancelling circuit is used, see §2.2.3D and Figure 2.29. Emitter followers are also used as voltage regulators, see §2.2.4 and Figures 2.21 and 2.22. A precision alternative is the *op-amp follower*, see §4.2.3 in Chapter 4.

### ¶O. Current Source (or Current Sink).

In contrast to the familiar *voltage source* (which delivers a constant voltage regardless of load current, think of a battery), a *current source* delivers a constant current regardless of the load's voltage drop, see §2.2.6 and Figure 2.31; there's no everyday "battery equivalent." Transconductance devices like BJTs, with their relatively constant collector output currents, are natural candidates for making current sources. For the simplest current source, the base is biased with a voltage, say  $V_b$ , with respect to a reference point (often ground), and the emitter is connected through a resistor to the same reference. For an *n-p-n* transistor with ground reference the output (sinking) current will be  $I_C = (V_b - V_{BE})/R_E$ , see Figure 2.32. For better stability and predictability the  $V_{BE}$  term can be cancelled, see Figure 2.33. The operating voltage range of a current source is called its *compliance range*, set on the low end by collector saturation, and on the high end by the transistor's breakdown voltage or by power-dissipation issues. Current sources are frequently created using current-mirror circuits, see ¶P below. Precise and stable current sources can

be made with op-amps (§4.2.5); there are also dedicated current-source integrated circuits (§9.3.14).

### ¶P. Current Mirrors.

A current mirror (§2.3.7) is a three-terminal current-source circuit that generates an output current proportional to an input "programming" current. In a typical configuration (Figures 2.55 and 2.58) the mirror attaches to a dc rail (or to ground), reflecting the programming current, the latter perhaps set by a resistor. The circuit often omits any emitter resistors, thus achieving compliance to within a fraction of a volt of the rail. Ordinarily you wouldn't attempt to apply exactly the right  $V_{BE}$  to generate a prescribed  $I_C$  (à la Ebers–Moll); but that's exactly what you're doing here. The trick is that one transistor ( $Q_1$ ) of the matched pair inverts Ebers–Moll, creating from the programming current  $I_P$  exactly the right  $V_{BE}$  to re-create the same current in the output transistor  $Q_2$ . Cute!

These circuits assume matched transistors, such as you would find inside an IC (recall from ¶G that even a 1 mV difference of  $V_{BE}$  produces a 4% change of current). Figure 2.62 graphs base-emitter voltage difference versus collector current ratio,  $\Delta V_{BE} = V_T \log_e(I_{C2}/I_{C1})$ . You can exploit this effect to generate a "ratio mirror," as discussed in Chapter 2x.

As nice as it looks, the basic current mirror of Figure 2.55 suffers from Early-effect change of output current when the output voltage changes. The effect is particularly serious with *p-n-p* transistors: in the example of a 2N5087 in ¶J above, the 4 mV change of  $V_{BE}$  (for a 10 V output change) would cause a 17% current error. One solution (Figure 2.60) is to add emitter degeneration resistors, at the expense both of compliance near the reference rail and of dynamic range. A more elegant solution is the Wilson mirror (Figure 2.61), which defeats Early effect by exploiting the ever-useful *cascode* configuration (Figure 2.84B). Cascode transistor  $Q_3$  passes output transistor  $Q_2$ 's collector current to the load, while  $Q_2$  operates with a fixed  $V_{CE}$  of one diode drop (its own  $V_{BE}$ ). The Wilson mirror's ingenious configuration also cancels base-current errors (an ordinary mirror with BJTs having  $\beta=100$  has a current error of 2%). Degeneration resistors can be added, as shown in circuit B, for additional suppression of Early effect, but they would be omitted in a "pure Wilson mirror." Linear ICs are full of Wilson mirrors. See Chapter 2x for a discussion of *bipolarity* current mirrors.

### ¶Q. Common-Emitter Amplifiers.

See §§2.2.7 and 2.3.4, and Figures 2.35, 2.48 and 2.50. The simplest form of BJT amplifier has a grounded emitter, a load resistor  $R_L$  from the collector to a supply  $V_+$ , and

a dc bias plus a small signal voltage applied to the base. The gain is  $G_V = -R_L/r_e$ . If the base bias is carefully set so that the collector current pulls the collector halfway to ground, then  $I_C = V_s/2R_L$ ,  $r_e = V_T/I_C = 2R_L V_T/V_s$ , and so the voltage gain (recall  $V_T \approx 25 \text{ mV}$ ) is  $G_V = -20V_s$ , where  $V_s$  is in units of volts. For  $V_s = 20 \text{ V}$ , for example, the voltage gain is  $-400$ .

That's a lot of gain! Unless the signals are small, however, there's a serious problem: the gain is inverse in  $r_e$ , thus proportional to  $I_C$ . But the latter changes as the output voltage swings up and down, producing first-order changes in gain, with resulting severe distortion (Figure 2.46). This can be alleviated (at the expense of gain) by adding *emitter degeneration* in the form of an emitter resistor  $R_E$ . The gain is then  $G_V = -R_L/(R_E + r_e)$ , with greatly reduced effect of varying  $r_e$ ; see Figure 2.47, where emitter degeneration was added to reduce the gain by a factor of ten ( $R_E = 9r_e$ ). This is also a form of negative feedback, see §2.3.4B and ¶W below. You can think of this circuit as a classic current source (¶O) driving a resistor as load; the voltage gain is the current source's transconductance multiplied by the load resistance,  $G_V = g_m R_L$ , where  $g_m = -1/r_e$ .

We've sidestepped the important issue of setting the base bias voltage to produce the desired quiescent collector current. But we don't know the appropriate voltage  $V_{BE}$ , and a small change has a big effect, see ¶G above (e.g., a  $60 \text{ mV}$  uncertainty in  $V_{BE}$ , which is about what you might encounter from different batches of a given transistor, produces a  $10\times$  error in  $I_C$ !). There are many circuit solutions, see §2.3.5, but the simplest involves adding emitter degeneration at dc, bypassed as necessary to produce higher gain at signal frequencies (Figure 2.50 and 2.51). Another approach is to use a matching transistor to set the bias, analogous to the current mirror (Figure 2.52); this method is inherent in the widely-used *differential amplifier* (Figure 2.65). A third approach is to exploit feedback to set the bias (Figures 2.53 and 2.54), a method that figures centrally in op-amp circuits (Chapter 4).

#### ¶R. Differential Amplifiers.

The differential amplifier (§2.3.8) is a symmetrical configuration of two matched transistors, used to amplify the difference of two input signals. It may include emitter degeneration (Figure 2.64), but need not (Figure 2.65). For best performance the emitter pulldown resistor is replaced by a current source, and (for highest gain) the resistive collector load is replaced by a current mirror (Figure 2.67). Differential amplifiers should reject strongly any common-mode input signal, achieving a good common-mode rejection ratio (CMRR, the ratio  $G_{\text{diff}}/G_{\text{CM}}$ ). Differential amplifiers can be used to amplify single-ended input signals (ground

the other input), where the inherent cancellation of  $V_{BE}$  offsets allows accurate dc performance (§2.3.8B). Ordinarily you use only one output from a differential amplifier; that is, it is used to convert a balanced input to a single-ended output. But you can use both outputs (a “fully-differential amplifier,” §5.17) to drive a balanced load, or to create a pair of signals  $180^\circ$  out of phase (a *phase splitter*). See also the sections on the emitter-input differential amplifier and on BJT amplifier distortion in Chapter 2x, and §5.13–§5.16 (precision differential and instrumentation amplifiers).

#### ¶S. Comparators.

A differential amplifier with lots of gain  $G_{\text{diff}}$  is driven into differential saturation with a small differential input (§2.3.8E). For example, just a few millivolts of input difference is adequate to saturate the output if  $G_{\text{diff}}=1000$  (easily accomplished with a current-mirror collector load). When operated in this way, the differential amplifier is a voltage *comparator*, a circuit used widely to sense thresholds or compare signal levels; it's the basis of analog-to-digital conversion, and figures importantly in Chapter 12 (see §12.3 and Tables 12.1 and 12.2).

#### ¶T. Push–Pull Amplifiers.

A single transistor conducts in one direction only (e.g., an *npn* transistor can only sink current from its collector, and source current from its emitter). That makes it awkward to drive a heavy load with alternating polarity (e.g., a loudspeaker, servomotor, etc.), although it can be done, wastefully, with a single-ended stage (“class-A”) with high quiescent current, see Figure 2.68. The push-pull configuration uses a pair of transistors connected to opposite supply rails (§2.4.1), an arrangement that can supply large output currents of either polarity with little or no quiescent current. Figure 2.69 shows a push-pull follower with complementary polarities, and with zero quiescent current (“class-B”); this produces some crossover distortion, which can be eliminated by biasing the pair into quiescent conduction (“class-AB,” Figure 2.71). The output transistors can be beta-boosting configurations like the Darlington or Sziklai (¶U), see for example Figure 2.78. The push-pull configuration is widely used in logic circuits (see Figure 10.25), gate driver ICs (see Figure 3.97), and in combination with op-amps to deliver greater output currents (see Figure 4.26).

#### ¶U. The Darlington and Sziklai Connections.

These simple combinations of two transistors create a 3-terminal equivalent transistor with  $\beta = \beta_1 \beta_2$ . The Darlington (Figures 2.74 and 2.75) cascades two transistors of the

same polarity and has a base-emitter drop of  $2V_{BE}$ ; the Sziklai (Figure 2.77) pairs opposite polarities, and has a single base-emitter drop (which is only weakly dependent on output current, thanks to  $R_B$ ). For either configuration a resistor  $R_B$  should be connected across the output transistor's base-emitter terminals. For more about this subject see the discussion in Chapter 2x.

#### ¶V. Miller Effect.

Like all electronic components, transistors have inter-terminal capacitances, designated (by terminal pairs)  $C_{be}$ ,  $C_{ce}$ , and  $C_{cb}$ .<sup>72</sup> While  $C_{be}$  and  $C_{ce}$  slow the input and output waveforms by creating lowpass filters with the source and load resistances, the effect of the feedback capacitance  $C_{cb}$  is more insidious: it creates an additional input capacitance to ground equal to  $C_{cb}$  multiplied by the stage's inverting voltage gain, thus its effective input capacitance becomes  $C_{eff} = (G_V + 1) C_{cb}$ . This is the infamous *Miller effect* (§2.4.5B), whose impact can be devastating in high-speed and wideband amplifiers. Some circuit solutions include

the grounded-base amplifier, the differential amplifier, and the cascode configuration (see the discussion of cascode in Chapter 2x).

#### ¶W. Negative Feedback.

If there were a Nobel prize for grand-concepts-in-circuit-design, it would surely go to Harold Black for his elegant elucidation of *negative feedback*. In its simplest form, it consists of subtracting, from the input signal, a fraction  $B$  of an amplifier's output signal  $V_{out}$  (Figure 2.85). If the amplifier's open-loop gain is  $A$ , then the closed-loop gain becomes (eq'n 2.16)  $G_{cl} = A / (1 + AB)$ . The quantity  $AB$ , which generally is large compared with unity, is called the *loop gain*, and it (more precisely the quantity  $1 + AB$ ) is the multiplier by which negative feedback improves the amplifier's performance: improved linearity and constancy of gain, and (in this *series feedback* circuit configuration) raised input impedance and lowered output impedance; see §2.5.3.

Feedback is the essence of linear design, and it is woven deeply into the DNA of op-amp circuits (the subject of Chapter 4), and power circuits (Chapter 9). With negative feedback you can make amplifiers with 0.0001% distortion, voltage sources with  $0.001 \Omega$  output impedance, and many other wonders too magnificent here to relate. Stay tuned. Better yet, *read on!*

<sup>72</sup> These have many aliases (a common set uses initials for "in" and "out" instead of "base" and "collector," thus  $C_{ie}$ ,  $C_{oe}$ , and  $C_{ob}$ , respectively), see the section on BJT bandwidth in Chapter 2x.

# FIELD-EFFECT TRANSISTORS

## CHAPTER 3

### 3.1 Introduction

Field-effect transistors (FETs) are different from the bipolar transistors<sup>1</sup> that we talked about in the last chapter. Broadly speaking, however, they are similar devices, which we might call *charge-control devices*: in both cases (Figure 3.1) we have a three-terminal device in which the conduction between two electrodes depends on the availability of charge carriers, which is controlled by a voltage applied to a third *control electrode*.

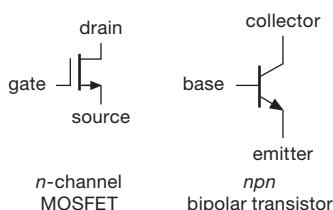


Figure 3.1. The *n*-channel MOSFET and its *npn* transistor analog.

Here's how they differ: in a bipolar transistor the collector-base junction is back-biased, so no current normally flows. Forward-biasing the base-emitter junction by  $\approx 0.6\text{ V}$  overcomes its diode "contact potential barrier," causing electrons to enter the base region, where they are strongly attracted to the collector. Although some base current results, most of these "minority carriers" are captured by the collector. This results in a collector current, controlled by a (much smaller) base current. The collector current is proportional to the rate of injection of minority carriers into the base region, which is an exponential function of  $V_{BE}$  (the Ebers-Moll equation). You can think of a bipolar transistor as a current amplifier (with roughly constant current gain  $\beta$ ) or as a transconductance device (Ebers-Moll: collector current programmed by base-emitter *voltage*).

In an FET, as the name suggests, conduction in a *channel* is controlled by an *electric field*, produced by a voltage applied to the *gate* electrode. There are no forward-biased junctions, so the gate draws no current. This is per-

haps the most important advantage of the FET. As with BJTs, there are two polarities, *n*-channel FETs (conduction by electrons) and *p*-channel FETs (conduction by holes). These two polarities are analogous to the familiar *npn* and *pnp* bipolar transistors, respectively. In addition, however, FETs tend to be confusing at first because they can be made with two different kinds of gates (thus JFETs and MOSFETs) and with two different kinds of channel doping (leading to *enhancement* and *depletion* modes). We'll sort out these possibilities shortly.

First, though, some motivation and perspective. The FET's nonexistent gate current is its most important characteristic. The resulting high input impedance (which can be greater than  $10^{14}\Omega$ ) is essential in many applications, and in any case it makes circuit design simple and fun. For applications like analog switches and amplifiers of ultra-high input impedance, FETs have no equal. They can be easily used by themselves or combined with bipolar transistors to make integrated circuits. In the next chapter we'll see how successful that process has been in making nearly perfect (and wonderfully easy to use) *operational amplifiers*, and in Chapters 10–14 we'll see how digital electronics has been revolutionized by MOSFET integrated circuits. Because many FETs using very low current can be constructed in a small area, they are especially useful for very large-scale integration (VLSI) digital circuits such as microprocessors, memory, and "application-specific" chips of the sort used in cellphones, televisions, and the like. At the other end of the spectrum, robust high-current MOSFETs (50 amps or more) have replaced bipolar transistors in many applications, often providing simpler circuits with improved performance.

### 3.1.1 FET characteristics

Beginners sometimes become catatonic when directly confronted with the confusing variety of FET types. That variety arises from the combined choices of polarity (*n*-channel or *p*-channel), form of gate insulation [semiconductor junction (JFET) or oxide insulator (MOSFET)], and channel doping (*enhancement* or *depletion* mode). Of the

<sup>1</sup> Often called BJTs, for "bipolar junction transistors," to distinguish them from FETs.

eight resulting possibilities, six *could* be made, and five actually are. Four of those five are of major importance.

It will aid understanding (and sanity), however, if we begin with one type only, just as we did with the *npn* bipolar transistor. Once comfortable with FETs, we'll have little trouble with their family tree.

### A. FET V-I curves

Let's look first at the *n*-channel enhancement-mode MOSFET, which is analogous to the *npn* bipolar transistor (Figure 3.2). In normal operation the drain ( $\sim$ collector) is more positive than the source ( $\sim$ emitter). No current flows from drain to source unless the gate ( $\sim$ base) is brought positive with respect to the source. Once the gate is thus "forward-biased," there will be drain current, all of which flows to the source. Figure 3.2 shows how the drain current  $I_D$  varies with drain-source voltage  $V_{DS}$  for a few values of controlling gate-source voltage  $V_{GS}$ . For comparison, the corresponding "family" of curves of  $I_C$  versus  $V_{BE}$  for an ordinary *npn* bipolar transistor is shown. Evidently there are a lot of similarities between *n*-channel MOSFETs and *npn* bipolar transistors.

Like the *npn* transistor, the FET has a high incremental drain impedance, giving roughly constant current for  $V_{DS}$  greater than a volt or two. By an unfortunate choice of language, this is called the "saturation" region of the FET (a better term is "current saturation") and corresponds to the "active" region of the bipolar transistor. Analogous to the bipolar transistor, a larger gate-to-source bias produces a larger drain current. And, analogous to bipolar transistors, FETs are not perfect transconductance devices (constant drain current for constant gate-source voltage): just as the ideal Ebers–Moll transconductance characteristic of bipolar transistors is degraded by the Early effect (§2.3.2D and §2x.8), there's an analogous departure from the ideal transconductance behavior for FETs, characterized by a finite drain output resistance  $r_o$  (more usually called  $1/g_{os}$ , see §3.3.2 and §3x.4).

So far, the FET looks just like the *npn* transistor. Let's look closer, though. For one thing, over the normal range of currents the saturation drain current increases rather modestly with increasing gate voltage ( $V_{GS}$ ). In fact, it is approximately proportional to  $(V_{GS} - V_{th})^2$ , where  $V_{th}$  is the gate threshold voltage at which drain current begins ( $V_{th} \approx 1.63\text{V}$  for the FET in Figure 3.2); compare this mild quadratic law with the steep exponential transistor law, as given to us by Ebers and Moll. Second, there is *zero* dc gate current, so you mustn't think of the FET as a device with current gain (which would be infinite). Instead, think of the FET as a transconductance device, with gate-source

voltage programming the drain current, as we did with the bipolar transistor in the Ebers–Moll treatment. Recall that the transconductance  $g_m$  is simply the ratio  $i_d/v_{gs}$  (using the convention of lowercase letters to indicate "small-signal" changes in a parameter; e.g.,  $i_d/v_{gs} = \delta I_D/\delta V_{GS}$ ). Third, the gate of a MOSFET is truly insulated from the drain-source channel; thus, unlike the situation for bipolar transistors (or JFETs, as we'll see), you can bring it positive (or negative) at least 10 V or more without worrying about diode conduction. Finally, the FET differs from the bipolar transistor in the so-called *linear* (low-voltage) region of the graph, where it behaves rather accurately like a resistor, *even for negative*  $V_{DS}$ ; this turns out to be quite useful because the equivalent drain-source resistance is, as you might guess, programmed by the gate-source voltage.

### B. Two examples

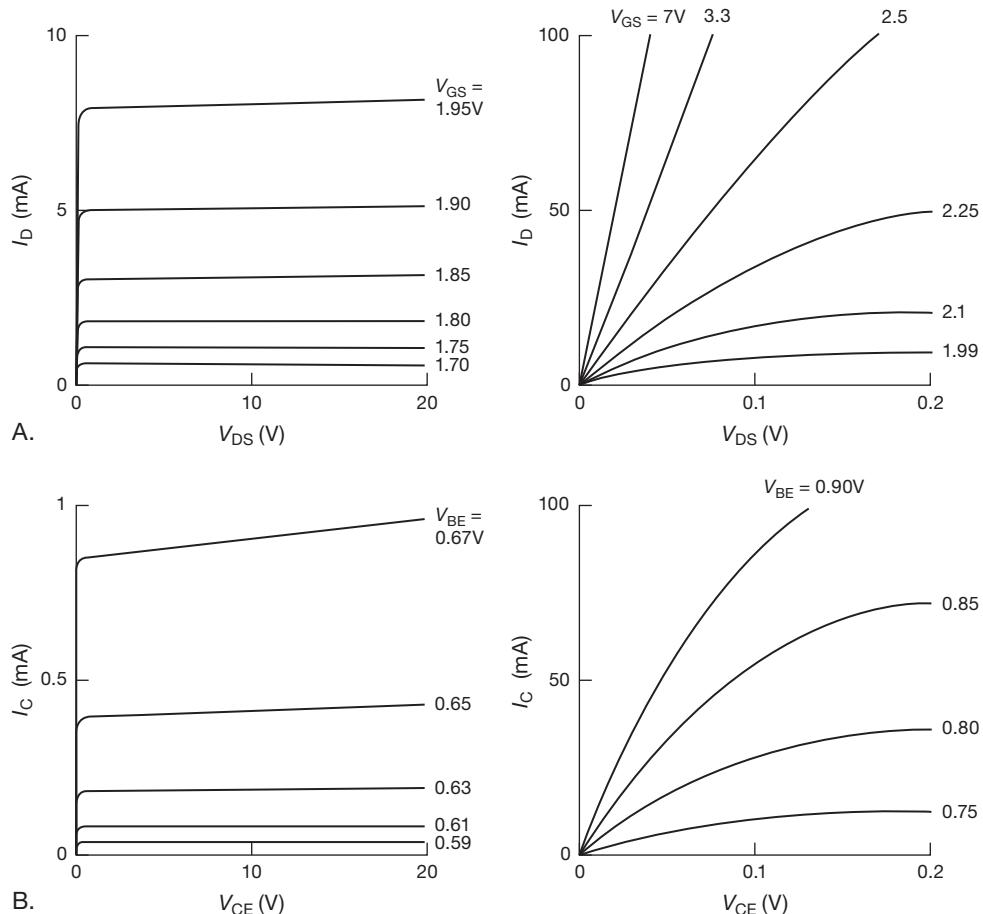
FETs will have more surprises in store for us. But before getting into more details, let's look at two simple switching applications. Figure 3.3 shows the MOSFET equivalent of Figure 2.5, our first saturated transistor switch. The FET circuit is even simpler, because we don't have to concern ourselves with the inevitable compromise of providing adequate base drive current (considering worst-case minimum  $\beta$  combined with the lamp's cold resistance) without squandering excessive power. Instead, we just apply a full-swing dc voltage drive to the cooperative high-impedance gate. As long as the switched-on FET behaves like a resistance that is small compared with the load, it will bring its drain close to ground; typical power MOSFETs have  $R_{ON} < 0.1\Omega$ , which is fine for this job.

We demonstrate this circuit in our electronics course, but we put a resistor in series with the gate. The students are surprised when they discover its resistance –  $10\text{ M}\Omega$  – which implies a "beta" of at least 100,000. They are even more surprised when they notice that the light stays on when the gate is then open-circuited: the gate voltage is held on the gate's capacitance, and will stay that way for the rest of the hour's lecture.<sup>2</sup> That implies that the gate current is well below a *picoampere*!

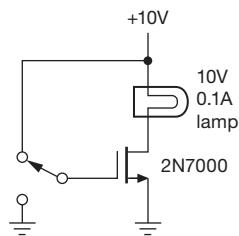
Figure 3.4 shows an "analog switch"<sup>3</sup> application, which cannot be done at all with bipolar transistors. The idea here is to switch the conduction of a FET from open-circuit (gate reverse-biased) to short-circuit (gate forward-biased), thus blocking or passing the analog signal (we'll see plenty of reasons to do this sort of thing later). In this

<sup>2</sup> The gate capacitance "remembers" whatever voltage was last applied. So you can have it stay on, stay off, or even stay at half-brightness, with no noticeable change even with the gate floating.

<sup>3</sup> Also called a "linear switch."

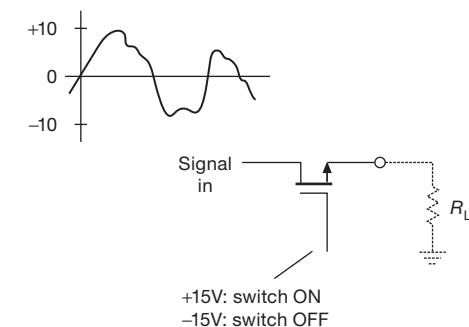


**Figure 3.2.** Measured MOSFET/transistor characteristic curves: A. VN0106 (similar to the popular 2N7000) *n*-channel MOSFET:  $I_D$  versus  $V_{DS}$  for various values of  $V_{GS}$ . B. 2N3904 *npn* bipolar transistor:  $I_C$  versus  $V_{CE}$  for various values of  $V_{BE}$ .



**Figure 3.3.** MOSFET power switch.

case we just arrange for the gate to be driven more negative than any input signal swing (switch *open*), or a few volts more positive than any input signal swing (switch *closed*). Bipolar transistors aren't suited to this application, because the base draws current and forms diodes with the emitter and collector, producing awkward clamping action. The MOSFET is delightfully simple by comparison, need-



**Figure 3.4.** MOSFET analog (signal) switch.

ing only a voltage swing into the (essentially open-circuit) gate.<sup>4</sup>

<sup>4</sup> It's only fair to mention that our treatment of this circuit has been some-

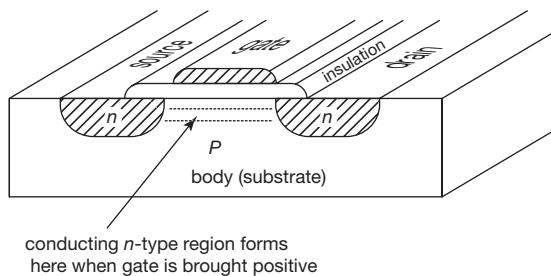


Figure 3.5. An *n*-channel “lateral” MOSFET.

### 3.1.2 FET types

#### A. *n*-channel, *p*-channel

Now for the family tree. First, FETs (like BJTs) can be fabricated in both polarities. Thus the mirror twin of our *n*-channel MOSFET is a *p*-channel MOSFET. Its behavior is symmetrical, mimicking *pnp* transistors: the drain is normally negative with respect to the source, and drain current flows if the gate is brought at least a volt or two negative with respect to the source. The symmetry isn’t perfect because the carriers are holes, rather than electrons, with lower mobility and minority carrier lifetime.<sup>5</sup> The consequence is worth remembering – *p*-channel FETs usually have poorer performance, manifested as a higher gate threshold voltage, higher  $R_{ON}$ , and lower saturation current.<sup>6</sup>

#### B. MOSFET, JFET

In a MOSFET (“Metal-Oxide-Semiconductor Field-Effect Transistor”) the gate region is separated from the conducting channel by a thin layer of  $\text{SiO}_2$  (glass) grown onto the channel (Figure 3.5). The gate, which may be either metal or doped silicon, is truly insulated from the source-drain circuit, with characteristic input resistance  $>10^{14}\Omega$ . It affects channel conduction purely by its electric field. MOSFETs are sometimes called *insulated-gate* FETs, or IGFETs. The gate insulating layer is quite thin, typically less than a wavelength of light, and can withstand gate voltages up to  $\pm 20\text{ V}$  in typical power MOSFETs (less for the

what simplistic, for instance ignoring the effects of gate-channel capacitance and the variation of  $R_{ON}$  with signal swing. We’ll have more to say about analog switches later.

<sup>5</sup> These are semiconductor parameters of importance in transistor performance.

<sup>6</sup> In the case of so-called “complementary pairs” (an *n*-channel and a *p*-channel part with similar voltage and current ratings), the *p*-channel part is usually built with a larger area in order to match the performance of the *n*-channel part. You can see the evidence in the datasheet in the form of greater capacitance for the *p*-channel part.

small MOSFETs in low-voltage integrated circuits). MOSFETs are easy to use because the gate can swing either polarity relative to the source without any gate current flowing. They are, however, quite susceptible to damage from static electricity; you can destroy a MOSFET device literally by touching it.

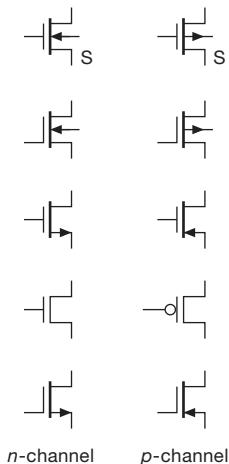


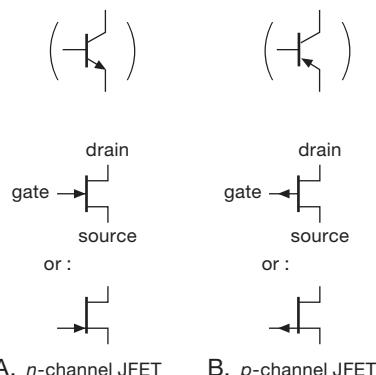
Figure 3.6. MOSFET schematic symbols.

The symbols for MOSFETs are shown in Figure 3.6. The extra terminal that is sometimes shown is the “body,” or “substrate,” the piece of silicon in which the FET is fabricated. Because the body forms a diode junction with the channel, it must be held at a nonconducting voltage. It can be tied to the source or to a point in the circuit more negative (positive) than the source for *n*-channel (*p*-channel) MOSFETs. It is common to see the body terminal omitted; furthermore, engineers often use the symbol with the symmetrical gate. Unfortunately, with what’s left you can’t tell source from drain; worse still, you can’t tell *n*-channel from *p*-channel! In this book we most often use the bottom pair of schematic symbols, which, though somewhat unconventional, are unambiguous and uncluttered.<sup>7</sup>

In a JFET (Junction Field-Effect Transistor) the gate forms a semiconductor junction with the underlying channel. This has the important consequence that *a JFET gate should not be forward biased with respect to the channel, to prevent gate current*. For example, diode conduction will occur as the gate of an *n*-channel JFET approaches  $+0.6\text{ V}$  with respect to the more negative end of the channel (which is usually the source). The gate is therefore operated reverse-biased with respect to the channel, and no

<sup>7</sup> In current practice, logic designers like to use the second pair up from the bottom, while power MOSFET users prefer the second pair down from the top.

current (except diode leakage) flows in the gate circuit. The circuit symbols for JFETs are shown in Figure 3.7. Once again, we favor the symbol with offset gate to identify the source (though JFETs and small integrated MOSFETs are symmetrical, power MOSFETs are quite asymmetrical, with very different capacitances and breakdown voltages).



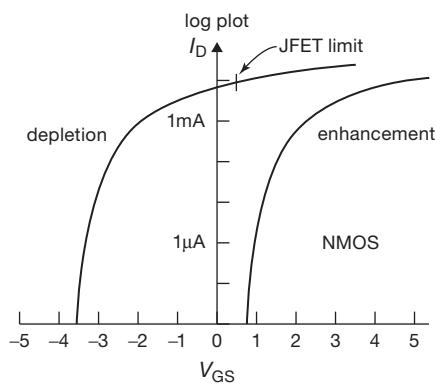
**Figure 3.7.** JFET schematic symbols: A. *n*-channel JFET. B. *p*-channel JFET.

### C. Enhancement, depletion

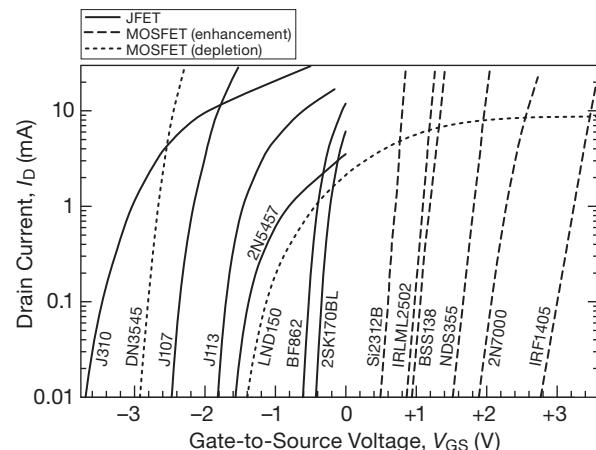
The *n*-channel MOSFETs with which we began the chapter were nonconducting, with zero (or negative) gate bias, and were driven into conduction by bringing the gate positive with respect to the source. This kind of FET is known as *enhancement mode*. The other possibility is to manufacture the *n*-channel FET with the channel semiconductor “doped” so that there is plenty of channel conduction even with zero gate bias, and the gate must be reverse-biased by a few volts to cut off the drain current. Such a FET is known as *depletion mode*. MOSFETs can be made in either variety, because the gate, being insulated from the channel, can swing either polarity. But JFETs, with their gate-channel diode, permit only reverse gate bias, and therefore are made only in depletion mode.

A graph of drain current versus gate-source voltage, at a fixed value of drain voltage, may help clarify this distinction (Figures 3.8 and 3.9). The enhancement-mode device draws no drain current until the gate is brought positive (these are *n*-channel FETs) with respect to the source, whereas the depletion-mode device is operating at nearly its maximum value of drain current when the gate is at the same voltage as the source. In some sense the two categories are artificial, because the two curves are identical except for a shift along the  $V_{GS}$  axis. In fact, it is possible to manufacture “in-between” MOSFETs. Nevertheless,

the distinction is an important one when it comes to circuit design.



**Figure 3.8.** Transfer characteristics ( $I_D$  versus  $V_{GS}$ ) for a JFET (depletion-mode) and a MOSFET (enhancement-mode) transistor. See also the measured curves in Figure 3.19.



**Figure 3.9.** Lending some authenticity to Figure 3.8’s notional sketch: measured  $I_D$  versus  $V_{GS}$  for a selection of *n*-channel FETs.

Note that JFETs are always depletion-mode devices and that the gate cannot be brought more than about 0.5 V more positive (for *n*-channel) than the source, since the gate-channel diode will conduct. MOSFETs can be either enhancement or depletion, but in practice the dominant species is enhancement, with a sprinkling of depletion-mode MOSFETs.<sup>8</sup> Most of the time, then, you need worry only about (a) depletion-mode JFETs and

<sup>8</sup> In the form of *n*-channel GaAs FETs, “dual-gate” cascodes for radiofrequency applications, and a selection of high-voltage depletion-mode power MOSFETs (such as the Supertex lateral LND150 or vertical DN3435, as well as offerings by six other manufacturers).

(b) enhancement-mode MOSFETs. Each come in the two polarities, *n*-channel and *p*-channel.

### 3.1.3 Universal FET characteristics

A family tree (Figure 3.10) and a map (Figure 3.11) of input–output voltage (source grounded) may help simplify things. The different devices (including garden-variety *npn* and *pnp* bipolar transistors) are drawn in the quadrant that characterizes their input and output voltages when they are in the active region with source (or emitter) grounded. You don't have to remember the properties of the five kinds of FETs, though, because they're all basically the same.

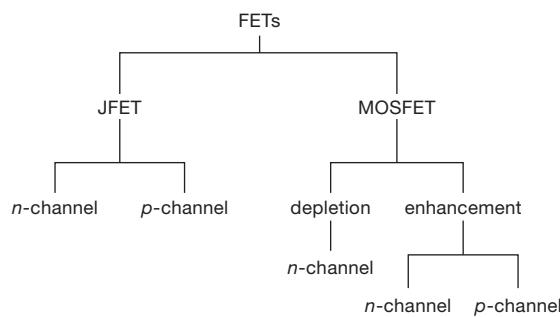


Figure 3.10. FET family tree.

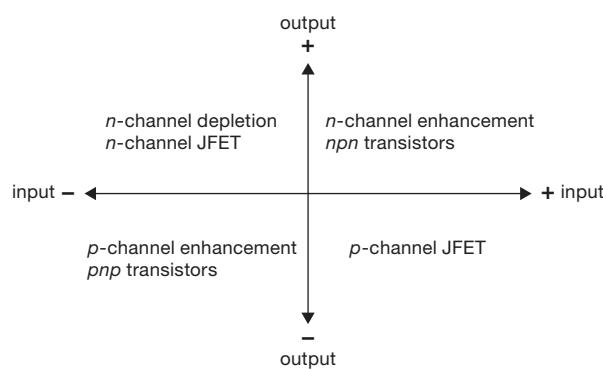


Figure 3.11. Transistor “polarity map.”

First, with the source grounded, a FET is turned on (brought into conduction) by bringing the gate voltage “toward” the active drain supply voltage. This is true for all five types of FETs, as well as the bipolar transistors. For example, an *n*-channel JFET (which is necessarily depletion mode) uses a positive drain supply, as do all *n*-type devices. Thus a positive-going gate voltage tends to turn on the JFET. The subtlety for depletion-mode devices is that the gate must be (negatively) back-biased for zero drain

current, whereas for enhancement-mode devices zero gate voltage is sufficient to give zero drain current.

Second, because of the near symmetry of source and drain, either terminal can act as the effective source (exception: not true for power MOSFETs, in which the body is internally connected to the source). When thinking of FET action, and for purposes of calculation, remember that the effective source terminal is always the one most “away” from the active drain supply. For example, suppose a FET is used to switch a line to ground, and both positive and negative signals are present on the switched line, which is usually selected to be the FET drain. If the switch is an *n*-channel MOSFET (therefore enhancement) and a negative voltage happens to be present on the (turned-off) drain terminal, then that terminal is actually the “source” for purposes of gate turn-on voltage calculation. Thus a negative gate voltage larger than the most negative signal, rather than ground, is needed to ensure turn-off.

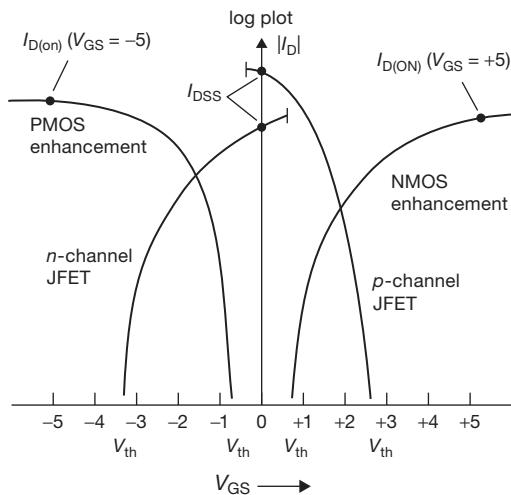
The graph in Figure 3.12 may help to sort out all these confusing ideas. Again, the difference between enhancement and depletion is merely a question of displacement along the  $V_{GS}$  axis, i.e., whether there is a lot of drain current or no drain current at all when the gate is at the same potential as the source. The *n*-channel and *p*-channel FETs are complementary in the same way as *npn* and *pnp* bipolar transistors.

In Figure 3.12 we have used standard symbols for the important FET parameters of saturation current and cut-off voltage. For JFETs the value of drain current with the gate shorted to source is specified on the datasheets as  $I_{DSS}$  and is nearly the maximum drain current possible. ( $I_{DSS}$  means current from drain to source with the gate shorted to the source. Throughout the chapter you will see this notation, in which the first two subscripted letters designate the pair of terminals and the third specifies the condition.) For enhancement-mode MOSFETs the analogous specification is  $I_{D(ON)}$ , given at some forward gate voltage (“ $I_{DSS}$ ” would be zero for any enhancement-mode device).

For JFETs the gate-source voltage at which the drain current is brought essentially to zero<sup>9</sup> is called the “gate-source cutoff voltage,”  $V_{GS(OFF)}$ , or (sometimes) the “pinch-off voltage,”  $V_P$ , and is typically in the range of  $-1\text{ V}$  to  $-5\text{ V}$  (positive for *p*-channel, of course). The analogous quantity is not normally specified for enhancement-mode MOSFETs;<sup>10</sup> instead, datasheets specify the

<sup>9</sup> Usually chosen to be  $10\text{ nA}$ ; a pinchoff-voltage test circuit is described in §4.3.4.

<sup>10</sup> We will use the symbol  $V_{th}$  to designate the analogous idealized “gate-source cutoff voltage” for MOSFETs, which we will need in some



**Figure 3.12.** Important gate voltages and drain currents.

“gate-source threshold voltage,”  $V_{GS(th)}$ , at which the onset of drain current has reached a small but arbitrary threshold value, typically 0.25 mA.  $V_{GS(th)}$  is typically in the range of 0.5–5 V, in the “forward” direction, of course.

With FETs it is easy to get confused about polarities. For example, *n*-channel devices, which usually have the drain positive with respect to the source, can have positive or negative gate voltage and positive (enhancement) or negative (depletion) threshold voltages. To make matters worse, the drain can be (and often is) operated negative with respect to the source. Of course, all these statements go in reverse for *p*-channel devices. In order to minimize confusion, we will always assume that we are talking about *n*-channel devices unless explicitly stated otherwise. Likewise, because MOSFETs are nearly always enhancement mode, and JFETs are always depletion mode, we’ll omit those designations from now on.

### 3.1.4 FET drain characteristics

In Figure 3.2 we showed a family of curves of  $I_D$  versus  $V_{DS}$  that we measured for a VN0106, an *n*-channel

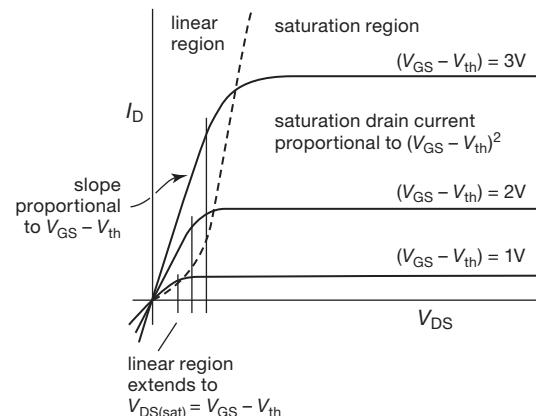
discussion that follows. In the electronics literature the symbol  $V_T$  is used for this quantity, called the “threshold voltage”; but we prefer to avoid the same symbol that is used for the “thermal voltage”  $V_T$  in the Ebers–Moll equation, where  $V_T = kT/q \approx 25$  mV. And don’t confuse  $V_{th}$  with  $V_{GS(th)}$ :  $V_{th}$  is obtained from extrapolating a  $\sqrt{I_D}$  versus  $V_{GS}$  plot; it’s not found in datasheets, but it’s quite useful. By contrast,  $V_{GS(th)}$  is not terribly useful, but it’s the quantity you find in datasheets.

enhancement-mode MOSFET.<sup>11</sup> We remarked that FETs behave like pretty good transconductance devices over most of the graph (i.e.,  $I_D$  nearly constant for a given  $V_{GS}$ ), except at small  $V_{DS}$ , where they approximate a resistance (i.e.,  $I_D$  proportional to  $V_{DS}$ ). In both cases the applied gate-source voltage controls the behavior, which can be well described by the FET analog of the Ebers–Moll equation. Let’s look now at these two regions a bit more closely; we’ll revisit this important subject in greater detail in §3.3 and again in the advanced-topics Chapter 3x.

Figure 3.13 shows the situation schematically. In both regions the drain current depends on  $V_{GS} - V_{th}$ , the amount by which the applied gate-source voltage exceeds the threshold (or pinch-off) voltage. The linear region, in which drain current is approximately proportional to  $V_{DS}$ , extends up to a voltage  $V_{DS(sat)}$ , after which the drain current is approximately constant. The slope in the linear region,  $I_D/V_{DS}$ , is proportional to the gate bias,  $V_{GS} - V_{th}$ . Furthermore, the drain voltage at which the curves enter the “saturation region,”  $V_{DS(sat)}$ , is approximately  $V_{GS} - V_{th}$ , making the saturation drain current,  $I_{D(sat)}$ , proportional to  $(V_{GS} - V_{th})^2$ , the quadratic law we mentioned earlier. For reference, here are the universal FET drain-current formulas:

$$I_D = 2\kappa[(V_{GS} - V_{th})V_{DS} - V_{DS}^2/2] \quad (\text{linear region}) \quad (3.1)$$

$$I_D = \kappa(V_{GS} - V_{th})^2 \quad (\text{saturation region}) \quad (3.2)$$



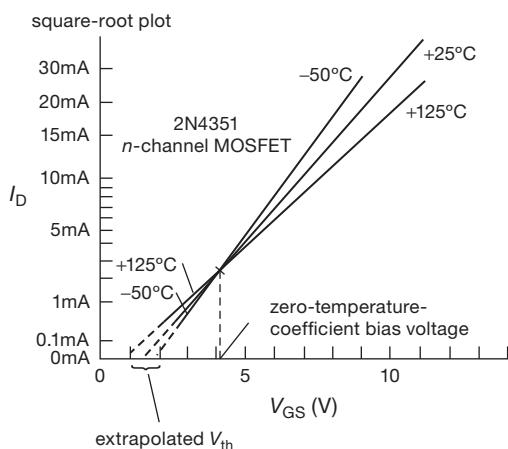
**Figure 3.13.** Linear and saturation regions of FET operation.

If we call  $V_{GS} - V_{th}$  (the amount by which the gate-source voltage exceeds the threshold) the “gate drive,” the

<sup>11</sup> The VN0106 is not widely available. It is similar to the very popular 2N7000 or BS170 (in the TO-92 package) and to the 2N7002, BSS138, or MMBF170 (in the SMT packages).

important results are that (a) the resistance in the linear region is inversely proportional to the gate drive, (b) the linear region extends to a drain-source voltage approximately equal to the gate drive, and (c) the saturation drain current is proportional to the square of the gate drive. These equations assume that the body is connected to the source. Note that the “linear region” is not really linear because of the  $V_{DS}^2$  term; we’ll show a clever circuit fix later.

The scale factor  $\kappa$  depends on particulars such as the geometry of the FET, oxide capacitance, and carrier mobility.<sup>12</sup> It has a temperature dependence  $\kappa \propto T^{-3/2}$ , which alone would cause  $I_D$  to decrease with increasing temperature. However,  $V_{th}$  also depends slightly on temperature ( $2\text{--}5\text{ mV}/^\circ\text{C}$ ); the combined effect produces the curve of drain current versus temperature, as shown in Figure 3.14.



**Figure 3.14.** The “threshold voltage”  $V_{th}$  is found by extrapolating a square-root plot of  $I_D$  to zero drain current. The FET saturation drain current has a negative tempco in the high-current regime.

At large gate voltages the negative temperature coefficient of  $\kappa$  causes the drain current to decrease with increasing temperature. As a consequence, FETs of a given type, operating in this high-current regime, can often be paralleled without the external current-equalizing (“emitter-ballasting”) resistors that you must use with bipolar transistors (see §3.6.3).<sup>13</sup> This same negative coefficient also prevents thermal runaway in local regions of the junction (an

effect known as “current hogging”), which severely limits the power capability of large bipolar transistors, as we’ll see when we discuss “second breakdown” and “safe operating area” in Chapter 9.

At small drain currents (where the temperature coefficient of  $V_{th}$  dominates),  $I_D$  has a positive tempco, with a point of zero temperature coefficient at some drain current in between. This effect is exploited in FET op-amps to minimize temperature drift, as we’ll see in the next chapter.

### A. Subthreshold region

Our expression given earlier for saturation drain current does not apply for very small drain currents. This is known as the “subthreshold” region, where the channel is below the threshold for conduction, but some current flows anyway because of a small population of thermally energetic electrons. If you’ve studied physics or chemistry, you probably know in your bones that the resulting drain current is exponential (with some scale factor) in the difference voltage  $V_{GS} - V_{th}$ .

We measured some MOSFETs over nine decades of drain current (1 nA to 1 A) and plotted the result as a graph of  $I_D$  versus  $V_{GS}$  (Figure 3.15). The region from 1 nA to 1 mA is quite precisely exponential; above this subthreshold region the curves enter the normal “quadratic” region. For the *n*-channel MOSFET (Supertex type VN01, similar to the ever-popular 2N7000) we checked out a sample of 20 transistors (from four different manufacturing runs spread over two years), plotting the extreme range to give you an idea of the variability (see next section). Note the somewhat poorer characteristics ( $V_{th}$ ,  $I_{D(ON)}$ ) of the “complementary” VP01 (similar to the popular BS250).

JFETs exhibit similar behavior, as illustrated in the measured data of Figure 3.16 (though  $V_{GS}$  is necessarily limited to reverse-bias voltage polarity, or at most to a forward-bias less than a diode drop). The quadratic region, where  $I_D \propto (V_{GS} - V_{th})^2$ , is seen most clearly by plotting the *square root* of drain current versus gate voltage; see Figure 3.14 and Figure 3.51 later in the chapter.

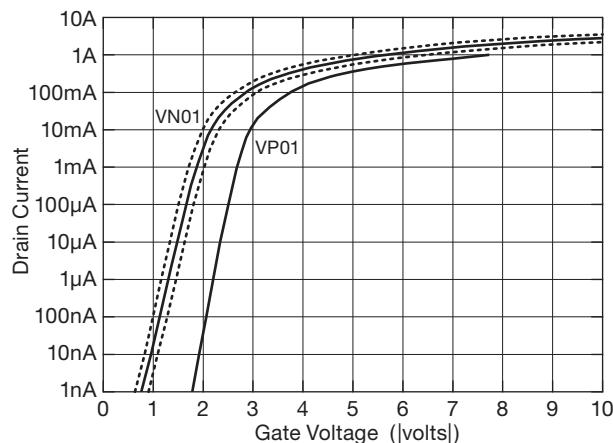
### 3.1.5 Manufacturing spread of FET characteristics

Before we look at some circuits, let’s take a look at the range of FET parameters (such as  $I_{DSS}$  and  $V_{GS(th)}$ ), as well as their manufacturing “spread” among devices of the same nominal type, in order to get a better idea of the FET. Unfortunately, many of the characteristics of FETs show a much greater process spread than the corresponding

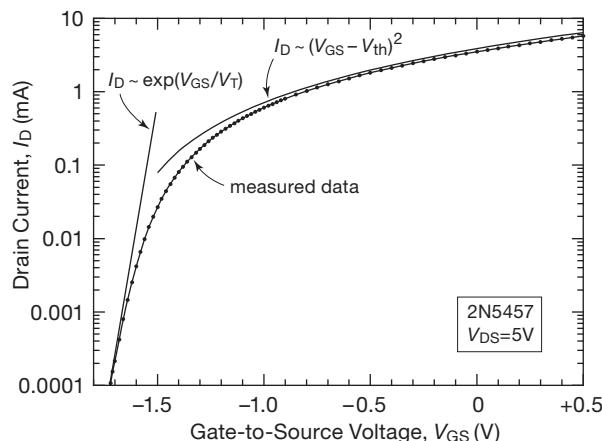
<sup>12</sup> You’ll commonly see the symbol  $k$  used here. We prefer  $\kappa$ , to avoid confusion with the Boltzmann constant  $k$  that figures in the Ebers–Moll equation for bipolar transistor behavior. The SPICE model for JFETs calls this parameter  $\beta$  (and for  $V_{th}$  it uses the parameter “VTO”).

<sup>13</sup> Some cautions apply, most notably with ordinary (“vertical”) power MOSFETs in linear applications, where they are operated at drain currents well below the region of negative temperature coefficient – see §3.5.1B and §3.6.3. In such applications (e.g., audio power amplifiers)

the alternative “lateral” MOSFET is popular, owing to its stabilizing negative coefficient.



**Figure 3.15.** Measured MOSFET saturation drain current versus gate-source voltage. For the VN01 the dotted curves are the extreme specimens, and the solid curve is the median, from a group of 20 MOSFETs.



**Figure 3.16.** Five decades of measured drain current versus gate-to-source voltage for the *n*-channel 2N5457 JFET. In the subthreshold region the drain current is exponential, like a BJT, with nearly the same scale factor  $V_T$  ( $kT/q$ , or 25.3 mV at room temperature); at higher currents it becomes quadratic (the calculated curve has been offset by +10% for clarity).

characteristics of bipolar transistors, a fact that the circuit designer must keep in mind. For example, the 2N7000 (a typical *n*-channel MOSFET) has a specified  $V_{GS(th)}$  of 0.8–3 V ( $I_D = 1 \text{ mA}$ ), compared with the analogous  $V_{BE}$  spread of 0.63–0.83 V (also at  $I_C = 1 \text{ mA}$ ) for a small *npn* bipolar transistor. Here's what you can expect:

### FET Characteristics: Manufacturing Spread

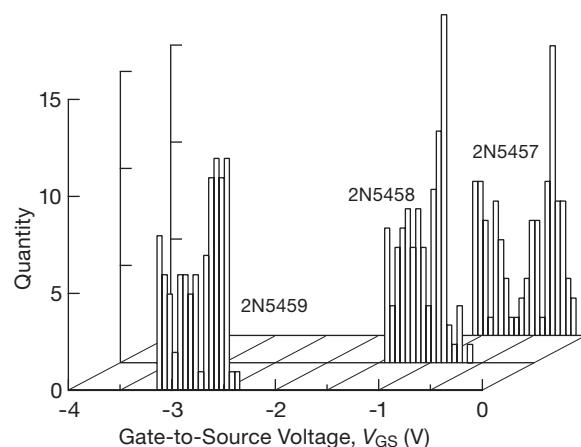
Characteristic	Available Range	Spread
$I_{DSS}, I_{D(ON)}$	1 mA to 500 A	$\times 5$
$R_{DS(ON)}$	$0.001\Omega$ to 10k	$\times 5$
$g_m @ 1 \text{ mA}$	500–3000 $\mu\text{S}$	$\times 5$
$V_P$ (JFETs)	0.5–10 V	5 V
$V_{GS(th)}$ (MOSFETs)	0.5–5 V	2 V
$BV_{DS(OFF)}$	6–1000 V	
$BV_{GS(OFF)}$	6–125 V	

$R_{DS(ON)}$  is the drain-source resistance (linear region, i.e., small  $V_{DS}$ ) when the FET is conducting fully, e.g., with the gate grounded in the case of JFETs or with a large applied gate-source voltage (usually specified as 10 V) for MOSFETs.  $I_{DSS}$  and  $I_{D(ON)}$  are the saturation-region (large  $V_{DS}$ ) drain currents under the same turned-on gate drive conditions.  $V_P$  is the pinch-off voltage (JFETs),  $V_{GS(th)}$  is the turn-on gate threshold voltage (MOSFETs), and the  $BV$ 's are breakdown voltages. As you can see, a JFET with a grounded source may be a good current source, but you can't predict very well what the current will be. Likewise, the  $V_{GS}$  needed to produce some value of drain current can vary considerably, in contrast to the predictable ( $\approx 0.6 \text{ V}$ )  $V_{BE}$  of bipolar transistors. Figure 3.17 illustrates this latter point graphically: we measured the  $V_{GS}$  values at a drain current of 1 mA for a hundred pieces each (hey, they're pretty cheap: about \$0.10 each) of three popular JFET types (the 2N5457–59 series, graded by their  $I_{DSS}$ ). The spread of gate-source voltages, within each type, is about 1 V. For comparison, look at the analogous plot for BJTs in Figure 8.44; there the spread is just 10–20 mV.

### A. Matching of characteristics

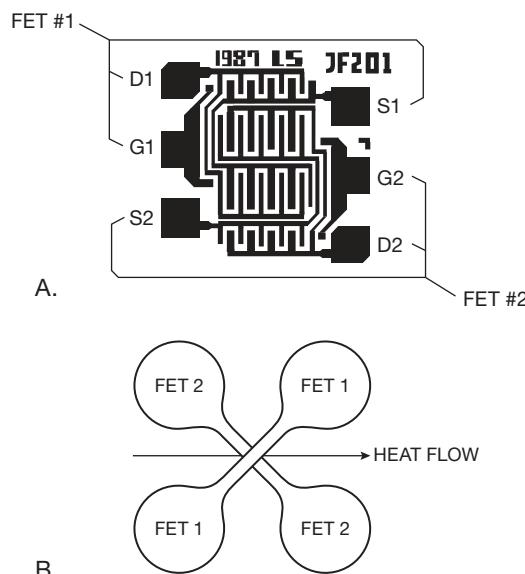
As you can see, FETs are inferior to bipolar transistors in  $V_{GS}$  predictability, i.e., they have a large spread in the  $V_{GS}$  required to produce a given  $I_D$ . Devices with a large process spread will, in general, have a larger offset (voltage unbalance) when used as differential pairs. For instance, typical run-of-the-mill bipolar transistors might show a spread in  $V_{BE}$  of 25 mV or so, at some collector current, for a selection of off-the-shelf transistors. The comparable “official” figure (as specified on datasheets) for MOSFETs is more like 1 V to 2 V!<sup>14</sup> Because FETs have some very desirable characteristics otherwise, it is worthwhile putting in some

<sup>14</sup> In practice, we've found considerably better matching within a single batch of MOSFETs, sometimes as tightly matched as 50 mV or so. On the other hand, a more typical spread within one batch is several hundred millivolts, as illustrated later in Figure 3.41. If matching is



**Figure 3.17.** We wore our fingers to the bone (a “digital” measurement?) collecting  $V_{GS}$  values (at  $V_{DS}=5$  V and  $I_D=1$  mA) for 300 JFETs in the popular 2N5457–59 series. Compare with the analogous histograms in Figure 8.44.

extra effort to reduce these offsets in specially manufactured matched pairs. IC designers use techniques like interdigitation (two devices sharing the same general piece of IC real estate) and thermal-gradient cancellation schemes to improve performance (Figure 3.18).



**Figure 3.18.** Techniques for transistor matching: A. Interdigitation (Courtesy of Linear Integrated Systems.) B. Temperature gradient cancellation.

important in some application (when several transistors are used in parallel, for example), you should measure the actual parts.

The results are impressive. Although FET devices still cannot equal bipolar transistors in  $V_{GS}$  matching, their performance is adequate for most applications. For example, the best previously available matched FET had<sup>15</sup> a voltage offset of 0.5 mV and tempco of 5  $\mu\text{V}/^\circ\text{C}$  (max), whereas the best bipolar pair has values of 25  $\mu\text{V}$  and 0.3  $\mu\text{V}/^\circ\text{C}$  (max), roughly 20 times better.

Operational amplifiers (the universal high-gain differential amplifiers we’ll see in the next chapter) are available in both flavors; you would generally choose one with bipolar innards for high precision (because of its close input-transistor  $V_{BE}$  matching), whereas a FET-input op-amp is the obvious choice for high-impedance applications (because its inputs – FET gates – draw no current). For example, the inexpensive JFET-input LF411 and LF412 that we will use as our all-around op-amp in the next chapter has a typical input (leakage) current of 50 pA and costs \$0.60; the popular MOSFET-input TLC272 costs about the same and has a typical input (leakage) current of only 1 pA! Compare this with a common bipolar op-amp, the LM324, with typical input (bias) current of 45,000 pA (45 nA).<sup>16</sup>

### 3.1.6 Basic FET circuits

Now we’re ready to look at FET circuits. You can usually find a way to convert a circuit that uses BJTs into one using FETs – but the new circuit may not be an improvement! For the remainder of the chapter we’d like to illustrate circuit situations that take advantage of the unique properties of FETs, i.e., circuits that work better with FETs or that you can’t build at all with bipolar transistors. For this purpose it may be helpful to group FET applications into categories; here are the most important, as we see it.

**High-impedance/low-current.** Buffers or amplifiers for applications in which the base current and finite input

<sup>15</sup> Sadly, these parts are no longer available. But the art of transistor matching is alive and well in the innards of op-amps, for which the best JFET specimen has an offset and tempco of 0.1 mV and 1  $\mu\text{V}/^\circ\text{C}$ , respectively, whereas the best BJT specimen has 0.01 mV and 0.1  $\mu\text{V}/^\circ\text{C}$ , i.e., 10 times better.

<sup>16</sup> BJT enthusiasts would cry “foul!,” and point out that you can use superbeta BJTs, combined with bias-current cancellation schemes, to bring the input current down to 25 pA; they would further point out that FET-input current (which is leakage) rises dramatically with temperature, whereas BJT-input current (which is honest bias current) is stable or even tends to decrease slightly (see Figure 3.48). FET enthusiasts would prevail, though, with the rejoinder that MOSFET-input amplifiers like the dual LMC6042 have typical input currents of 2 *femtoamps* (that’s 0.000002 nA!).

impedance of BJTs limit performance. Although you can build such circuits with discrete FETs, current practice favors using integrated circuits built with FETs. Some of these use FETs as a high-impedance front-end for an otherwise bipolar design, whereas others use FETs throughout. When available FET ICs do not provide adequate performance, a hybrid approach (discrete JFET front-end, assisted by an op-amp) can push the performance envelope.

**Analog switches.** MOSFETs are excellent voltage-controlled analog switches, as we hinted in §3.1.1B. We'll look briefly at this subject. Once again, you should generally use dedicated "analog switch" ICs, rather than building discrete circuits.

**Digital logic.** MOSFETs dominate microprocessors, memory, special-purpose VLSI, and most high-performance digital logic. They are used exclusively in micropower logic and low-power portable devices. Here, too, MOSFETs make their appearance in integrated circuits. We'll see why FETs are preferable to BJTs.

**Power switching.** Power MOSFETs are usually preferable to ordinary bipolar power transistors for switching loads, as we suggested in our first circuit of the chapter. For this application you use *discrete* power FETs.

**Variable resistors; current sources.** In the "linear" region of the drain curves, FETs behave like voltage-controlled resistors; in the "saturation" region they are voltage-controlled current sources. You can exploit this intrinsic behavior of FETs in your circuits.

**Generalized replacement for bipolar transistors.** You can use FETs in oscillators, amplifiers, voltage regulators, and radiofrequency circuits (to name a few), where bipolar transistors are also normally used. FETs aren't guaranteed to make a better circuit – sometimes they will, sometimes they won't. You should keep them in mind as an alternative.

Now let's look at these subjects. We'll adopt a slightly different order, for clarity.

## 3.2 FET linear circuits

**A note to the reader:** This section and the next (§§3.2 and 3.3) deal primarily with *JFETs*, which are well suited to linear applications such as current sources, followers, and amplifiers. If you need a low-noise amplifier with extremely high input impedance, the JFET is your friend (and maybe your *only* friend). Readers wishing to move directly to MOSFETs, starting with FET switches, may wish to

**Table 3.1 JFET Mini-table<sup>a</sup>** (see also JFET Table 3.7)

Part #	Curve	$V_{GS(off)}$			measured at 1mA			$C_{rss}$ typ (pF)	$R_{on}$ typ (Ω)
		$I_D$	$I_{DSS}$ (mA)	min (V)	max (V)	$V_{GS}$ (V)	$g_m$ (mS)		
2N5484	A	1–5	-0.3	-3	-0.73	2.3	180	1	-
2N5485	B	4–10	-0.4	-4	-1.7	2.1	110	1	-
2N5486	C	8–20	-2	-6	-2.4	2.1	50	1	-
2N5457	D	1–5	-0.5	-6	-0.81	2.0	200	1.5	-
2N5458	E	2–9	-1	-7	-2.3	2.3	170	1.5	-
2N5459	F	4–16	-2	-8	-2.8	2.0	100	1.5	-
BF862	G	10–25	-0.3	-1.2	-0.40	12	250	1.9	-
J309	H	12–30	-1	-4	-1.6	4.2	300	2	50
J310	J	24–60	-2	-6.5	-3.0	4.3	100	2	50
J113	K	2–	-0.5	-3	-1.5	5.7	140	3	50
J112	L	5–	-1	-5	-3.3	5	100	3	30
PN4393	M	5–30	-0.5	-3	-0.83	6.2	100	3.5	100
PN4392	N	25–75	-2	-5	-2.6	5.4	130	3.5	60
LSK170B	P	6–12	-0.2	-2	-0.09	11	160	5	-
J110	Q	10–	-0.5	-4	-1.2	6.1	220	8	18
J107	R	100–	-0.5	-4.5	-2.6	8.2	340	35	8
J105	-	500–	-4.5	-10	-8.7	6.4	60	35	3
IF3601	S	30–	-0.04	-3	-0.24	27	1400	300	-

**Notes:** (a) sorted by family  $C_{rss}$ , and within each family by increasing  $I_{DSS}$ . (b)  $G_{max}=g_m/g_{os}$ , the maximum grounded-source voltage gain into a current source as drain load;  $G_{max}$  is proportional to  $V_{DS}$  (tabulated values are at  $V_{DS}=5V$ ), and for most JFETs  $G_{max}$  is relatively constant over varying  $|I_D|$ .

skip over these JFET materials,<sup>17</sup> and proceed directly to §3.4 on page 171, where we launch into the MOSFET-dominated subjects of signal switching, digital logic, and power switching.

### 3.2.1 Some representative JFETs: a brief tour

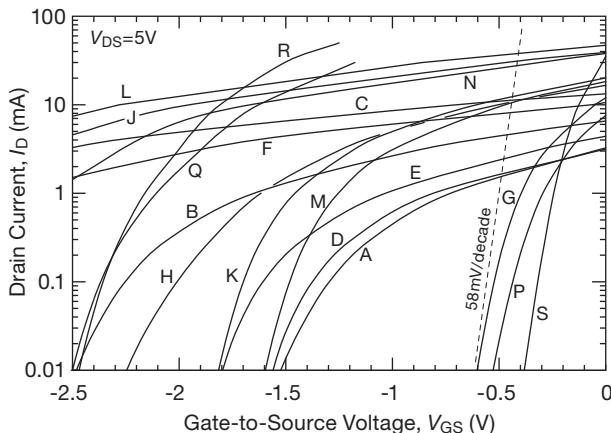
Table 3.1 lists a small selection of representative *n*-channel JFETs.<sup>18</sup> Let's take a look at what you get.

This selection includes only *n*-channel JFETs, the dominant polarity. Complements with similar characteristics are sometimes available, for example the *p*-channel 2N5460–62 for the *n*-channel 2N5457–59; see Table 3.7 on page 217 for additional examples.

Many JFETs come in families of three or four parts, graded by  $I_{DSS}$  and  $V_{GS(off)}$ , which alleviates somewhat

<sup>17</sup> You'll want to study this material, though, if you want to understand MOSFET linear amplifiers, because we address topics like the significance of an FET's transconductance and output conductance, and their variation with drain voltage and current.

<sup>18</sup> The expanded Table 3.7 on page 217 includes many more JFETs; later in this chapter there are analogous tables of MOSFETs (Tables 3.4a and 3.4b, pages 188–191, Table 3.5, page 206, and Table 3.6, page 210).



**Figure 3.19.** Measured drain current versus gate-source voltage for the JFETs in Table 3.1 on the previous page.

the annoying circuit design problems created by the wide spread of those parameters. But even those graded families may present a spread of as much as 5:1 (or more). Note also that JFETs intended for switching applications (the ones specifying  $R_{on}$ ) may specify only a *minimum* value of  $I_{DSS}$ : what can you say, for example, about the likely value for a J110 (specified as  $I_{DSS} = 10 \text{ mA}$ , minimum)? Answer: not much – our sample measured 122 mA!

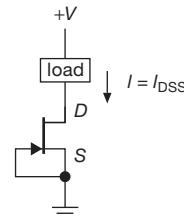
In many applications (amplifiers, followers) you want lots of transconductance gain,  $g_m$ . JFET datasheets usually specify  $g_m$  at the part's  $I_{DSS}$ , but that's not terribly useful if you don't know what  $I_{DSS}$  is. Moreover, the listed  $g_m$  at  $I_{DSS}$  is afflicted with the usual specification spread, typically 5:1 or so. Unlike BJTs, for which the transconductance is predictably given by  $g_m = 1/r_e = I_C/V_T$  (where  $V_T = kT/q \approx 25.3 \text{ mV}$ ), the transconductance of different JFET types can vary by an order of magnitude, even when each is operated at the same drain current. In Table 3.1 on the previous page we've listed measured values of  $g_m$ , all at a standard current of 1 mA.<sup>19</sup> At these currents their transconductance is much less than that of a BJT (where  $g_m = 40 \text{ mS}$  at 1 mA), though they compete well at very low currents (the subthreshold region). This behavior can be seen in the different slopes of the measured  $I_D$  versus  $V_{GS}$  curves of Figure 3.19.

The column labeled  $G_{\max}$  lists the voltage gain when used as a grounded-source amplifier with current-source load; in that case the effective load resistance is related to a quantity called  $g_{os}$ , the output conductance seen looking into the drain with the gate voltage held constant (analogous

to the Early effect in BJTs; more on this in Chapter 3x). Here, too, there is a wide spread among JFET types.

A parameter that is important in low-level amplification is a JFET's input noise voltage, not listed here but treated in detail in Chapter 8. The standout happens to be the IF3601 (an amazing  $e_n = 0.3 \text{ nV}/\sqrt{\text{Hz}}$ ), but your pact with the devil is the high 300 pF capacitance of the large-area junction.<sup>20</sup>

There's much more to say about the inhabitants of the JFET zoo, as we'll see in connection with Table 3.7 on page 217. Chapter 8 discusses JFETs in connection with noise (§§8.6 and 8.6.5), with a table of relevant parts (Table 8.2 on page 516).



**Figure 3.20.** An *n*-channel JFET current sink.

### 3.2.2 JFET current sources

JFETs are used as current sources within integrated circuits (particularly op-amps), and also sometimes in discrete designs. The simplest JFET current source is shown in Figure 3.20; we chose a JFET, rather than a MOSFET, because it needs no gate bias (it's depletion mode). From a graph of FET drain characteristics (Figure 3.21) you can see that the current will be reasonably constant for  $V_{DS}$  larger than a couple of volts. However, because of  $I_{DSS}$  spread, the current is unpredictable. For example, the MMBF5484 (a typical *n*-channel JFET) has a specified  $I_{DSS}$  of 1–5 mA. Still, the circuit is attractive because of the simplicity of a two-terminal constant-current device. If that appeals to you, you're in luck. You can buy "current-regulator diodes" that are nothing more than JFETs with gate tied to source, sorted according to current. They're the current analog of a zener (voltage-regulator) diode. Here are the characteristics of the 1N5283–1N5314 series:<sup>21</sup>

<sup>20</sup> See Table 8.2 for the IF3601 and IF3602 (dual). Runners-up in the low-noise competition are the LSK170B and the BF862, with considerably lower capacitances.

<sup>21</sup> Said to be available from several manufacturers. Alternatives include the MSS283, MV5283, and MX5283 series from Microsemi; the SST502–SST511 and CR160–CR470 series from Vishay; and the J500–J511, J553–J557, and U553–U557 series from InterFET. Alternative sources: Central Semiconductor and Linear Integrated Systems.

<sup>19</sup> In the normal "quadratic" region of drain current, transconductance varies approximately as  $\sqrt{I_D}$ , see §3.3.3.

Characteristic	Value
Currents available	0.22–4.7 mA
Tolerance	$\pm 10\%$
Temperature coefficient	$\pm 0.4\%/\text{ }^\circ\text{C}$
Voltage range	1–2.5 V min, 100 V max
Current regulation	5% typical
Impedance	1M typ (for 1 mA device)

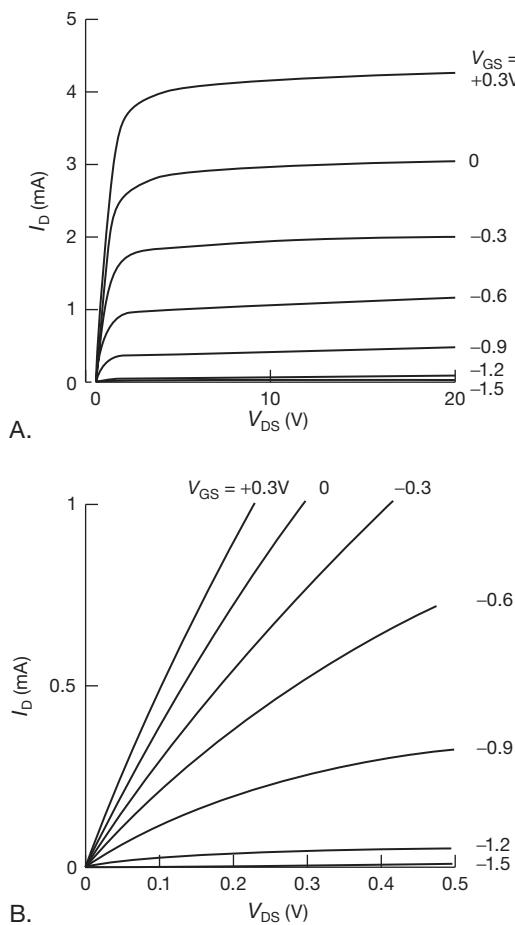


Figure 3.21. Measured JFET characteristic curves. 2N5484 n-channel JFET:  $I_D$  versus  $V_{DS}$  for various values of  $V_{GS}$ . See also Figure 3.47.

We measured  $I$  versus  $V$  for a 1N5294 (rated at 0.75 mA), applying 1 ms voltage pulses at 100 ms intervals to prevent heating. Figure 3.22A shows good constancy of current up to the breakdown voltage ( $\sim 145$  V for this particular specimen). You can see also the effect of heating when voltage is applied continuously in a dc measurement, caused by the negative temperature coefficient of drain current. Figure 3.22B shows that the device reaches full cur-

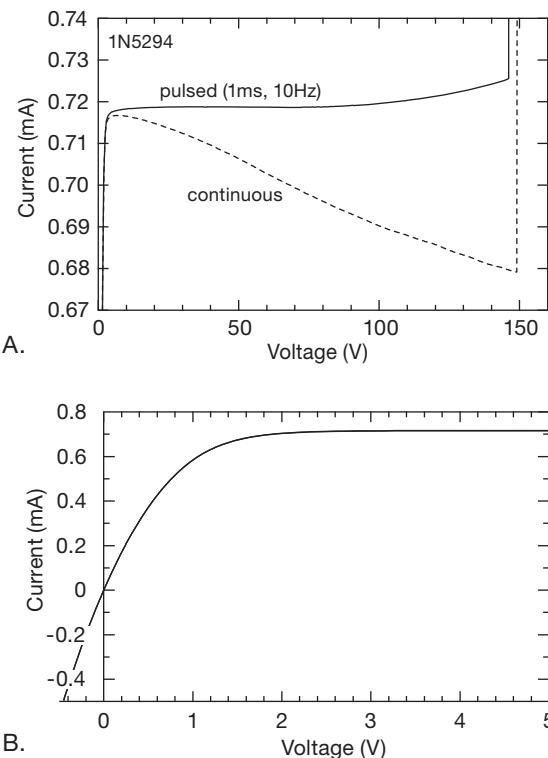


Figure 3.22. 1N5294 current regulator diode.

rent with somewhat less than 1.5 V across it (here both pulsed and dc curves are plotted, demonstrating negligible thermal effects with less than 0.4 mW dissipation). We'll show how to use these devices to make a cute triangle-wave generator in §7.1.3E. And we'll have much more to say about current sources in §4.2.5 and §9.3.14.

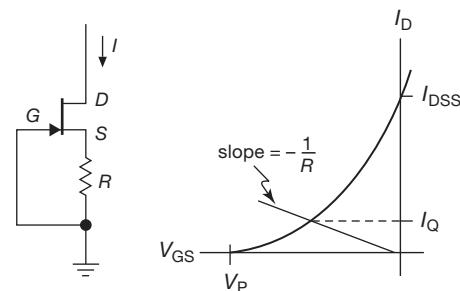


Figure 3.23. JFET current sink ( $I = V_{GS}/R$ ) for  $I_D < I_{DSS}$ .

#### A. Source self-biasing

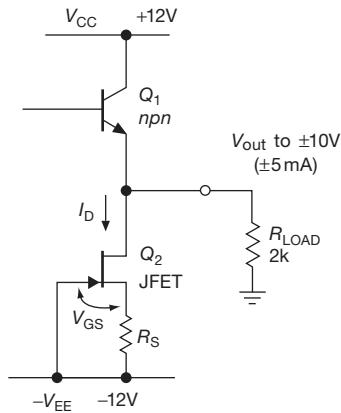
A variation of the previous circuit (Figure 3.23) gives you an adjustable current source. The self-biasing resistor  $R$  back-biases the gate by  $I_D R$ , reducing  $I_D$  and bringing the

JFET closer to pinch-off. You can estimate the value of  $R$  from the drain curves for the particular JFET. This circuit allows you to set the current (which must be less than  $I_{DSS}$ ), as well as to make it more predictable. Furthermore, the circuit is a better current source (higher impedance) because the source resistor provides “current-sensing feedback” (which we’ll learn about in §4.2.5A). (There’s a nice demonstration of this in the figure “Measured  $I_D$  versus  $V_{DS}$  for four JFETs” in §3x.4.3, where you will find drain-current versus drain-voltage curves both with and without a source self-biasing resistor.) Remember, though, that actual curves of  $I_D$  for some value of  $V_{GS}$  obtained with a real FET may differ markedly from the values read from a set of published curves, owing to manufacturing spread. (This is illustrated nicely by the examples of Figures 3.25 and 3.41 on pages 145 and 156, using actual measured drain characteristics from a batch of JFETs.) You may therefore want to use an adjustable source resistor, if it is important to have a specific current.

**Exercise 3.1.** Use the 2N5484 measured curves in Figure 3.21 to design a JFET current source to deliver 1 mA. Now ponder the fact that the specified  $I_{DSS}$  of a 2N5484 is 1 mA (min), 5 mA (max).

### B. Example: emitter follower pull-down

Let’s look at an example to explore further this problem of unpredictability of JFET zero-bias drain current,  $I_{DSS}$  (or, equivalently, the difficulty of predicting the gate-source bias needed to produce a desired drain current).



**Figure 3.24.** Design example: *npn* emitter follower with JFET current sink.

Figure 3.24 shows a BJT emitter follower, running between  $\pm 12$  V split supplies, with a current-sinking JFET pull-down to the negative rail. We specify that the circuit must be able to deliver a full swing of  $\pm 10$  V into a  $2\text{k}\Omega$

load (that’s  $\pm 5$  mA load current). You might at first think of using a simple pull-down resistor  $R_E$  to the  $-12$  V rail. But the output swing requirement makes things difficult, because you would need to keep  $R_E$  less than  $400\Omega$  (we might choose  $365\Omega$ , a standard 1% value) to get a full negative swing; and that low resistance would produce a relatively high quiescent current (at 0 V output) of 33 mA (thus  $\sim 400$  mW quiescent dissipation both in  $Q_1$  and in  $R_E$ ), compared with the 5 mA peak current delivered to the load (recall the discussion in §2.4.1). Worse still, a resistive pulldown also greatly degrades linearity, owing to variations in the follower’s  $r_e$  caused by the large variation of collector current (65 mA at the top of the swing, dropping to 0.5 mA at the bottom, thus corresponding  $r_e$  of  $0.4\Omega$  and  $50\Omega$  into a combined load resistance of  $\sim 300\Omega$ ). Finally, the small resistive pull-down resistor (compared with the minimum load resistance) undesirably reduces the circuit’s input impedance by a factor of 6.

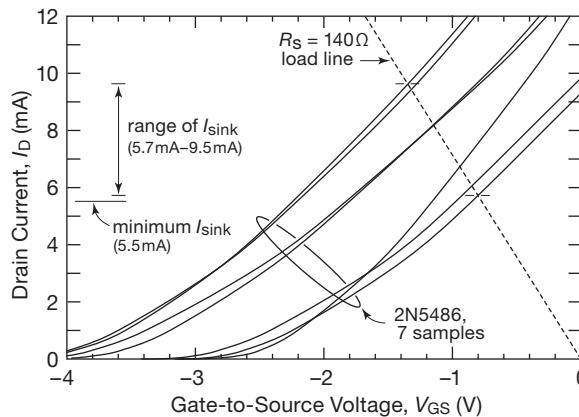
So an active current sink is the way to go. A first possibility is to choose a JFET whose minimum specified  $I_{DSS}$  is at least equal to our required 5.5 mA. Only the 2N5486 member of the 2N5484–86 family satisfies this requirement ( $8\text{mA} \leq I_{DSS} \leq 20\text{mA}$ , see Table 3.1 on page 141). But these currents are rather more than we’d like, and a part with  $I_{DSS} = 20$  mA produces too much heat: the worst-case dissipation is 440 mW in either the JFET (at peak positive swing) or the BJT (at peak negative swing with no load), too much for a transistor in a TO-92 or SOT-23 package without a heatsink.

So, let’s add a source resistor so we can tailor the JFET’s drain current; we’ll aim for a 5.5 mA minimum sink current, so that we retain 0.5 mA reserve at full negative swing. The 2N5486’s minimum  $I_{DSS}$  of 8 mA guarantees that a source self-biased circuit can sink the required 5.5 mA current. Now we need only choose the source resistor  $R_S$ .

The problem is that datasheet curves of  $I_D$  versus  $V_{GS}$  (called “transfer characteristics”), when provided at all, do not show the full range of possibilities; instead they show curves typical of parts with two or three selected values of  $I_{DSS}$  within the allowable range. And sometimes all you get are tabulated limits for  $I_{DSS}$  and for  $V_{GS(\text{off})}$ .<sup>22</sup> But you can measure some JFETs to get a sense of things. We did this, and Figure 3.25 shows measured  $I_D$  versus  $V_{GS}$  curves for seven 2N5486s from different manufacturers

<sup>22</sup> It is possible to extrapolate from published (or measured) curves by estimating  $k$  and  $V_{th}$  for the simple quadratic law  $I_D = k(V_{GS} - V_{th})^2$ . See the discussion in Chapter 3x.

and batches.<sup>23</sup> Assuming this represents the full range of variability (it doesn't, quite, as seen from the minimum  $I_{DSS}$  of 9.2 mA) then we can swing a load line up from the origin until the lowest intersection is above  $I_D = 5.5$  mA. That's an  $R_S$  of  $140\Omega$  (shown), for which the range of drain current is 5.7 mA (minimum) to 9.5 mA (maximum).



**Figure 3.25.** Choosing a source resistor  $R_S$  to bias a JFET current sink to produce  $I_{sink} \geq 5.5$  mA.

The good news is that the circuit will work; the bad news is that the range of sink currents is nearly 2:1 (taking into account the possibility of production parts whose curves span a somewhat wider range than seen in these seven parts). But the good news, again, is that even for a JFET at the high end of the range (thus  $I_{sink} \approx 10$  mA), the follower's worst-case dissipation is limited to 220 mW (at peak negative swing, no load), and the JFET's worst-case dissipation is likewise limited to 220 mW (at peak positive swing). This is well within allowable dissipation for a TO-92 transistor (350 mW at 25°C ambient).

### C. Current sinks for JFET amplifiers

Stepping back a bit, one might ask whether a JFET current sink, with its 2:1 spread of quiescent current, was a good choice. True, it works. But you can do better with a simple BJT current sink, five versions of which are shown in Figure 3.26. These use more parts, but sink a predictable current. And if you *really* care about minimizing parts count, then you can always use the alternative of a JFET selected for a narrow range of  $I_{DSS}$ , with no self-biasing resistor,

i.e., Figure 3.24 with  $R_S = 0$  (the 2N5485 specifies  $I_{DSS}$  as 4 mA–10 mA; you might select parts from 5.5–8 mA).<sup>24</sup>

This example illustrates the down side of the loose drain current (and corresponding gate voltage) specifications characteristic of all JFETs. As attractive as it may seem to drop in a JFET when you need a current source, it's problematical. But JFETs come into their own when you need an amplifier with high input impedance and low noise – although the loosey-goosey specs are still challenging, the results are worth the bother. We'll see examples presently.

### D. Imperfect current source

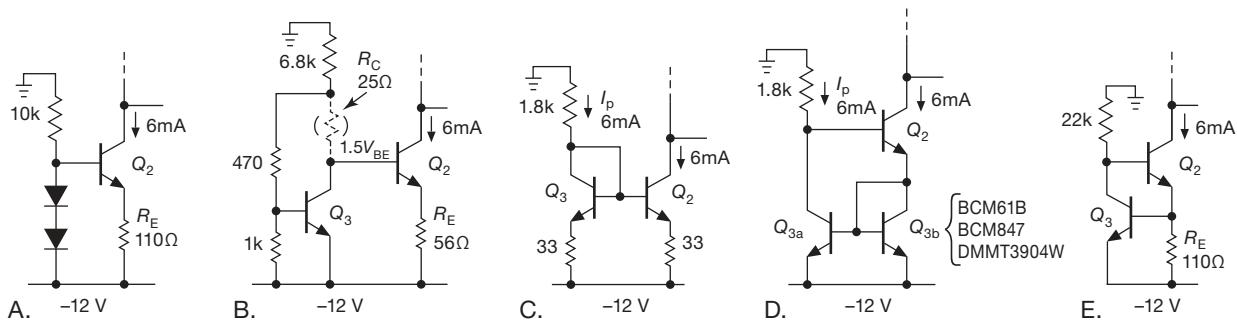
A JFET current source, even if built with a source resistor, shows some variation of output current with output voltage; i.e., it has finite output impedance, rather than the desirable infinite  $Z_{out}$ .<sup>25</sup> The measured curves of Figure 3.21, for example, suggest that, over a drain voltage range of 5–20 V, a 2N5484 shows a drain current variation of 5% when operated with gate tied to source (i.e.,  $I_{DSS}$ ). This might drop to 2% or so if you use a source resistor. An elegant solution is the use of a cascode transistor to suppress drain voltage variations in the current-setting transistor. This can be used both for BJT current sources (it's shown in §2x.3) and for JFET current sources, as shown in Figure 3.27. The idea (as with BJTs) is to use a second JFET to hold constant the drain-source voltage of the current source.  $Q_1$  is an ordinary JFET current source, shown in this case with a source resistor.  $Q_2$  is a JFET of larger  $I_{DSS}$ , connected "in series" with the current source. It passes  $Q_1$ 's (constant) drain current through to the load, while holding  $Q_1$ 's drain at a fixed voltage – namely the gate-source voltage that makes  $Q_2$  operate at the same current as  $Q_1$ . Thus  $Q_2$  shields  $Q_1$  from voltage swings at its output; since  $Q_1$  doesn't see drain voltage variations, it just sits there and provides constant current. If you look back at the Wilson mirror (Figures 2.61, 3.26D), you'll see that it uses this same voltage-clamping idea.

You may recognize this JFET circuit as the "cascode," which is normally used to circumvent Miller effect (§2.4.5). A JFET cascode is simpler than a BJT cascode,

<sup>24</sup> An expensive alternative is to use a pre-sorted two-terminal "current regulator diode" like those in the footnote 3.2.2 on page 142. These appear to be a vanishing species, and the range of currents is quite limited. (A double complaint, reminiscent of the dialog "The food there is so bad." "Yes, and such small portions.")

<sup>25</sup> This is important also for JFET amplifiers (§§3.2.3A and 3.3.2). For further details look at the discussion in Chapter 3x.

<sup>23</sup> Making a mere cameo appearance here, compared with their full performance in Figures 3.55 and 3.56 and associated discussion).



**Figure 3.26.** Alternatives to the JFET pulldown of Figure 3.24. A. Classic BJT current sink, base biased to  $\sim 2V_{BE}$ ; you can substitute a red LED for the diode pair. B.  $Q_3$  creates a  $1.5V_{BE}$  base bias for current sink  $Q_2$ ; adding optional resistor  $R_C$ , chosen equal to  $Q_3$ 's  $r_e$ , cleverly compensates for the latter's change in  $V_{BE}$  with supply-voltage variations (i.e.,  $I_C$  variations through the 5.6k resistor). This configuration is useful if the current-sink output must operate very close to the negative rail – down to a few hundred mV, say, if configured as a  $1.25V_{BE}$  bias. C. Current mirror with  $\approx 200$  mV emitter ballasting (needed to equalize collector currents in view of  $V_{BE}$  mismatch, and to suppress the Early effect output-current variations). D. Wilson mirror with matched pair; no emitter resistors needed. E. “Ring-of-two” current source. See Figure 2.32 for other current source circuits.

however, because you don't need a bias voltage for the gate of the upper FET: because it's depletion mode, you can simply connect the upper gate to the lower source terminal (compare with Figure 2.84);  $Q_2$ 's gate-to-source voltage at the operating current (set by  $Q_1$  with its  $R_S$ ) then sets  $Q_1$ 's drain-to-source operating voltage:  $V_{DS1} = -V_{GS2}$ . A nice additional benefit is that the resultant circuit is a *two-terminal* current source.

It is important to realize that a good bipolar transistor current source will give far better predictability and stability than a JFET current source. Furthermore, the op-amp-assisted current sources we'll see in the next chapter are better still. For example, a FET current source might vary 5% over a typical temperature range and load-voltage variation, even after being set to the desired current by trimming the source resistor, whereas an op-amp/transistor (or op-amp/FET) current source is predictable and stable to better than 0.5% without great effort.

### 3.2.3 FET amplifiers

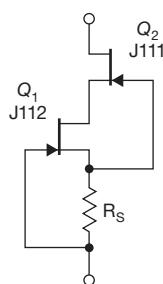
Source followers and common-source FET amplifiers are analogous to the emitter followers and common-emitter amplifiers made with bipolar transistors that we talked about in the previous chapter. However, the absence of dc gate current makes it possible to realize very high input impedances. Such amplifiers are essential when dealing with the high-impedance signal sources encountered in measurement and instrumentation. For some specialized applications you may want to build followers or amplifiers with discrete FETs; most of the time, however, you can take advantage of FET-input op-amps. In either case it's worth knowing how they work.

With JFETs it is convenient to use the same self-biasing scheme as with JFET current sources (§3.2.2), with a single gate-biasing resistor to ground (Figure 3.28); MOSFETs require a divider from the drain supply, or split supplies, just as we used with BJTs. The gate-biasing resistors can be quite large (a megohm or more), because the gate leakage current is measured in picoamps to nanoamps.

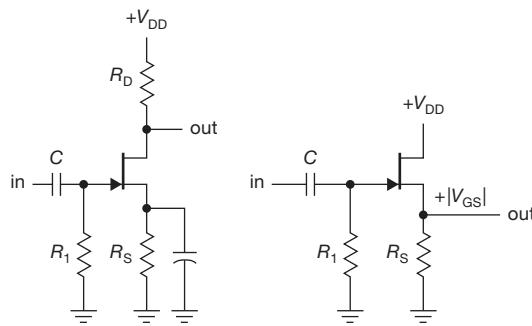
#### A. Transconductance

The absence of gate current makes *transconductance* (the ratio of output current to input voltage:  $g_m = i_{out}/v_{in}$ ) the natural gain parameter for FETs. This is in contrast to bipolar transistors in the last chapter, where we at first flirted with the idea of current gain, or beta ( $i_{out}/i_{in}$ ), then introduced the transconductance-oriented Ebers–Moll model: it's useful to think of BJTs either way, depending on the application.

FET transconductance can be estimated from the



**Figure 3.27.** Cascode JFET current sink.



**Figure 3.28.** Common-source amplifier and source follower. For both configurations the source voltage is above ground, because of the source current flowing through  $R_S$ , with a quiescent point  $V_S = V_{GS} = R_S I_D(V_{GS})$ .

characteristic curves, either by looking at the increase in  $I_D$  from one gate-voltage curve to the next on the family of curves (Figures 3.2 or 3.21), or, more simply, from the slope of the  $I_D$  versus  $V_{GS}$  transfer characteristics curve (Figures 3.15 or 3.51). The transconductance depends on drain current (we'll see how, shortly) and is, of course,

$$g_m(I_D) = i_d/v_{gs}.$$

(Remember that lowercase letters indicate quantities that are small-signal variations.) From this we get the voltage gain,

$$G_{\text{voltage}} = v_d/v_{gs} = -R_D i_d/v_{gs}$$

i.e.,

$$G = -g_m R_D, \quad (3.3)$$

just the same as the bipolar-transistor result in §2.2.9, with load resistor  $R_C$  replaced with  $R_D$ . Typically, small-signal FETs have transconductances in the neighborhood of 10 millisiemens<sup>26</sup> (mS) at a few millamps.<sup>27</sup> Because  $g_m$  depends on drain current, there will be some variation of gain (nonlinearity) over the waveform as the drain current varies, just as we have with grounded-emitter amplifiers (where  $g_m = 1/r_e$ , proportional to  $I_C$ ).

In the following discussion we'll be using the concept of FET gate drive,  $V_{GS} - V_{th}$ . Recall that  $V_{th}$  is the extrapolated gate threshold voltage we discussed in §§3.1.3 and 3.1.4.

The variation of  $g_m$  with drain current is easy to calculate and highly useful when designing JFET followers and amplifiers. For operation above subthreshold ( $I_D >$

<sup>26</sup> Formerly millimhos, or mΩ.

<sup>27</sup> This is substantially less than that of a BJT at the same current; the latter has  $g_m = 40$  mS at 1 mA, and 200 mS at 5 mA, for example. There's further discussion in §3.3.3 and §3x.2.

$I_{DSS}/25$ , say), we've seen that the drain current is quadratic in the gate drive

$$I_D = \kappa(V_{GS} - V_{th})^2, \quad (3.4)$$

from which the transconductance ( $g_m = i_d/v_{gs} = \partial I_D / \partial V_{GS}$ ) is seen to be

$$g_m = 2\kappa(V_{GS} - V_{th}) = 2\sqrt{\kappa I_D}. \quad (3.5)$$

In other words, in the “quadratic region” of drain current,  $g_m$  is proportional to gate drive, increasing approximately linearly from pinchoff to its specified value at  $I_{DSS}$ ; alternatively, you can say that it is proportional to the square root of drain current.<sup>28</sup> This is a helpful rule, particularly because the datasheets specify  $g_m$  only at its maximum value, at  $I_{DSS}$ ; we'll use it shortly.<sup>29</sup>

As an example, if (as is often the case) you're operating a JFET in its quadratic region and you want to estimate the transconductance at some drain current  $I_D$ , then if you know  $g_m$  at some other drain current  $I_{D0}$  (which may be  $I_{DSS}$ ) you can exploit the square root dependence on drain current in eq'n 3.5 to find, simply

$$g_m/g_{m0} = (I_D/I_{D0})^{1/2}. \quad (3.6)$$

FETs in general have considerably lower transconductance than bipolar transistors,<sup>30</sup> which makes them less impressive as amplifiers and followers; we treat this in more detail in §3x.2. However, their outstanding characteristic of extremely low input (gate) current, often of order a picoampere or less, makes it worthwhile to develop circuit solutions that circumvent the problems of low gain (e.g., current source as drain load), or that enhance their effective transconductance (“transconductance enhancer”).

At this point it's helpful to see some JFET amplifier examples.

<sup>28</sup> Be careful about signs: in these equations  $V_{th}$  and  $V_{GS}$  are negative (for n-channel JFETs), but  $V_{th}$  is more negative, thus a positive value for  $g_m$ . As long as you respect signs, these expressions work for n-channel or p-channel and for enhancement or depletion modes. Note that the value  $\kappa$  is not given on datasheets, but can be determined empirically for a given part type and manufacturer. Generally speaking, within a given batch or type of JFET you'll find variations in  $V_{th}$ , with  $\kappa$  being relatively constant. Thus a measurement of  $I_{DSS}$  and  $V_{th}$  allows you to calculate  $\kappa$  from eq'n 3.4, under the assumption that the quadratic region of drain current extends all the way to  $I_{DSS}$  (it usually does).

<sup>29</sup> See §3x.2 for further discussion of transconductance versus drain current.

<sup>30</sup> Except in the low drain-current (“subthreshold”) region; see Figure 3.54 and analogous figures in §3x.2.

## B. JFET amplifier configurations

Figure 3.29 shows the basic configurations for a JFET common-source amplifier stage. In circuit A the JFET is running at its  $I_{DSS}$ , with  $R_D$  sized small enough so that the drain is at least a volt or two above ground for the maximum specified  $I_{DSS}$ . (This is often an annoying constraint, given the loose ratio of specified  $I_{DSS(\max)}/I_{DSS(\min)}$  – commonly 5:1 for most JFETs, see Tables 3.1 on page 141, 8.2 on page 516, and 3.7 on page 217; presently we'll see ways to handle this awkward situation.) The resistor across the input can be very large – 100 MΩ or more – with an input blocking capacitor (for an ac-coupled amplifier); or it can be omitted altogether for a dc-coupled signal near ground. For this circuit the ideal voltage gain is  $G=g_m R_D$ , where  $g_m$  is the transconductance at the operating drain current; it is analogous to the BJT grounded-emitter amplifier of Figure 2.44.<sup>31</sup>

To illustrate actual component values and performance we've chosen the exemplary BF862, because of its high transconductance (45 mS typical at  $I_{DSS}$ ) and tight  $I_{DSS}$  spec (10–25 mA); it happens also to be a low-noise part, as we'll see in Chapter 8. The drain resistor  $R_D$  is sized to maintain a minimum 2.5 V across  $Q_1$  (for specified  $I_{DSS(\max)}$ ); the typical voltage gain is then  $G=-g_m R_D \approx -13$  (inverting).

By adding a source resistor  $R_S$  circuit B lets you run at a drain current less than  $I_{DSS}$ , in the manner of Figures 3.23 and 3.25. But the source degeneration reduces the gain, to  $G=-R_D/(R_S+1/g_m)$ . This is analogous to the degenerated BJT common-emitter amplifier of Figure 2.49 (but with simpler self-biasing because the gate-source junction is back-biased), with  $1/g_m$  replacing  $r_e$  (you can think of  $1/g_m$  as an “intrinsic source impedance” of the JFET, analogous to the intrinsic emitter resistance  $r_e$  of the BJT).<sup>32</sup> Illustrating with the same BF862, we aim for a drain current of 2 mA by choosing a source self-bias resistor  $R_S=200\Omega$ , estimating that ~0.4 V of gate back-bias is about right.<sup>33</sup> Estimating that  $g_m \approx 20 \text{ mS}$  at this drain current,<sup>34</sup> we ar-

rive at an estimated voltage gain (from the above equation) of  $G \approx -8$ .

Circuit C bypasses the source resistor at signal frequencies, so you can run at the same dc drain current as circuit B, but with the higher gain as in circuit A (where  $g_m$  is the transconductance at the actual drain current, here reduced from the  $I_{DSS}$  of circuit A); it's analogous to the BJT circuit of Figure 2.48. You can throttle back the gain by adding a gain-setting resistor  $R_{S'}$  in series with the capacitor (circuit D), for a signal-frequency gain of  $G=-R_D/(R_S||R_{S'}+1/g_m)$ ; this is analogous to the BJT circuit of Figure 2.50. Or you can step on the throttle by adding a second stage of voltage gain, as in circuit E, with the second stage common-emitter amplifier multiplying the first-stage gain of any of the previous single-stage circuits by a factor  $R_C I_C / V_T$  (i.e.,  $R_C / r_e$ ), where as always  $V_T = kT/q \approx 25 \text{ mV}$ . This approximation assumes  $Q_2$  is driven by a voltage source, i.e.,  $R_D \ll \beta r_e$ ; most of the gain of the combination comes from the BJT, with its high transconductance. The similar-looking circuit F creates a three-terminal hybrid creature, with the BJT's  $g_m$  contributing to achieve a high effective overall transconductance; in this configuration the BJT is a “transconductance enhancer.” This configuration is closely analogous to the BJT complementary Darlington (Sziklai) of Figure 2.77, and is treated in more detail in §3x.2.

## C. Adding a cascode

The last four circuits show how to implement a drain-clamping cascode to the common-source stage. This elegant configuration is usually featured as a way to circumvent the Miller effect (the effective multiplication of drain-gate capacitance by the stage's voltage gain); that's the way it was presented in Figure 2.84, where it made its first appearance. Here it indeed accomplishes that (it's a “Miller killer”), which is helpful in keeping the input impedance high. But it's better than that: (a) it also lets you keep the drain-to-source voltage low (avoiding the precipitous rise in “impact-ionization” gate current, §3.2.8); and (b) by clamping the drain-source voltage it circumvents the “ $g_{os}$  effect” (finite output impedance  $r_o$ , caused by  $I_D$  dependence on  $V_{DS}$ ); the voltage gain is then not degraded, and is simply  $G=g_m R_D$ . This latter benefit is reminiscent of the use of a cascode transistor in the Wilson current mirror (Figure 2.61); it's an “Early killer.” There's more detail in Chapter 3x (§3x.4), including experimental results from four JFET exemplars whose  $g_m$  and  $g_{os}$  were measured; then their gain predictions were compared with measured amplifier gains (both with and without cascode).

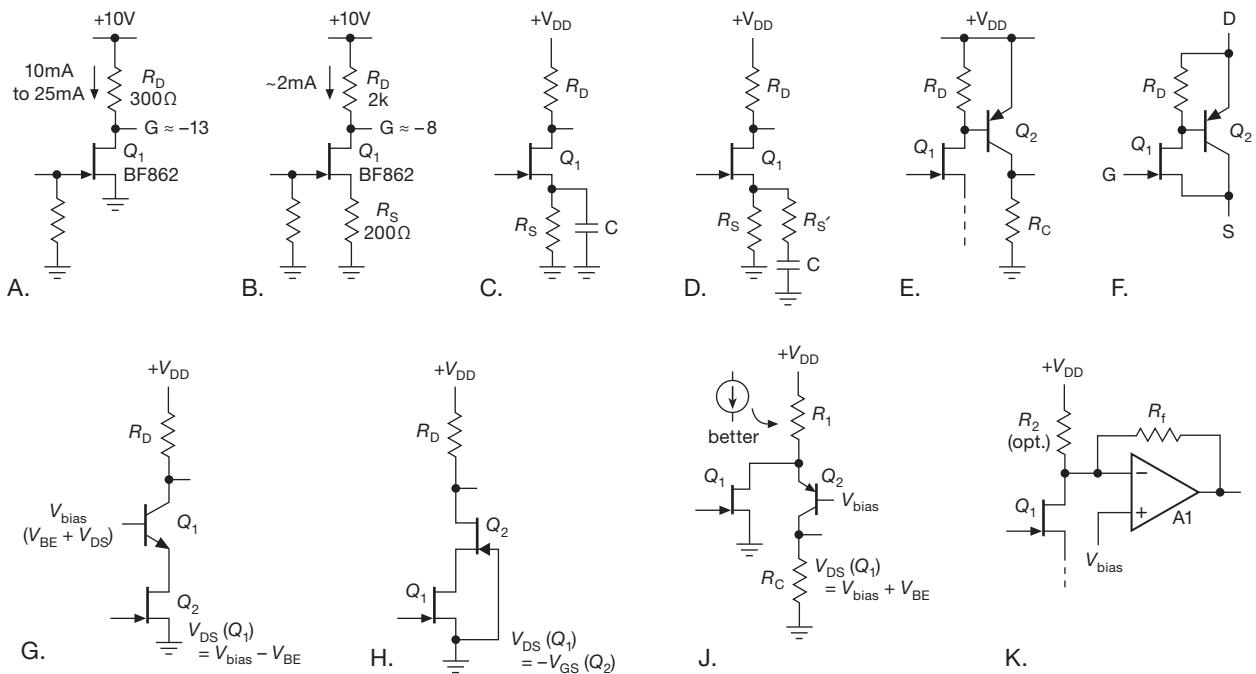
In circuit G the BJT's base bias sets the JFET's

<sup>31</sup> However, owing to the finite output impedance of the JFET (called  $r_o$ , or  $1/g_{os}$ ), the drain load resistor effectively sees a parallel resistance of  $r_o$ , so the gain is reduced to  $G=g_m(R_D||r_o)$ ; this has a negligible effect for the component values here. This is analogous to the Early effect in BJTs, and becomes important for large values of  $R_D$ , or (especially) when  $R_D$  is replaced by a current source. Lots of discussion in §3x.4.

<sup>32</sup> This time we're ignoring the JFET's finite output impedance; see §3x.4 if you're curious why and for much more detail (if it's detail you really want . . . be careful what you wish for).

<sup>33</sup> Our confidence is bolstered, a bit, by having measured a sample's  $I_D$  versus  $V_{GS}$ .

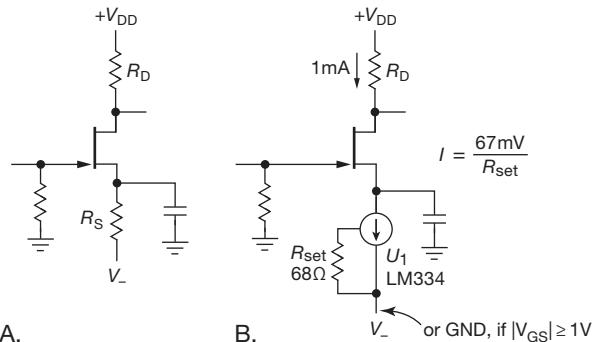
<sup>34</sup> We measured this, too.



**Figure 3.29.** Simple single-supply common-source JFET amplifiers.

drain-to-source operating voltage. It's simpler to use a second JFET ( $Q_2$  in circuit H), which must be chosen to have a larger  $V_{GS}$  back-bias than  $Q_1$  at the same drain current, though the generally loose gate voltage specifications make this an uncertain proposition. Circuit J is an “inverted cascode,” in which  $Q_1$ 's drain current variations divert current from  $Q_2$ ; it's a helpful circuit to know about when you find yourself bumping into the positive rail with a conventional cascode. Finally, in circuit K an op-amp transimpedance stage (current-to-voltage; see §4.3.1) substitutes for the drain-clamping cascode transistor  $Q_2$ : feedback through  $R_f$  maintains its inverting input (the “ $-$ ” input) at the bias voltage while producing an output voltage  $V_o = -I_D R_f + V_{bias}$ ; the optional resistor  $R_2$  lets you add an offset to reposition the quiescent output voltage according to your whim.

The circuits in Figure 3.29 operate from a single positive supply voltage; they're simple, but, given the characteristically loose  $I_{DSS}$  and  $V_{GS}$  specifications of JFETs, they suffer from significant uncertainty of operating current. If you have available a negative supply voltage as well, there are several ways to rig things up to ensure predictable biasing. Look at Figure 3.30A, where the operating current of the  $n$ -channel JFET is set by the source pulldown resistor,  $I_D = -(V_- + V_{GS})/R_S$ , or approximately  $V_-/R_S$  for negative supply voltages large compared with



**Figure 3.30.** A negative supply rail allows predictable source-pulldown biasing of JFET common-source amplifiers.

the JFET's gate-source voltage. As in Figure 3.29C, the bypass capacitor lets signal frequencies partake of the full JFET gain, i.e.,  $G_V = -g_m R_D$ , where  $g_m$  is the transconductance at the operating current. A more elegant solution is the use of a current-sink pulldown, as in Figure 3.30B. The LM334 is an inexpensive ( $\sim \$0.50$ ) resistor-programmable current source (see §9.3.14B), here configured for 1 mA ( $I \approx 0.067/R_{set}$ ). With this circuit there's no uncertainty about the operating current (it does not depend on  $V_{GS}$ ); better still, the LM334 operates down to 1 V drop, so you can operate with a single positive supply if the JFET's

minimum specified gate-source voltage at the programmed current is at least a volt.<sup>35</sup>

#### D. Series-feedback (“current-feedback”) pair

The JFET amplifiers just illustrated have admirably high input impedance, but they suffer from rather low (and not terribly predictable) gain. Bipolar transistors give you predictable gain, and lots of it; but you pay the price in terms of input current. You can have the best of both worlds, though, by combining a JFET front-end of modest (and not terribly predictable) gain with some serious second-stage gain. That way you get the ultra-low input current (high input impedance) of a JFET, but with enough overall open-loop circuit gain so that negative feedback can close the loop to produce predictable gain.

The next circuit is a low-power ( $660\ \mu\text{A}$ ) battery-operated amplifier. We'll explore it in more detail than usual, introducing some new concepts along the way. Figure 3.31 shows the first of several JFET amplifier examples that exploit the JFET's ultra-low input current, combined with an additional gain stage (and feedback) to achieve predictable and stable voltage gain. It is similar to the bipolar series-feedback pair illustrated in Figure 2.92:  $Q_1$  is a common-source amplifier, with BJT  $Q_2$  providing second-stage voltage gain to the output (via follower  $Q_3$ , whose base-emitter drop across  $R_3$  sets  $Q_2$ 's collector current). That provides the needed voltage gain (which the JFET's low  $g_m$  is unable to provide). Negative feedback closes the loop via voltage divider  $R_6$  and  $R_5||R_1$  (at signal frequencies) and biasing via  $R_6$  and  $R_1$  at dc. This configuration is variously known as “series feedback” or “current feedback.”

The spread of specified  $I_{DSS}$  (or, equivalently, of  $V_{GS(\text{off})}$ ) creates a problem in any JFET design. To deal with that, we choose a JFET with a tight  $V_{GS(\text{off})}$  specification ( $-1.2$  to  $-2.7$  V), and we run it at a drain current well below  $I_{DSS}$  (10 mA minimum) so that the gate-source voltage is close to  $V_{GS(\text{off})}$ . The feedback path sets the signal gain. With some careful thought (and some juggling and iteration), the same feedback path can be made to establish the (dc-coupled) bias condition.

Here's how it goes. The gate is at ground; we start by assuming a source voltage of approximately 1.7 V, and choose  $R_1$  for  $500\ \mu\text{A}$ . Of that, about  $300\ \mu\text{A}$  comes from the JFET's drain current ( $Q_2$ 's  $V_{BE}$  across  $R_2$ ); thus

<sup>35</sup> The effective capacitance of an LM334 current source/sink is 10 pF, small enough to ignore for most purposes. We calculated this from the slew-rate plot in the datasheet. TI's application note LB-41 has additional useful information about the LM334.

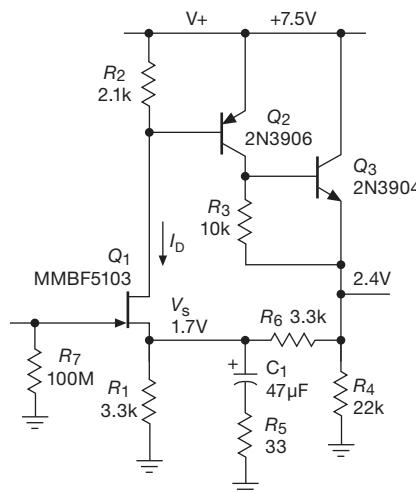


Figure 3.31. JFET–BJT series-feedback pair.

$\sim 200\ \mu\text{A}$  is coming from  $R_6$ . This puts the output at about +2.4 V and sets  $Q_3$ 's emitter current at about  $300\ \mu\text{A}$  ( $110\ \mu\text{A}$  through  $R_4$ , plus  $200\ \mu\text{A}$  through  $R_6$ , minus  $60\ \mu\text{A}$  from  $R_3$ ).

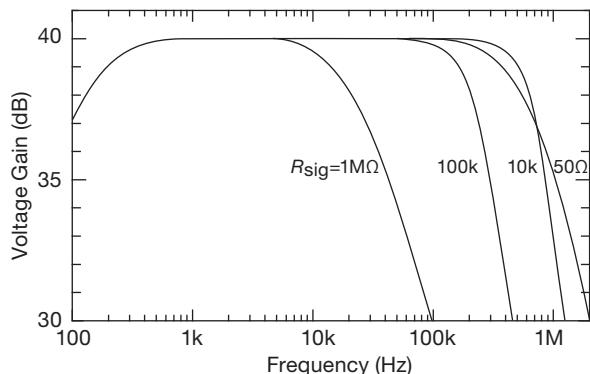
That's the self-consistent situation, under the assumed  $V_{GS}$  of  $-1.7$  V. For a different  $V_{GS}$ , the dc output voltage would change accordingly; it could range from +1.3 to +4 V over the extreme specified range of gate threshold voltage. This would degrade the maximum possible output swing, but that would usually be OK for an amplifier handling small signals (if not,  $R_1$  could be selected for different  $V_{GS}$  ranges of JFET parts).<sup>36</sup>

The nominal gain at signal frequencies is approximately 100, set by  $R_5$  (blocked by  $C_1$ ):  $G=1+R_6/(R_5||R_1)$ . The low-frequency  $-3\text{ dB}$  point is at 100 Hz (where the reactance of  $C_1$  equals  $R_5$ ). The high-frequency  $-3\text{ dB}$  point is not so easily calculated, but a SPICE model puts it at approximately 800 kHz (it measured 720 kHz in our breadboard, where there are some added parasitic capacitances). The latter is due primarily to the  $RC$  rolloff of  $Q_1$ 's output signal impedance of  $2.1\text{ k}\Omega$  (i.e.,  $R_2$ ) driving  $Q_2$ 's input capacitance of  $\sim 4\text{ pF}$ , with the latter greatly magnified by the Miller effect.

For large signal driving impedances  $R_{\text{sig}}$  the amplifier's

<sup>36</sup> A better way to ensure proper biasing is to replace  $R_1$  with a  $0.5\text{ mA}$  current sink. A JFET comes to mind (we've got JFETs on the mind!), but, given their unpredictable dc characteristics, a much better choice would be one of the BJT current sinks illustrated in Figure 3.26. Another way of handling this thorny problem is to use a slow feedback loop to stabilize  $I_D$  to a desired value less than the specified minimum  $I_{DSS}$ .

bandwidth is reduced,<sup>37</sup> owing to an input capacitance of  $\sim 5 \text{ pF}$ ; see Figure 3.32. This is due primarily to the JFET's drain-to-gate capacitance (plus wiring capacitance), given that the source terminal is bootstrapped by feedback. There are numerous tricks to deal with this effect (if more bandwidth is desired), including a cascode in the JFET's drain (which can be bootstrapped to further suppress its input capacitance) and a cascode in the  $Q_2$  BJT gain stage. Some of these techniques are discussed in Chapter 3x.



**Figure 3.32.** Measured gain versus frequency for the amplifier of Figure 3.31. The  $f_{3\text{dB}}$  with  $R_{\text{sig}}=1 \text{ M}\Omega$  shows that  $C_{\text{in}}=7 \text{ pF}$ .

#### Design equations and design hints

Collecting it together in one place:

$$G = 1 + \frac{R_6}{R_5 \parallel R_1} \approx 1 + \frac{R_6}{R_5} \quad (\text{ac gain}),$$

$$G_{\text{OL}} = g_{m1} R_2 g_{m2} R_3 g_{m3} (R_4 \parallel R_6), \quad (\text{open-loop gain}^{38})$$

$$I_D = \frac{V_{\text{BE}2}}{R_2} \approx \frac{0.7}{R_2} \quad (\text{JFET bias}),$$

$$I_{C2} = \frac{V_{\text{BE}3}}{R_3} \approx \frac{0.65}{R_3} \quad (Q_2 \text{ bias}),$$

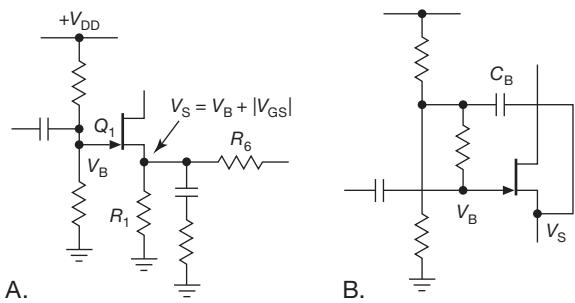
$$V_{\text{out}} = V_S \left( 1 - \frac{R_6}{R_1} \right) + \frac{R_6}{R_2} V_{\text{BE}2} + \frac{R_6}{R_1} |V_{\text{EE}}| \quad (\text{output bias}).$$

<sup>37</sup> Though paradoxically for  $R_{\text{sig}}$  values of a few  $\text{k}\Omega$  it is *extended* somewhat, owing to some response “peaking.”

<sup>38</sup> This expression overestimates the open-loop gain by neglecting the gain-limiting Early effect in  $Q_2$ , the stage where most of the circuit's overall gain resides. The measured 2N3906 Early voltage  $V_A \approx 25 \text{ V}$  (§2x.8) implies a maximum voltage gain of the  $Q_2 Q_3$  stage of  $\sim 1000$  (compared with its ideal  $G \approx 2500$ ), thus an overall open-loop gain of  $\sim 5000$ . This is ample for the modest  $\times 100$  closed-loop gain.

For single-supply operation the last term is zero. Basically,  $R_2$  sets  $I_D$ , and the ratio  $R_6/R_1$  sets  $V_{\text{out}}$ . For single-supply operation ( $V_{\text{EE}}=0$ ) use a small value for  $R_6$  if the JFET has substantial  $V_{\text{GS}}$  at its operating current and larger values for  $R_6$  for lower  $V_{\text{GS}}$  parts. The latter is tricky, because the “leverage” of  $R_6/R_1$  can push  $V_{\text{out}}$  all over the map. Choose  $R_4$  to help set  $I_{C3}$  after you've dealt with  $V_{\text{out}}$ . It may be necessary to select  $R_1$  to go with batches of parts with similar  $V_{\text{GS}}$ . A negative  $V_{\text{EE}}$  supply helps with biasing and also permits output swings both sides of ground.

Another way to handle the uncertainty in  $V_{\text{out}}$  is to apply a positive bias to the gate, as in Figure 3.33A. This adds a positive offset of  $V_B$  at the source terminal (whose voltage is now  $V_S = V_B - V_{\text{GS}}$ , where  $V_{\text{GS}}$  is negative for an  $n$ -channel JFET), making  $V_{\text{GS}}$  less important, and thus a smaller fractional uncertainty in  $V_S$ . While you're at it, you can easily bootstrap the gate bias divider,<sup>39</sup> as in Figure 3.33B, to raise the input impedance.



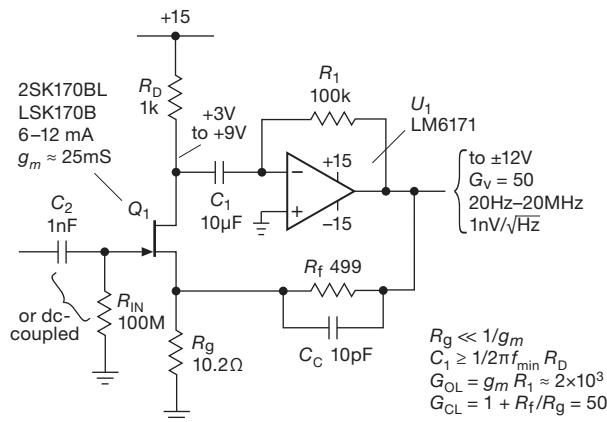
**Figure 3.33.** A. A positive bias at the gate of  $Q_1$  in Figure 3.31 improves the predictability of  $V_{\text{out}}$ . B. Add a bootstrap to raise  $R_{\text{in}}$ .

#### E. Simple “hybrid” JFET amplifier

With the assistance of an op-amp (the magnificent centerpiece of Chapter 4) you can do wonders. Put simply, an op-amp is a “very high-gain difference amplifier in a bottle,” intended to be fodder for feedback as the universal core of pretty much any analog circuit. It’s *pure engine*: a turbocharged Harley unicycle with dual intakes. This example and the next show a couple of ways to use an op-amp’s properties in support of a JFET amplifier. Look first at Figure 3.34. Here we’ve chosen the excellent 2SK170B (with LSK170B as a second source) for the front-end: it has lots of transconductance (about 25 mS at its  $I_{\text{DSS}}$  of 6–12 mA), along with very low noise voltage ( $\sim 1 \text{ nV}/\sqrt{\text{Hz}}$ ). We run it at zero gate voltage; and we deal with the 2:1

<sup>39</sup> You can, of course, bootstrap the gate resistor even when the gate is biased at ground, via  $R_7$  in Figure 3.31.

spread of specified  $I_{DSS}$  (a tighter specification than provided for most JFETs) by choosing the drain load resistor  $R_D$  small enough to avoid dc saturation even at  $I_{DSS}$  (max). The actual drain voltage is unimportant because we use ac coupling to the second stage (via  $C_1$ ). Ignoring for the moment the second stage (and setting  $R_g = 0$ ), the front-end voltage gain would be  $G = g_m R_D$ , or roughly  $G \approx 25$ , with perhaps  $\pm 25\%$  uncertainty from JFET manufacturing process variations.



**Figure 3.34.** Hybrid JFET amplifier: high-Z, low-noise, wideband amplifier. A response peak (for  $R_{sig} \sim 1\text{k}\Omega$ ) can be tamed by adding a  $10\text{-}20\text{ pF}$  capacitor across the input, see Figure 3.35.

But the load seen by the drain (at signal frequencies) is in fact the low-impedance input of the second stage, which is an op-amp configured here as a current-to-voltage (“transresistance”) converter (see §4.3.1C). Its “gain” (ratio of output voltage to input current, thus units of resistance) is just  $R_1$ , making the overall open-loop gain  $G = g_m R_1$  (again assuming no feedback, and  $R_g = 0$ ). So, for the circuit values shown, the open-loop gain  $G_{OL} \approx 2500$ .

Now we close the loop via  $R_f$ , subtracting from the input a fraction  $R_g/(R_g + R_f)$ , for an ideal closed-loop gain  $G_{CL} = 1 + R_f/R_g = 50$ . The loop gain (ratio of open-loop to closed-loop gains) is about 50, adequate for good linearity and predictability of gain (see §2.5.3). Note the low resistance of  $R_g$ : it should be small enough so that the open-loop gain is not much reduced (thus  $R_g < 1/g_m$ ); and it should also be small enough so that its Johnson noise contribution is insignificant (§8.1). For a JFET transconductance of 25 mS, the first constraint limits  $R_g$  to somewhat less than  $40\Omega$ ; and for this JFET’s noise voltage of  $\sim 1\text{nV}/\sqrt{\text{Hz}}$  the second constraint limits  $R_g$  to somewhat less than  $25\Omega$ . Thus we choose the low  $\sim 10\Omega$  value. With the values shown, the open-loop gain is reduced by  $\sim 20\%$ , the in-

put noise voltage is increased by  $\sim 8\%$ , and the feedback network loads the op-amp to  $\pm 20\text{ mA}$  at full swing.<sup>40</sup>

The op-amp here was chosen for its wide bandwidth (it falls to unity gain at  $\sim 100\text{ MHz}$ ), so the circuit’s closed-loop gain rolls off at about  $20\text{ MHz}$ ,<sup>41</sup> as seen in the measured data of Figure 3.35. A bonus is the impressive output drive capability: up to  $100\text{ mA}$ , and full  $\pm 10\text{ V}$  swing to nearly  $10\text{ MHz}$ . The small compensation capacitor  $C_c$  enhances stability: with no compensation we measured  $5\text{ dB}$  of peaking at  $16\text{ MHz}$ ; adding  $C_c$  resulted in an insignificant  $0.1\text{ dB}$  peaking at  $10\text{ MHz}$  and a high-frequency rolloff of  $-3\text{ dB}$  at  $22\text{ MHz}$ .<sup>42</sup>

An alternative circuit that exploits the JFET’s high input impedance is an input JFET source follower (§3.2.6) driving a stage of voltage gain. That’s a perfectly good configuration, particularly if a current sink is used for the source pulldown. But the circuit of Figure 3.34 excels in achieving both lower noise and better linearity.

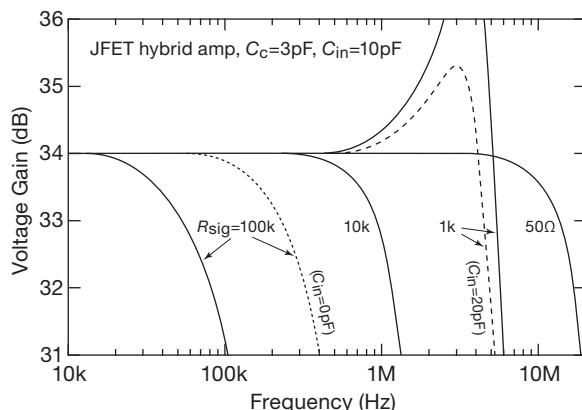
### 3.2.4 Differential amplifiers

Thus far we’ve dealt with the JFET’s uncertain  $I_D$  versus  $V_{GS}$  relationship by restricting ourselves to ac-coupled amplifier designs. But we can do better: a JFET matched pair lets us make dc-coupled amplifiers of respectable performance. And of course their very low input current means that these circuits can serve as high-input-impedance frontend stages for bipolar differential amplifiers, as well as for the important op-amps and comparators we’ll meet in the next chapter. As we mentioned earlier, the substantial  $V_{GS}$  offsets of FETs will generally result in larger input voltage offsets and offset drifts than with a comparable amplifier

<sup>40</sup> If you’re not happy with that, you could double or triple  $R_f$  and  $R_g$ , at the expense of slightly higher amplifier noise. See §8.6 for much more on low-noise JFET design.

<sup>41</sup> When driven with a low signal impedance. With a higher impedance signal source, the rolloff is dominated by the circuit’s input capacitance:  $Q_1$  has a drain-to-gate capacitance  $C_{TSS}$  of  $6\text{ pF}$ , thus the observed  $-3\text{ dB}$  bandwidth of  $\sim 400\text{ kHz}$  with a  $100\text{k}\Omega$  source. Happily, the larger gate-to-source capacitance ( $C_{iss} \approx 30\text{ pF}$ ) is bootstrapped into insignificance by feedback. “And what of the evil Miller effect?” you wonder. That’s suppressed here, because the op-amp holds its  $(-)$  input fixed (a “virtual ground”) via feedback through  $R_1$ ; see §4.3.1C.

<sup>42</sup> In any feedback circuit there’s the possibility of some “peaking” in the response at some frequency (or, in the worst case, a full-blown oscillation). This circuit exhibits modest peaking for input impedances in the range of a few  $\text{k}\Omega$ , as seen in Figure 3.35. Its amplitude can be tamed by adding  $\sim 10\text{-}20\text{ pF}$  of input shunt capacitance, with some consequent reduction of bandwidth.



**Figure 3.35.** Measured gain versus frequency for the amplifier of Figure 3.34. The solid curves are with a 10 pF shunt capacitor across the input, a compromise value for good performance over a wide range of input signal impedances (omit the capacitor for high source impedance, increase it for the worst-case  $R_{\text{sig}} \approx 1\text{k}\Omega$ ).

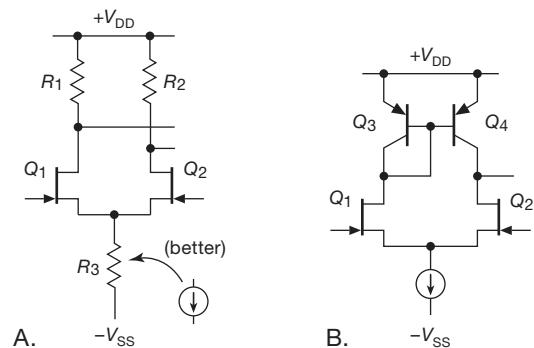
constructed entirely with bipolar transistors, but of course the input impedance will be raised enormously.

Figure 3.36 shows the simplest configurations, analogous to the simple BJT differential amplifiers of Figures 2.63 and 2.67. The differential gain of the classic long-tailed pair in Figure 3.36A (defined as  $\Delta V_{\text{out}}/\Delta V_{\text{in}}$ , with  $R_1 = R_2$  and differential output as shown) is just  $G = g_m R_1$ ; the common-mode rejection is greatly improved with a current sink substituting for the source resistor  $R_S$ . A drawback of this circuit is the uncertainty in gain (owing to the uncertainty in transconductance); and the gain is modest, owing to the limited transconductance of JFETs in general. You can circumvent the gain limitation by replacing the drain load resistor(s) with a current mirror, as in Figure 3.36B. This circuit fragment is not bias stable, however: it must be accompanied by a following stage that is configured to provide dc feedback.

These circuits suffer also from the Miller effect (§2.4.5), whose multiplying action on the feedback capacitance  $C_{\text{rss}}$  acts to increase the effective input capacitance, thus (in combination with the signal's source impedance) reducing the bandwidth. The current mirror in circuit (B) clamps  $Q_1$ 's drain, but the Miller effect is present still at  $Q_2$ 's input.<sup>43</sup> As with BJT amplifier stages, an effective method of eliminating the Miller effect is the use of a cascode transistor (either JFET or BJT) in the drain(s), a desirable configuration discussed later and in Chapter 3x in the section

“Bandwidth of the Cascode.”. By clamping the drain voltage, the cascode also eliminates a reduction in gain seen in simple circuits like that of Figure 3.36A: that’s because the JFET’s drain-current depends somewhat on drain voltage (the upward slope in plots of drain-current versus drain-source voltage, which you can think of as a finite output impedance), which can reduce the ideal  $g_m R_1$  gain by as much as 25%.

For this and other reasons the cascode is highly recommended, even when bandwidth is not an issue. Curious about those “other reasons”? Consider this: the gate current in JFETs, normally down in the picoamperes, rises precipitously with drain-to-source voltage – see Figure 3.49 on page 164 where picoamps become microamps! A cascode lets you clamp the drain to a low operating voltage, suppressing this effect. This is nicely illustrated in the figure toward the end of §3x.4 in Chapter 3x.



**Figure 3.36.** Simplest JFET differential amplifiers.

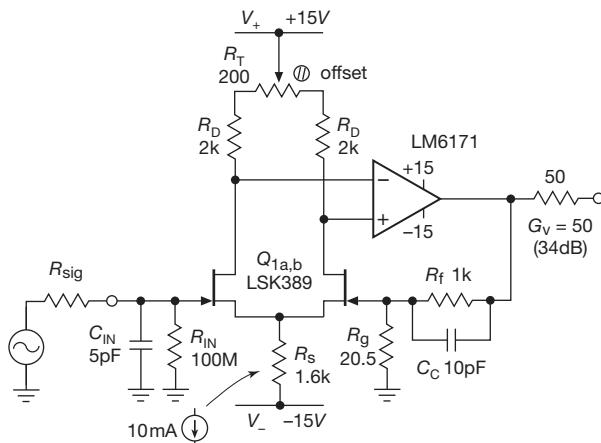
#### A. Example: a dc-coupled hybrid JFET amplifier

In the earlier hybrid amplifier design example (Figure 3.34) we dodged the issue of unpredictable JFET drain current by settling for an ac-coupled amplifier. That’s fine for something like an audio or RF amplifier; but sometimes you’d like response all the way down to dc.

You can achieve that by exploiting a matched-pair JFET differential input stage in a fully dc-coupled arrangement, as in Figure 3.37. The overall configuration is a common-source differential amplifier with the input signal connected to one side. Its differential voltage output drives a wideband op-amp, whose output (divided by the gain-setting voltage divider) provides negative feedback to the other terminal of the input pair. As with the previous circuit, the closed-loop gain is  $G_{\text{CL}} = 1 + R_f/R_g = 50$ , with a small loop compensation capacitor  $C_c$  chosen for best response without peaking.

The devil’s in the details, which, if you’re lucky,

<sup>43</sup> Unless that output drives a transimpedance stage that clamps its voltage, as in Figure 3.31 or 3.34.



**Figure 3.37.** An op-amp closes the loop around a JFET matched pair to create a dc-coupled wideband low-noise amplifier with high input impedance. Replace source bias resistor  $R_s$  with a current sink (see Figure 3.26) for enhanced supply rejection.

may come together in a symphonic harmony. This circuit worked out nicely; thus our enthusiasm for a bit of discussion. Here goes.

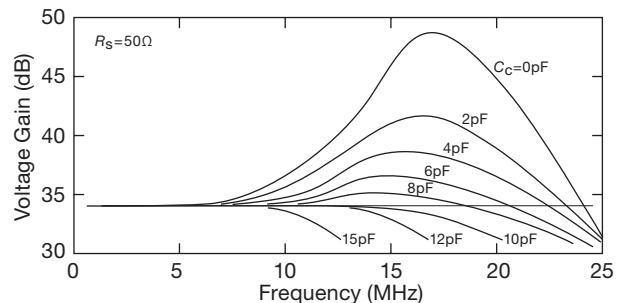
The LSK389 is a very low-noise ( $\sim 1 \text{ nV}/\sqrt{\text{Hz}}$ ) monolithic matched JFET pair,<sup>44</sup> available in three  $I_{DSS}$  grades. We chose the midcurrent -B suffix part (with a specified  $I_{DSS}$  between 6 mA and 12 mA), and we force each JFET of the pair to run at 5 mA by sinking 10 mA from the pair of source terminals. That puts the drains at +5 V (the op-amp enforces equality), with an open-loop gain (from single-ended input to differential output) of approximately  $G = g_m(R_D + 0.5R_T) \approx 40$ . To this the op-amp contributes its substantial open-loop gain (90 dB, i.e.,  $\times 30,000$ ), in a scary-looking configuration whose stability might seem seriously to be in doubt. But not to worry – the  $\div 50$  divider in the return loop safely limits the loop gain,<sup>45</sup> with any tendency toward instability easily tamed with appropriate choice of  $C_c$ .

According to the datasheet, the input pair  $Q_{1ab}$  has “Tight Matching.” But that’s on the scale of JFETs, not BJTs – here that “matching” is a whopping  $\pm 20 \text{ mV}$  maximum (about 100× the matching of a good BJT pair), which the circuit’s  $G = 50$  would amplify to an output offset of 1 V! Hence the offset trim  $R_T$ , with enough range to balance the worst-case input offset.

Our initial design included a bypass capacitor from the drain of the input transistor to ground, to suppress the

Miller effect. A nice thought, but the reality is that the Miller effect is almost entirely absent owing to the op-amp’s clamping effect on the drain pair. And the bypass capacitor introduces two problems: it unbalances the transistor pair, so the circuit is sensitive to noise on the positive supply rail; and it introduces a phase shift within the loop, causing some undesirable peaking, thus requiring a larger value of  $C_c$  and therefore reduced bandwidth.

Now for the feedback stability “compensation.” In a single-ended circuit like Figure 3.34 you can put a small capacitor across either  $R_1$  or across  $R_f$ . Here, though, we want to maintain input-stage symmetry, so  $C_c$  has to go across  $R_f$ . On the bench we found that 10 pF was required to eliminate peaking in the frequency response, when tested with a low-level sinewave input; Figure 3.38 shows the data.



**Figure 3.38.** Selecting the compensation capacitor  $C_c$  for the amplifier of Figure 3.37. A value of 8 or 10 pF works well.

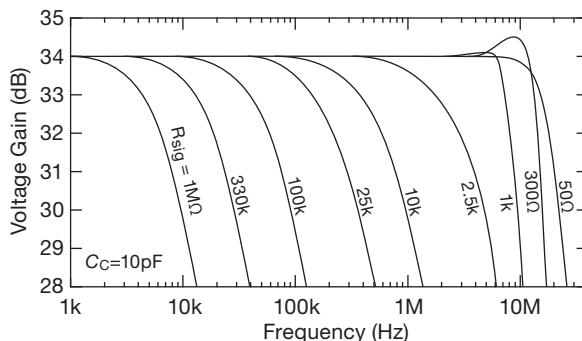
A final bit of compensation is the shunt capacitor  $C_{in}$  across the input to suppress some modest peaking that is seen for signal sources of  $\sim 1\text{k}\Omega$  impedance ( $R_{sig}$ ). Here 5 pF did the job nicely (though it increased the circuit’s input capacitance to  $\sim 20 \text{ pF}$ ). Figure 3.39 shows the measured gain versus frequency for the finished circuit (with the component values shown in Figure 3.37), for nine values of  $R_{sig}$  (spanning four decades of resistance).

As with the previous circuit, the LM6171 op-amp provides a full  $\pm 10 \text{ V}$  output swing to nearly 10 MHz. The  $50\Omega$  resistor in series with the output ensures stability into capacitive loads; it also provides “back termination” into  $50\Omega$  coax cable (see the Appendix H on Transmission Lines). This op-amp is not particularly quiet ( $e_n \sim 12 \text{ nV}/\sqrt{\text{Hz}}$ ), but that’s good enough: the circuit’s input-stage gain of  $\sim 40$  reduces the op-amp’s noise contribution to  $\sim 0.3 \text{ nV}/\sqrt{\text{Hz}}$  when referred to the input. The amplifier’s overall noise is  $\sim 2 \text{ nV}/\sqrt{\text{Hz}}$ .<sup>46</sup>

<sup>44</sup> A replacement for Toshiba’s legendary (and discontinued) 2SK389.

<sup>45</sup> To approximately the same value as that of the op-amp alone, were it connected as a stable unity-gain follower.

<sup>46</sup> If it’s *quiet* you want, you can substitute the large-die IF3602 dual



**Figure 3.39.** Measured gain versus frequency for the amplifier of Figure 3.37, for a range of values of signal source impedance  $R_{\text{sig}}$ .

This is a pretty good amplifier! With a bit of refinement (primarily substituting a low-noise current sink for  $R_S$ ), two such amplifiers, combined in a so-called “instrumentation amplifier” (INA) configuration, will outperform any available integrated INA in terms of both noise and speed; see Figure 8.49 in §8.6.3.

### B. Comparison with JFET-input op-amps

We can hear it already: “Yeah, yeah... you guys like to show off your clever circuits, with lots of discrete components. But nowadays that kind of circuit art is obsolete, because you can get all that performance, and more, in readily available integrated circuits – op-amps, in particular.”

But, *can* you? Table 3.2 lists the currently available op-amps with JFET inputs that have a chance of competing. How does their performance compare with that of our dc-coupled hybrid amplifier of Figure 3.37? Let’s see...

**Output swing** Only the first three op-amps can swing over the full  $\pm 15$  V range; OK, but...

**Bandwidth** ...the fastest of those “high-voltage” op-amps has 80 MHz GBW, thus for  $G=50$  a bandwidth of less than 2 MHz; and the two fastest op-amps, which can match our amplifier’s bandwidth, can swing only  $\pm 4$  V or so. By contrast, our amplifier’s GBW is 4 GHz (40 times the op-amp’s  $f_T=100$  MHz), and thus excess gain (e.g., 400 at 10 MHz) with lower resulting distortion.

**Noise** Our amplifier’s noise of  $\sim 2$  nV/ $\sqrt{\text{Hz}}$  is 6 dB quieter than the best of Table 3.2’s offerings.

**Cost** \$5–10 for the op-amp solution, roughly the same for the better-performing hybrid amplifier (\$2.50 for the LM6171, \$3.25 for the LSK389 dual JFET)

JFET. That will reduce the input noise to  $\sim 0.7$  nV/ $\sqrt{\text{Hz}}$ , but with a greatly increased input capacitance (about 300 pF!). (And thermal noise from the feedback network will degrade this unless  $R_g$  is decreased to  $\sim 5$  Ω; see §8.1 and Figure 8.80A.)

**Table 3.2 Selected Fast JFET-input Op-amps<sup>a</sup>**

Part #	Supply		$I_{\text{bias}}$ 25°C typ	$e_n$ 1kHz typ	GBW typ	Slew Rate typ	Cost qty 25 (\$US)
	Voltage range (V)	$I_Q$ typ (mA)	(pA)	(nV/ $\sqrt{\text{Hz}}$ )	(MHz)	(V/ $\mu\text{s}$ )	
OPA604A	9–50	5	50	10	20	25	2.93
OPA827A	8–40	5	15	4	22	28	9.00
ADA4637	9–36	7	1	6	80 <sup>d</sup>	170	10.12
OPA656	9–13	14	2	7 <sup>b</sup>	230	290	5.59
OPA657	9–13	14	2	7	1600 <sup>d</sup>	700	10.01
ADA4817	5–10.6	19	2	4 <sup>c</sup>	1050	870	4.93

Notes: (a) candidates for wideband low-noise amplifier.

(b) low  $e_nC$  noise:  $C_{\text{in}}=2.8$  pF. (c) lowest  $e_nC$  noise:  $C_{\text{in}}=1.5$  pF. (d) decomp,  $G_{\text{CL}}>7$ .

The contest so far? The hybrid amplifier is winning, in terms of the combined performance metric of bandwidth, output swing, and noise voltage. But we’re not done yet...

**Offset voltage** The op-amps win, here, with out-of-the-box  $V_{\text{os}}$  values of 2 mV (max) for the three fastest parts; the hybrid amp requires manual trimming of its 20 mV worst-case untrimmed offset, if better is needed.

**Parts count** A win, again, for the op-amps.

**Input capacitance** Just 1.5 pF for the ADA4817, versus 10 pF or more for the hybrid (the price we paid for 2× lower noise).

**Input current** 20 pA (max) for the ADA4817 (but that’s a low-voltage part), versus 200 pA for the hybrid (*unfair!* – that’s specified at a large negative bias,  $V_{\text{GS}} = -30$  V)

The verdict? A split decision: the JFET op-amp solution is simple and can deliver plenty of speed (or plenty of swing, but not both), along with untrimmed accuracy and very low-input capacitance (thus low “ $e_nC$ ” noise; see Chapter 8). The hybrid approach delivers speed and swing and lowest noise voltage; but it requires manual trim, it’s more complicated, and it has more input capacitance. Note also that an op-amp is a more flexible building block in general, providing, for example, a wide common-mode input-voltage range that our hybrid circuit does not have; but that is not needed here, because the input is always close to ground (owing to the circuit’s gain of 50).

### 3.2.5 Oscillators

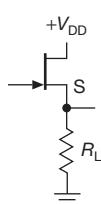
In general, FETs have characteristics that make them useful substitutes for bipolar transistors in almost any circuit that can benefit from their uniquely high input impedance and low bias current. A particular instance is the use of a JFET amplifier stage to implement a high-stability *LC* or crystal oscillator; we’ll show examples in §7.1.5D.

### 3.2.6 Source followers

Because of the relatively low transconductance of FETs, it's often better to use a FET "source follower" (analogous to an emitter follower) as an input buffer to a conventional BJT amplifier, rather than trying to make a common-source FET amplifier directly. You still get the high input impedance and zero dc input current of the FET, and the BJT's large transconductance lets you achieve high single-stage gain. Furthermore, discrete FETs (i.e., those that are not part of an integrated circuit) tend to have higher inter-electrode capacitance than BJTs, leading to greater Miller effect (§2.4.5B) in common-source amplifiers; the source-follower configuration, like the emitter follower, has no Miller effect.

FET followers, with their high input impedance, are commonly used as input stages in oscilloscopes as well as in other measuring instruments. There are many applications in which the signal-source impedance is intrinsically high, e.g., capacitor microphones, pH probes, charged-particle detectors, or microelectrode signals in biology and medicine. In these cases, a FET input stage (whether discrete or part of an integrated circuit) is a good solution. Within circuits there are situations in which the following stage must draw little or no current. Common examples are analog "sample-and-hold" and "peak detector" circuits, in which the level is stored on a capacitor and will "droop" if the next amplifier draws significant input current. In all these applications the advantage of negligible input current of a FET more than compensates for its low transconductance, making source followers (or even common-source amplifiers) attractive alternatives to the bipolar emitter follower.

Figure 3.40 shows the simplest source follower, which ideally should produce an accurate replica of the input waveform while drawing essentially zero input current. Let's figure out important things like its quiescent operating point, its exact voltage gain, its output impedance, and the voltage offset from input to output.



**Figure 3.40.** *N*-channel JFET source follower. Unlike the *npn* BJT emitter follower (in which the output trails the input by  $V_{BE} \approx 0.6$  V), the output is here more positive than the input.

### A. Quiescent operating point

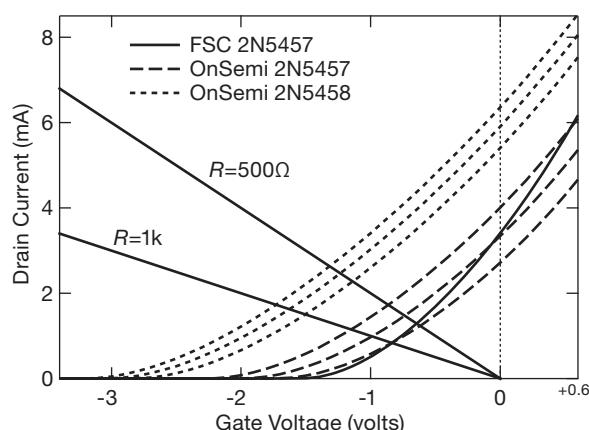
The JFET source follower is not as simply analyzed as the analogous BJT emitter follower, in which the emitter voltage simply trails the base voltage by a relatively constant (and predictable)  $V_{BE} \approx 0.6$  V. That is because the FET has less abrupt (and far less predictable) transfer characteristics ( $I_D$  versus  $V_{GS}$ ) – the same issue we've just wrestled with in connection with the JFET current source (§3.2.2) and the JFET amplifier (§3.2.3).

We could use the same iterative approach here, seeking the quiescent source voltage  $V_S$  (and therefore  $V_{GS} = V_S$ ) that produces a source current  $I_S$  (and therefore  $I_D = I_S$ ) consistent with that  $V_S$ . And we could do this by interpolating between curves like those in Figure 3.21A (a family of  $I_D$  versus  $V_{DS}$  for several  $V_{GS}$ 's), or by sliding up and down a transfer characteristic curve like that in Figure 3.41 (a single curve of  $I_D$  versus  $V_{GS}$ , at some fixed  $V_{DS}$ ), until we find the point at which  $I_D R_L = -V_{GS}$ .

But there is an elegant graphical method, used extensively during the days of vacuum tubes, that lets you find the operating point immediately: the method of "load lines."

### B. Load lines

To find the operating point for the source follower in Figure 3.40, we simply notice that the load resistor  $R_L$  imposes its rules on allowable  $V_{GS}$  versus  $I_S$ , namely, Ohm's law:  $I_S R_L = -V_{GS}$ . We can plot this constraint on the same graph as the transfer curve of Figure 3.41, as a straight line with slope  $-1/R_L$ ; note that it goes "backward,"



**Figure 3.41.** Measured transfer curves for a set of 2N5457 and 2N5458 *n*-channel JFETs at  $V_{DS}=10$  V. These measurements extend beyond  $I_{DSS}$ , with  $V_{GS}$  taken 0.6 V into positive territory. The Onsemi curves show the parts with the lowest, middle, and highest  $I_{DSS}$  from a batch of ten each.

because  $V_S = -V_{GS}$ . The operating point has to be consistent with this constraint and simultaneously with the transfer characteristic of the JFET. In other words, the operating point is the intersection of the two curves. In this case, with  $R_L = 1\text{k}\Omega$ , the quiescent point is at  $V_S = +1.6\text{ V}$  (and, from the lowest 2N5458 curve,  $I_D = 1.6\text{ mA}$ ).

Lest one be tempted to fall too quickly in love with this technique, we hasten to point out that the characteristic curves for a particular type of JFET exhibit a large spread. For the 2N5458 illustrated in Figure 3.41, for example, the specification allows  $I_{DSS}$  to be anywhere between 2 and 9 mA (and the pinchoff voltage  $V_{GS(\text{off})}$  can range from  $-1.0$  to  $-7.0\text{ V}$ ). In practice, it is rare to find devices at the extremes, and there tends to be good consistency in a single manufacturing batch (as indicated by the date code stamped on the parts); for example, by measuring a batch of 10 2N5458's (Figure 3.41) we determined that the quiescent point in this circuit would range from 1.52 to 1.74 V.

### C. Output amplitude and voltage gain

We can figure out the output amplitude, as we did for the emitter follower in Section 2.3.3, using the transconductance. We have

$$v_s = R_L i_d$$

since  $i_g$  is negligible; but

$$i_d = g_m v_{gs} = g_m (v_g - v_s),$$

so

$$v_s = \left[ \frac{R_L g_m}{(1 + R_L g_m)} \right] v_g.$$

That is, the gain is

$$G = \frac{1}{1 + \frac{1}{g_m R_L}}. \quad (3.7)$$

For  $R_L \gg 1/g_m$  it is a good follower ( $v_s \approx v_g$ ), with gain approaching, but always less than, unity. We are not near that limit in this example, in which the measured value  $g_m = 1.9\text{ mS}$  implies a voltage gain of  $G_V = 0.66$  into the  $1\text{k}\Omega$  load, far from the ideal of unity gain. Furthermore, the variation of transconductance over the signal swing results in undesirable nonlinearity. One solution is to use a JFET with higher transconductance, or (better) add a BJT transconductance enhancer (Figure 3.29F and §3x.2). But in situations in which the external load impedance is high, an elegant solution is to use a current sink as an active load, as we'll see presently (§3.2.6F).

### D. Input impedance

Our hope that JFET source followers have infinite input impedance is largely fulfilled, but they do have some gate leakage current (see §3.2.8) and input capacitance (see Table 3.7 on page 217). Gate leakage can become a problem at drain-gate voltage greater than about 5 V (Figure 3.49), so be sure to check the JFET's datasheet and, if necessary, consider adding a cascode to limit  $V_{DG}$ .

A follower's frequency response when driven by signals of high source impedance is limited by input capacitance,  $f_{3\text{dB}} = 1/2\pi C_{in}$ , where  $C_{in} = C_{iss} + C_{rss} + C_{stray}$ . The gate-source capacitance  $C_{iss}$  is generally about two to five times higher than the gate-drain capacitance  $C_{rss}$ , but fortunately it's bootstrapped by the follower action and is effectively reduced to  $(1 - G_V)C_{iss}$ . If you follow our advice (below) so that  $G_V$  is nearly 1.0, only the JFET's  $C_{rss}$  remains to limit the bandwidth. But it's possible to bootstrap the drain and reduce the effect of  $C_{rss}$  by a factor of 5. This leaves  $C_{stray}$  as "the last man standing" to limit the bandwidth – but you may be able to knock him down too by "guarding" most of the input wiring capacitance (i.e., using the follower's output signal to drive the cable's shield; see the discussion of signal guarding in §5.15.3).

### E. Output impedance

The preceding equation for  $v_s$  is precisely what you would predict if the source follower's output impedance were equal to  $1/g_m$  (try the calculation, assuming a source voltage of  $v_g$  in series with  $1/g_m$  driving a load of  $R_L$ ). This is exactly analogous to the emitter-follower situation, in which the output impedance was  $r_e = 25/I_C$ , or  $1/g_m$ . It can be easily shown explicitly that a source follower has output impedance  $1/g_m$  by figuring the source current for a signal applied to the output with the gate grounded (Figure 3.42). The drain current is

$$i_d = g_m v_{gs} = g_m v,$$

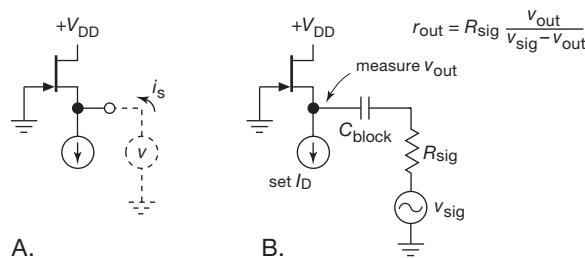
so

$$r_{out} = v/i_d = 1/g_m, \quad (3.8)$$

typically a few hundred ohms at currents of a few milliamperes.<sup>47</sup>

In general, FET source followers aren't nearly as stiff as emitter followers. The exception is at very low currents,

<sup>47</sup> In practice, a more convenient way to measure the follower output impedance is to inject a signal *current* and measure the resulting source voltage, as in Figure 3.42B. Get the current from a signal generator, with a series resistor  $R_{sig}$  much larger than  $r_{out}$ , taking care to keep  $v_{out}$  small, say  $\sim 50\text{ mV}$ ; then the equation in the figure gives you  $r_{out}$ .



**Figure 3.42.** Calculating source-follower output impedance.

in the subthreshold region, where the transconductance of some JFETs approaches that of a BJT operated at the same current; see Figure 3.54 on page 168.

In this example, with  $g_m = 1.9 \text{ mS}$ , the impedance looking back into the JFET's source is  $r_{\text{out}} = 525 \Omega$ , which combines with the parallel  $1\text{k}$  source load resistor to produce an output impedance of  $345 \Omega$ , quite a bit higher than the analogous value of  $r_e = 16 \Omega$  for a BJT operating at the same  $1.6 \text{ mA}$ .

We were able to calculate the voltage gain and output impedance reasonably accurately in this example because we took the trouble to measure the  $I_D$  versus  $V_{GS}$  characteristic curves. It's worth pointing out, however, that the manufacturer's datasheet for the 2N5458 gives us little help here: it gives no characteristic curves for the 2N5458, only for the lower-current part (2N5457); and for the 2N5458 it specifies  $g_m$  only at  $I_{DSS}$ , where it gives a range of  $1.5 \text{ mS}$  to  $5.5 \text{ mS}$ . From these limits, along with the  $I_{DSS}$  and  $V_{GS(\text{OFF})}$  limits above, we would not be able to form a good estimate of the operating transconductance, because the operating point with a fixed value of source load resistor is ill-determined. We could do better by assuming that we adjust  $R_S$  to make  $I_D=1.6 \text{ mA}$ , say; then, using the fact that  $g_m \propto \sqrt{I_D}$ , the specified limits of  $I_{DSS}$  and of  $g_m$  (at  $I_{DSS}$ ) guarantee that  $g_m$  lies in the range  $0.6 \text{ mS}$  to  $4.9 \text{ mS}$ .<sup>48</sup> Our measured value of  $g_m$  falls nicely within this range, being rather close to the geometric mean of these limits.

There are two drawbacks to this circuit.

1. The relatively high output impedance means that the output swing may be significantly less than the input swing, even with high load impedance, because  $R_L$  alone forms a divider with the source's output impedance. Furthermore, because the drain current is changing over the signal waveform,  $g_m$  and therefore the output impedance will

vary, producing some nonlinearity (distortion) at the output. The situation is improved if FETs of high transconductance are used, of course, but a combination FET–BJT follower (or FET–BJT “ $g_m$  enhancer,” Figure 3.29F) is often a better solution.

2. Because the  $V_{GS}$  needed to produce a certain operating current is a poorly controlled parameter in FET manufacture, a source follower has an unpredictable dc offset, a serious drawback for dc-coupled circuits.

(An additional problem is caused by the fact that an FET's drain current depends to some degree on drain-to-source voltage. You might call this the “ $g_{os}$  effect,” which also acts to reduce the gain from the ideal  $G = 1$ . It is discussed later in §§3.3.2 and in the Chapter 3x section “Bandwidth of the Source Follower with a Capacitive Load.”)

Perhaps this is a good place to pause and realize that many of the circuits we've been considering would be easier to implement, and would work better, if we had access to a negative supply voltage. But often that's not the case; so, in the spirit of real-world circuit design constraints (and as a useful learning exercise) we're slogging through the extra difficulties posed by single-supply JFET follower design. But if you do have a negative supply available, by all means use it!

## F. Active load

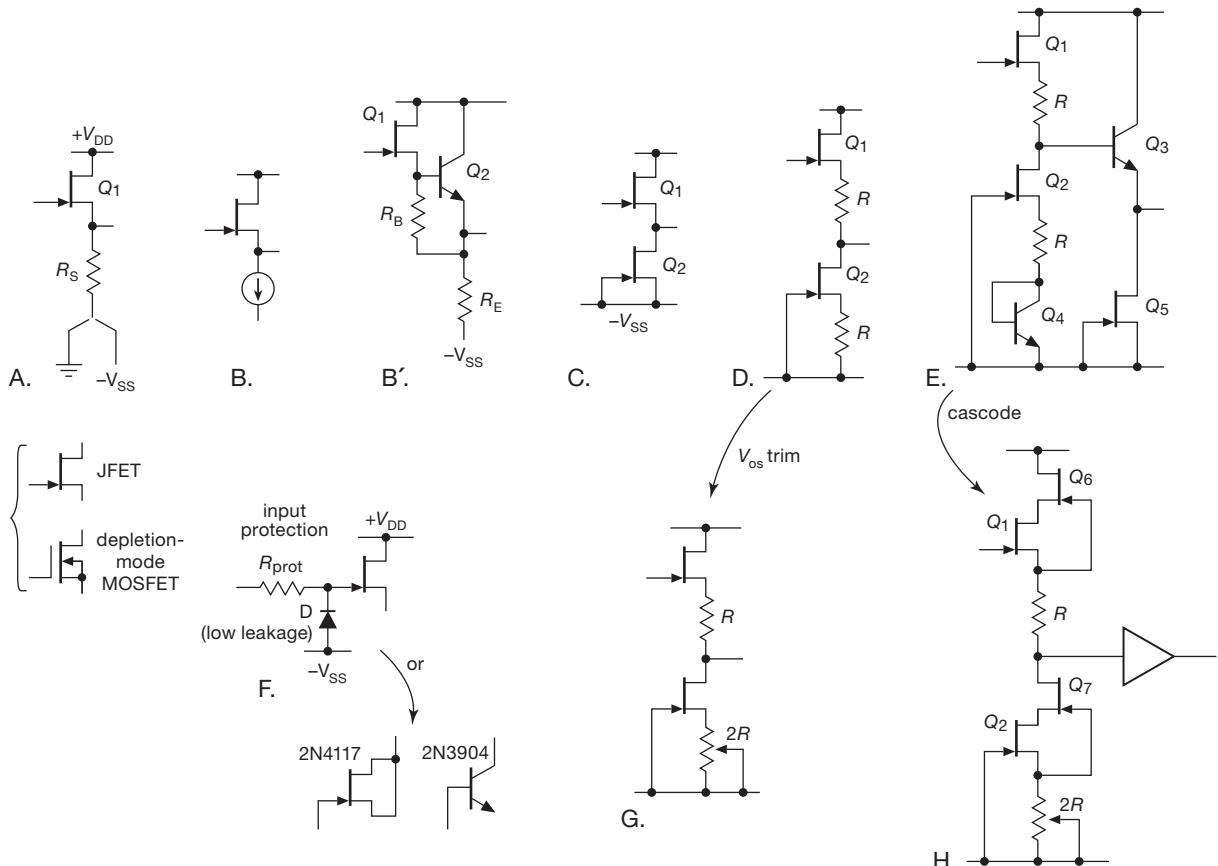
The addition of a few components improves the source follower enormously. Work with us here, as we take it in stages (Figure 3.43).

First we replace the load resistor (called  $R_S$  in Figure 3.43A) with a (pull-down) current sink (circuit B). (You can think of this as the previous case, with infinite  $R_S$ .) The constant source current makes  $V_{GS}$  approximately constant, thus reducing nonlinearities. A nice trick (circuit B') has a BJT follower doing double duty, both providing low output impedance while sinking a (roughly) constant current of  $V_{BE}/R_B$ .

We still have the problem of unpredictable (and therefore nonzero) offset voltage (from input to output) of  $V_{GS}$  (or  $V_{GS} + V_{BE}$ , for circuit B'). Of course, we could simply adjust  $I_{\text{sink}}$  to the particular value of  $I_{DSS}$  for the given FET (in the first circuit) or adjust  $R_B$  (in the second). This is a poor solution for two reasons: (a) it requires individual adjustment for each FET; and (b) even so,  $I_D$  may vary by a factor of two over the normal operating temperature range for a given  $V_{GS}$ .

A better circuit uses a matched FET pair to achieve zero offset (circuit C).  $Q_1$  and  $Q_2$  are a matched pair, on a single chip of silicon, for example the excellent LSK389 (see

<sup>48</sup> In fact, one can narrow that estimate somewhat, because  $g_m$  and  $I_{DSS}$  are correlated: a JFET sample with unusually high  $g_m$  will lie at the high end of the  $I_{DSS}$  distribution as well.



**Figure 3.43.** JFET unity-gain source followers – from simplest to best.

Table 3.7 on page 217), with  $Q_2$  sinking a current of  $I_{DSS}$ , i.e., its drain-current corresponding to  $V_{GS}=0$ . But the JFETs are matched, so  $V_{GS}=0$  for both transistors: voilà,  $Q_1$  is a follower with zero offset. Because  $Q_2$  tracks  $Q_1$  in temperature, the offset remains near zero, independent of temperature.

You usually see the preceding circuit with source resistors added (circuit D). A little thought should convince you that the upper resistor  $R$  is necessary and that equal-value resistors guarantee that  $V_{out} = V_{in}$  if  $Q_1$  and  $Q_2$  are matched. This circuit modification gives better  $I_D$  predictability, it allows you to set the drain current to some value less than  $I_{DSS}$ , and the source degeneration gives improved linearity. The variation of circuit G lets you trim the (already small) residual offset voltage caused by imperfect matching of the  $Q_1Q_2$  pair; the LSK389, for example, specifies a worst-case mismatch (at 1 mA drain current) of  $\Delta V_{GS} = 20 \text{ mV}$ .<sup>49</sup>

Circuit E adds a BJT output follower ( $Q_3$ ), with a JFET

current sink pull-down ( $Q_5$ ). Transistor  $Q_4$  adds a compensating  $V_{BE}$  in  $Q_2$ 's source to maintain approximately zero dc offset from input to output.

Circuits A–D all share a problem, namely, the drain-to-source voltage across  $Q_1$  varies directly with input signal. This can cause several undesirable effects. For example, imagine that circuit C is run between  $\pm 10 \text{ V}$  supply rails and that the input signal swings between  $+5$  and  $-5 \text{ V}$ .

the two JFETs, but in circuit D there is a  $V_{DS}$  mismatch that depends on the input signal voltage relative to the supply rails. To estimate the resulting offset voltage of the follower you need to know the JFETs' output conductance ( $g_{os}$ ), which causes an input–output offset in the follower that is proportional to the  $V_{DS}$  mismatch. That parameter is not specified on this JFET's datasheet, but from our measurements (see Table 3.7 on page 217) we know that  $g_{os} \approx 100 \mu\text{S}$ , which causes a follower offset of  $\Delta V = \Delta V_{DS} / G_{max}$ ; here that amounts to  $\approx 60 \text{ mV}$  for a  $10 \text{ V}$  difference in  $V_{DS}$ , quite a bit larger than the  $20 \text{ mV}$  maximum untrimmed offset of the JFET pair (when the  $V_{DS}$ 's are balanced). The cure? Circuit H, a cascode in each JFET to hold each  $V_{DS}$  constant. Thunderous applause, yet again, for the remarkable cascode.

<sup>49</sup> But a tricky “gotcha”: the offset specification assumes equal  $V_{DS}$  for

At the positive signal peak,  $Q_1$  has less than 5 V from drain to source, while  $Q_2$  has more than 15 V. Because a FET's drain current (at fixed  $V_{GS}$ ) varies slightly with drain-to-source voltage (discussed in §§3.3.2 and 3x.4), consequences here are a departure from strict unity gain and (worse) a potential nonlinearity; another consequence is that the input gate current can rise dramatically at drain-to-source voltages greater than 5 V (see Figure 3.49 on page 164), seriously degrading the otherwise admirably low input current.

An excellent solution to these problems (and others!) is the cascode configuration, as in circuit H. Here we've added JFETs  $Q_6$  and  $Q_7$ , which need not be matched, but which must be chosen to have a  $V_{GS}$  larger than the minimum desired  $V_{DS}$  of  $Q_1$  and  $Q_2$ . The cascode transistors bootstrap  $V_{DS}$  of  $Q_1$  and  $Q_2$  to a voltage equal to the  $V_{GS}$  of  $Q_6$  and  $Q_7$  while passing through the drain currents. So  $Q_1$  and  $Q_2$  operate at constant (and low)  $V_{DS}$ , with the cascode transistors taking up the slack as the signal swings, thus addressing both problems described in the preceding paragraph. The results are dramatic, as we'll see presently in a low-distortion "case study."

A further improvement on these JFET follower circuits is the addition of a *pnp* transconductance enhancer in  $Q_1$ 's drain circuit (as in Figure 3.29, where the greatly increased transconductance rescues an otherwise mediocre amplifier stage); this is particularly helpful if the follower is to drive a relatively low load impedance. This is fleshed out in more detail in §3x.2.

JFETs can handle plenty of forward gate current, but they are easily damaged by reverse breakdown. When that possibility exists, it's a good idea to add gate protection, as in circuit F. The series resistor  $R_{prot}$  limits current through the clamp diode  $D$  (which should be a low-leakage part like the 1N3595, if low input current is important). You can use the base-collector junction of an ordinary BJT, or the gate-channel diode of a JFET; see the plot of measured diode reverse leakage currents in §1x.7. But there's a compromise here: a large value of  $R_{prot}$  safely limits the clamp current, but it introduces excessive Johnson (thermal) noise, a serious issue in low-noise applications. The use of a depletion-mode MOSFET current limiter solves this problem elegantly; see §5.15.4 for details.

Note that the JFETs in these examples can be replaced with depletion-mode MOSFETs, which are available with voltage ratings to 1000 V; in that case it's necessary to protect the gate against both forward and reverse overvoltages greater than  $\pm 20$  V.

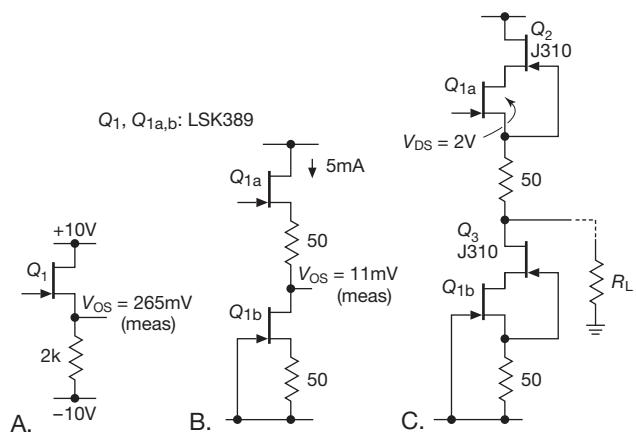
In a further variation of these circuits, you can use the output signal to drive an inner "guard" shield in order to

effectively eliminate the effects of shielded-cable capacitance, which would otherwise be devastating for the high source impedances that you might see with this sort of high-impedance input buffer amplifier.

### G. Case study: low-distortion JFET follower

To explore quantitatively the improvement you get with a current-sink pull-down, and further with a cascode arrangement, we wired up the three follower circuits in Figure 3.44, each with an LSK389 dual JFET; these correspond to circuits A, D, and H of Figure 3.43. To seriously challenge these circuits' linearity, we drove each with a clean 1 kHz sinewave,<sup>50</sup> at signal amplitudes that pushed precariously close to the supply rails.

The simple resistor pull-down circuit (Figure 3.44A) exhibited an expected dc offset (about 0.25 V at the quiescent point), with measured distortion (Figure 3.45) going from 0.02% (at 1 Vrms) to about 0.14% (at 5 Vrms). That's pretty decent performance, especially given that this circuit is entirely open-loop (no feedback); it's better than we expected. The distortion was almost entirely second harmonic (i.e., at  $2f_{in}$ ).

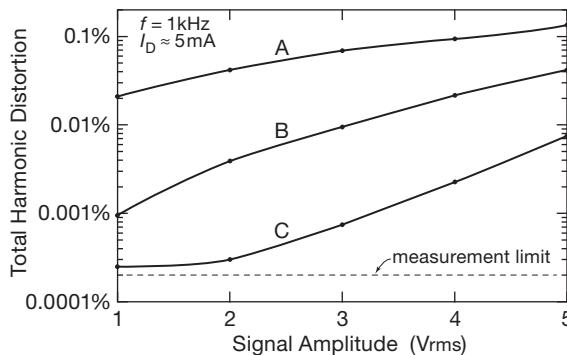


**Figure 3.44.** Three candidates for the JFET follower low-distortion medal-of-honor.

Adding an LSK389 matched JFET current sink with source degeneration (Figure 3.44B) made a nice improvement: a dc offset of about 10 mV and a measured distortion reduced by a factor of ten (20 dB), with the distortion now almost entirely third harmonic ( $3f_{in}$ ). We're in serious audiophile territory here. Finally, adding a cascode (the

<sup>50</sup> From an SRS DS360 "Ultra-low distortion function generator": distortion less than 0.0003%. We measured the output distortion with a ShibaSoku 725B distortion analyzer.

J310's  $V_{GS}$  is much larger than that of  $Q_1$ , so that the latter runs at  $V_{DS} \approx 2$  V) improves the linearity by another 20 dB, bumping into the measurement floor of our modest apparatus.<sup>51</sup> The low drain-to-source voltage across  $Q_{1a}$  imposed by the cascode also ensures a low input gate current.<sup>52</sup>



**Figure 3.45.** Measured distortion versus signal amplitude for the JFET followers of Figure 3.44, with  $R_L=1$  M.

### 3.2.7 FETs as variable resistors

Figure 3.21 showed the region of JFET characteristic curves (drain current versus  $V_{DS}$  for a small family of  $V_{GS}$  voltages), both in the normal ("saturated") regime and in the "linear" region of small  $V_{DS}$ . We showed the equivalent pair of graphs for a MOSFET at the beginning of the chapter (Figure 3.2). The  $I_D$ -versus- $V_{DS}$  curves are approximately straight lines for  $V_{DS}$  smaller than  $V_{GS} - V_{th}$ , and they extend in both directions through zero, i.e., the device can be used as a voltage-controlled resistor for small signals of either polarity. From our equation for  $I_D$  versus  $V_{GS}$  in the linear region (§3.1.4, eq'n 3.1) we easily find the ratio ( $I_D/V_{DS}$ ) to be

$$\frac{1}{r_{DS}} = 2\kappa \left[ (V_{GS} - V_{th}) - \frac{V_{DS}}{2} \right]. \quad (3.9)$$

The last term represents a nonlinearity, i.e., a departure from resistive behavior (resistance shouldn't depend on signal voltage). However, for drain voltages substantially less than the amount by which the gate is above threshold ( $V_{DS} \rightarrow 0$ ), the last term becomes unimportant, and the FET behaves approximately like a resistance:

<sup>51</sup> Oscillator courtesy of eBay; distortion analyzer courtesy of the MIT Flea Market.

<sup>52</sup> These impressively low distortions were measured into a high impedance. If you want to drive a substantial load, you may need to add a " $g_m$  enhancer" to  $Q_{1a}$ ; see §§3x.2 and 3x.4 in Chapter 3x.

$$r_{DS} \approx 1/[2\kappa(V_{GS} - V_{th})]. \quad (3.10)$$

Because the device-dependent parameter  $\kappa$  isn't a quantity you are likely to know, it's more useful to write  $r_{DS}$  as

$$r_{DS} \approx r_{G0}(V_{G0} - V_{th})/(V_G - V_{th}), \quad (3.11)$$

where the resistance  $r_{DS}$  at any gate voltage  $V_G$  is written in terms of the (known) resistance  $r_{G0}$  at some gate voltage  $V_{G0}$ .

**Exercise 3.2.** Derive the preceding "scaling" law.

From either formula you can see that the conductance ( $=1/r_{DS}$ ) is proportional to the amount by which the gate voltage exceeds threshold. Another useful fact is that  $r_{DS}=1/g_m$ , i.e., the channel resistance in the *linear* region is the inverse of the transconductance in the *saturated* region. This is a handy thing to know, because either  $g_m$  or  $r_{DS}$  is a parameter nearly always specified on FET data sheets.

**Exercise 3.3.** Show that  $r_{DS}=1/g_m$  by finding the transconductance from the saturation drain-current formula in §3.1.4.

Typically, the values of resistance you can produce with FETs vary from a few tens of ohms (as low as  $0.001\Omega$  for power MOSFETs) all the way up to an open circuit. A typical application might be an automatic-gain-control (AGC) circuit in which the gain of an amplifier is adjusted (by means of feedback) to keep the output within the linear range. In such an AGC circuit you must be careful to put the variable-resistance FET at a place in the circuit where the signal swing is small, preferably less than 200 mV or so.

The range of  $V_{DS}$  over which the FET behaves like a good resistor depends on the particular FET and is roughly proportional to the amount by which the gate voltage exceeds threshold. Typically you might see nonlinearities of about 2% for  $V_{DS} < 0.1(V_{GS} - V_{th})$ , and perhaps 10% nonlinearity for  $V_{DS} \approx 0.25(V_{GS} - V_{th})$ . Matched FETs make it easy to design a ganged variable resistor to control several signals at once. You can also find some JFETs specifically intended for use as variable resistors (e.g., the InterFET 2N4338-41 series, and VCR series), with nominal ON-resistances specified at some  $V_{GS}$  (usually 0V).

#### A. Linearizing trick

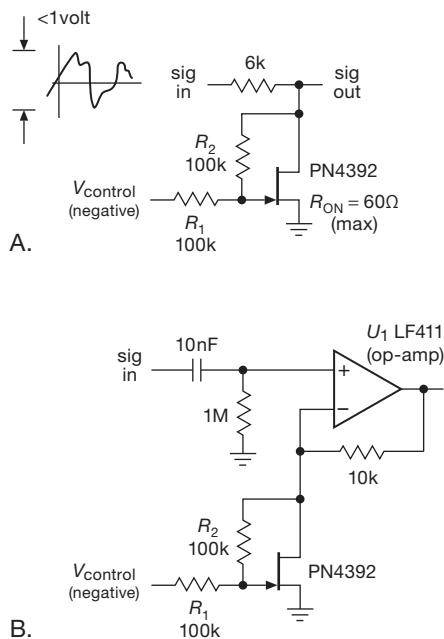
It is possible to improve the linearity and, simultaneously, the range of  $V_{DS}$  over which a FET behaves like a resistor, by a simple compensation scheme. Look at the expression 3.9 for  $1/r_{DS}$ ; you can see that the linearity will be nearly

perfect if you can add to the gate voltage a voltage equal to one half the drain-source voltage. Figure 3.46 shows two circuits to do exactly that.

In the first, the JFET forms the lower half of a resistive voltage divider, thus forming a voltage-controlled attenuator (or “volume control”).  $R_1$  and  $R_2$  improve the linearity by adding a voltage of  $0.5V_{DS}$  to  $V_{GS}$ , as just discussed. The JFETs shown have an ON-resistance (gate grounded) of  $60\ \Omega$ (max), giving the circuit an attenuation range of 0 to 40 dB. In the second circuit, the JFET’s controllable resistance forms the lower leg of the gain-setting feedback divider in an op-amp noninverting voltage amplifier (the voltage gain is  $G=[10k/R_{FET}]+1$ ).

The linearization of  $r_{DS}$  with a resistive gate-divider circuit, as above, is remarkably effective. In Figure 3.47 we’ve compared actual measured curves of  $I_D$  versus  $V_{DS}$  in the linear (low- $V_{DS}$ ) region for FETs, both with and without the linearizing circuit. The linearizing circuit is essential for low-distortion applications with signal swings of more than a few millivolts. We used it for amplitude control in the oscillator circuit of Figure 7.22, where the JFET was combined with a series resistor to create a low-distortion gain trimmer; the measured distortion of just 0.0002%.

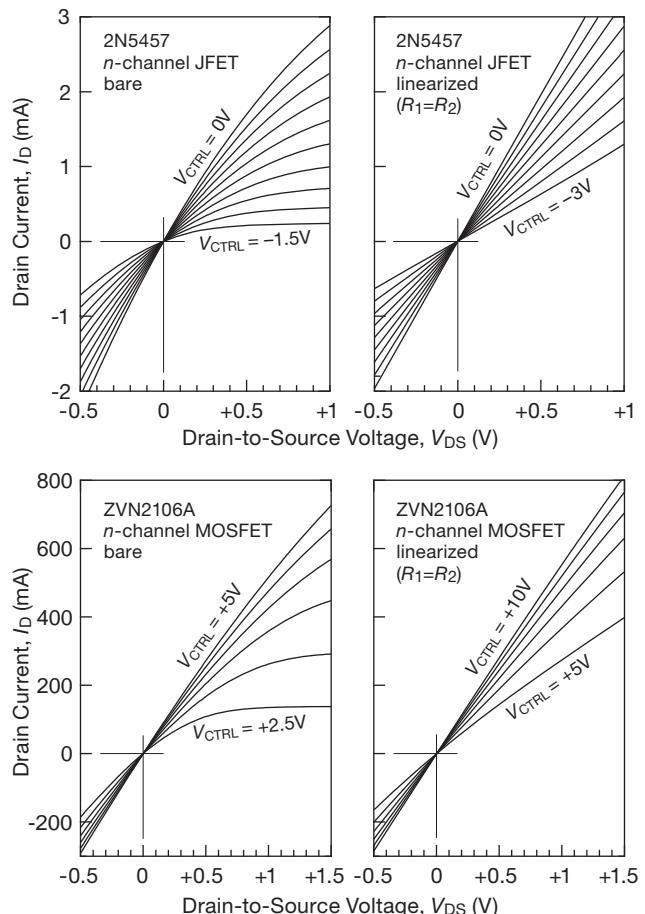
When considering FETs for an application requiring a gain control, e.g., an AGC or “modulator” (in which the amplitude of a high-frequency signal is varied at an audio rate, say), it is worthwhile to look also at “analog-



**Figure 3.46.** Linearizing the JFET variable resistor.

multiplier” ICs. These are high-accuracy devices with good dynamic range that are normally used to form the product of two voltages. One of the voltages can be a dc control signal, setting the multiplication factor of the device for the other input signal, i.e., the gain. Analog multipliers exploit the  $g_m$ -versus- $I_C$  characteristic of bipolar transistors [ $g_m = I_C(\text{mA})/25$  siemens], in a configuration known as the “Gilbert cell,” using matched arrays to circumvent problems of offsets and bias shifts. At very high frequencies (100 MHz and above), passive “balanced mixers” are often the best devices to accomplish the same task.

It is important to remember that a FET in conduction at low  $V_{DS}$  behaves like a good resistor all the way down to zero volts from drain to source (and even a bit beyond, in the opposite quadrant). There are no diode drops, sat-



**Figure 3.47.** Measured curves of  $I_D$  versus  $V_{DS}$  for both (top) a JFET and (bottom) a MOSFET, showing the linearizing effect of a resistor pair (as in Figure 3.46). Note the relatively high current scale for the MOSFET.

uration voltages, or the like to worry about. We will see op-amps and digital logic families (CMOS) that take advantage of this nice property, giving outputs that saturate cleanly to the power supplies.

### 3.2.8 FET gate current

We said at the outset that FETs in general, and MOSFETs in particular, have essentially zero steady-state gate current. This is perhaps the most important property of FETs, and it was exploited in the high-impedance amplifiers and followers in the previous sections. It will prove essential, too, in applications to follow – most notably analog switches and digital logic.

Of course, at some level of scrutiny we might expect to see some gate current. It's important to know about gate current, because a naive zero-current model is guaranteed to get you in trouble sooner or later. In fact, finite gate current arises from several mechanisms. (a) Even in MOSFETs the silicon dioxide gate insulation is not perfect, leading to leakage currents in the picoampere range. (b) In JFETs the gate “insulation” is really a back-biased diode junction, with the same impurity and junction leakage-current mechanisms as ordinary diodes. (c) Furthermore, JFETs (*n*-channel in particular) suffer from an additional effect known as “impact-ionization” gate current, which can reach astounding levels. (d) Finally – and most important for high-speed circuits – both JFETs and MOSFETs have *dynamic* gate current, caused by ac signals driving the gate capacitance; this can cause Miller effect, just as with bipolar transistors.<sup>53</sup> We'll deal with this important topic later, in §§3.5 and 3.5.4.

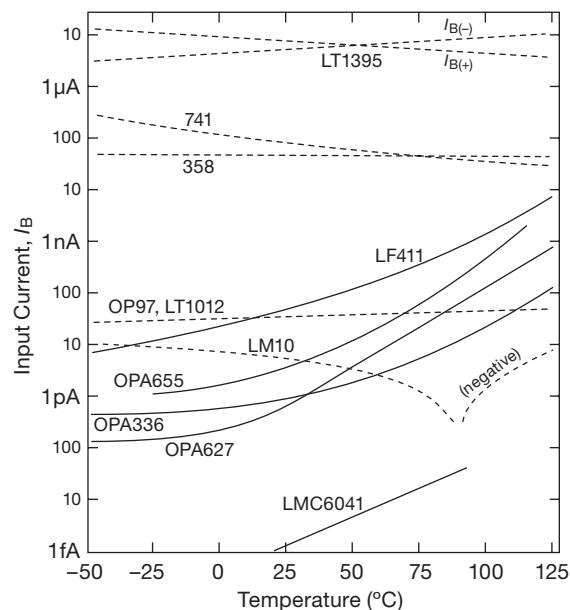
In most cases gate input currents are negligible in comparison with BJT base currents. However, there are situations in which a FET may actually have *higher* input current. Let's look at the numbers.

#### A. Gate leakage

The low-frequency input impedance of a FET amplifier (or follower) is limited by gate leakage. JFET datasheets usually specify a breakdown voltage,  $BV_{GSS}$ , defined as the voltage from gate to *channel* (source and drain connected together) at which the gate current reaches  $1\ \mu A$ . For smaller applied gate-channel voltages, the gate leakage current,  $I_{GSS}$ , again measured with the source and drain connected together, is considerably smaller, dropping

quickly to the picoampere range for gate-drain voltages well below breakdown. With MOSFETs you must never allow the gate insulation to break down; instead, gate leakage is specified as some maximum leakage current at a specified gate-channel voltage. Integrated circuit amplifiers with FETs (e.g., FET op-amps) use the misleading term “input bias current,”  $I_B$ , to specify input leakage current; it's usually in the picoampere range.

The good news is that these leakage currents are in the picoampere range at room temperature. The bad news is that they increase rapidly (in fact, exponentially) with temperature, roughly doubling every  $10^\circ C$ . By contrast, BJT base current isn't leakage, it's bias current, and in fact tends to *decrease* slightly with increasing temperature. The comparison is shown graphically in Figure 3.48, a plot of input current versus temperature for several IC amplifiers (op-amps). The FET-input op-amps have the lowest input currents at room temperature (and below), but their input current rises rapidly with temperature, crossing over the curves for amplifiers with carefully designed BJT input stages like the LM10 and LT1012. These BJT op-amps, along with “premium” low-input-current JFET op-amps like the OPA111 and OPA627, are fairly expensive. However, we also included everyday “jellybean” op-amps like the bipolar LM358 and JFET LF411/2 in the figure to give



**Figure 3.48.** The input current of a FET amplifier is gate leakage, which doubles every  $10^\circ C$ . The FET-input amplifiers in this plot (solid lines) are easily spotted by their characteristic upward slope.

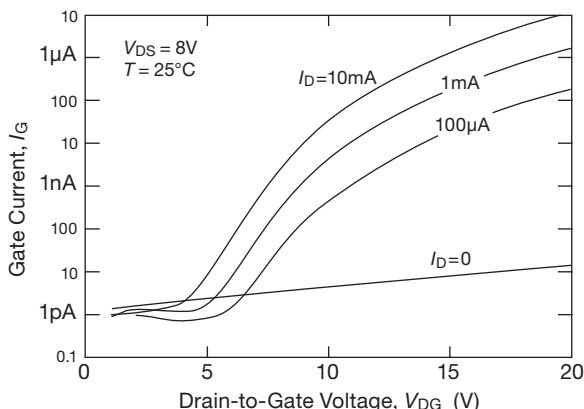
<sup>53</sup> In extreme cases, for example in high-voltage power switching, it can require amperes of gate drive current to switch a large MOSFET on the time scale of nanoseconds. This is not a trivial effect!

an idea of input currents you can expect with inexpensive (less than a dollar) op-amps.

### B. JFET impact-ionization current

In addition to conventional gate leakage effects, *n*-channel JFETs suffer from rather large gate leakage currents when operated with substantial  $V_{DS}$  and  $I_D$  (the gate leakage specified on datasheets is measured under the unrealistic conditions that both  $V_{DS}=0$  and  $I_D=0$ ). Figure 3.49 shows what happens. The gate leakage current remains near the  $I_{GSS}$  value until you reach a critical drain-gate voltage, at which point it rises precipitously. This extra “impact-ionization” current is proportional to drain current, and it rises exponentially with voltage and temperature. The onset of this current occurs at drain-gate voltages of about 25% of  $BV_{GSS}$ , and it can reach gate currents of a microamp or more. Obviously a “high-impedance buffer” with a microamp of input current is worthless. That’s what you would get if you used a BF862 as a follower, running 1 mA of drain current from a 20 volt supply.

This extra gate leakage current afflicts primarily *n*-channel JFETs, and it occurs at higher values of drain-gate voltage. Some cures are to (a) operate at low drain-gate voltage, either with a low-voltage drain supply or with a cascode, (b) use a *p*-channel JFET, where the effect is much smaller, or (c) use a MOSFET. The most important thing is to be aware of the effect so that it doesn’t catch you by surprise.



**Figure 3.49.** JFET gate leakage increases disastrously at higher drain-gate voltages and is proportional to drain current, as seen in the datasheet curves for the otherwise excellent BF862 *n*-channel JFET.

### C. Dynamic gate current

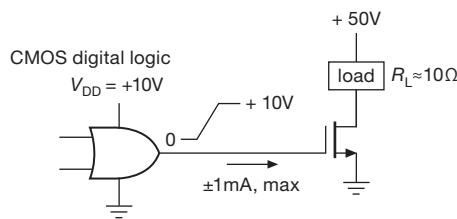
Gate leakage is a dc effect. Whatever is driving the gate must also supply an *ac* current, because of gate capacitance. Consider a common-source amplifier. Just as with bipolar transistors, you can have the simple effect of input capacitance to ground (called  $C_{iss}$ ), and you can have the capacitance-multiplying Miller effect (which acts on the feedback capacitance  $C_{rss}$ ). There are two reasons why capacitive effects are more serious in FETs than in bipolar transistors. First, you use FETs (rather than BJTs) because you want very low input current; thus the capacitive currents loom relatively larger for the same capacitance. Second, FETs often have considerably larger capacitance than equivalent bipolar transistors.

To appreciate the effect of capacitance, consider a FET amplifier intended for a signal source of 100k source impedance. At dc there’s no problem, because the picoampere currents produce only microvolt drops across the signal source’s internal impedance. But at 1 MHz, say, an input capacitance of 5 pF presents a shunt impedance of about 30k, seriously attenuating the signal. In fact, *any* amplifier is in trouble with a high-impedance signal at high frequencies, and the usual solution is to operate at low impedance (50 Ω is typical) or use tuned *LC* circuits to resonate away the parasitic capacitance. The point to understand is that the FET amplifier doesn’t look like a  $10^{12}\Omega$  load at signal frequencies.

As another example, imagine switching a 5 amp high-voltage load with a power MOSFET (there aren’t any high-power *JFETs*), in the style of Figure 3.50. One might naively assume that the gate could be driven from a digital logic output with low current-sourcing capability, for example the so-called 4000-series CMOS logic, which can supply an output current of the order of 1 mA with a swing from ground to +10 V. In fact, such a circuit would be a disaster, because 1 mA of gate drive into the 200 pF average feedback capacitance of the IRF740 would stretch the output switching speed to a leisurely 50  $\mu s$ .<sup>54</sup>

Even worse, the dynamic gate currents ( $I_{gate}=C dV_D/dt$ ) would force currents back into the logic device’s output, possibly destroying it by means of a perverse effect known as “SCR latchup” (more of which in Chapters 10 and 11);

<sup>54</sup> Our model is unacceptably crude, because feedback capacitance varies rapidly with drain voltage (see §3.5.4A). It’s OK to use a constant value for feedback capacitance  $C_{rss}$  for small-signal calculations; but for a switching application like this, you need to go to the datasheet for values of *gate charge*, which take into account the nonlinear behavior of the capacitances. In this example the datasheet specifies  $Q_G \approx 40\text{nC}$  of gate charge, producing a switching time of  $t = Q_G/i = 40\mu s$  at a (dynamic) drive current  $i$  of 1 mA.



**Figure 3.50.** Dynamic gate current example: driving a fast switching load.

for this and other reasons a series resistor (not shown in the figure) is usually added between the driving device and the MOSFET's gate. Bipolar power transistors have somewhat lower capacitances and therefore somewhat lower dynamic input currents (but still in the same ballpark); but when you design a circuit to drive a 5 amp power BJT, you're *expecting* to provide a few hundred millamps or so of base drive (via a Darlington or whatever), whereas with a FET you tend to take for granted low input current. In this example – in which you would have to supply a couple of *amperes* of gate drive current to bring about the 25 ns switching speed that the MOSFET is capable of – the ultra-high-impedance FET has lost some of its luster.

**Exercise 3.4.** Estimate the switching times for the circuit of Figure 3.50, with 1 amp of gate drive current, assuming either (a) an average feedback capacitance of 200 pF, or (more accurately) (b) a required gate charge of 40 nC.

### 3.3 A closer look at JFETs

In §3.1.4 we introduced the landscape of FET operating regions: for drain voltages of a volt or more (to get beyond the “linear” resistive region) there is the conventional operating region in which the saturation drain current<sup>55</sup>  $I_D$  is proportional to  $(V_{GS} - V_{th})^2$  and (at much lower drain currents) the subthreshold region in which  $I_D$  is exponential in  $V_{GS}$ .

That's the simple picture. Because JFETs are the devices of choice for accurate or low-noise (or both) circuits with high input impedance, it's worth looking more closely at their idiosyncrasies, ideally with measurements of actual devices.

<sup>55</sup> The term “saturation” can be confusing: for FETs it is used to denote *current* saturation, the region of drain voltages more than a volt or so where the drain current is approximately constant. By contrast, for BJTs the term “saturation” denotes *voltage* saturation (an ON switch), in which the collector voltage is close to zero. It never hurts to add the qualifying adjective (though people seldom do).

We've undertaken an exhaustive (and exhausting!) review of most of the available JFETs, collecting sample batches of each, often from multiple manufacturers. Table 3.7 on page 217 includes most of them, with some measured parameters ( $I_{DSS}$ , and  $g_{os}$ ,  $g_m$ , and  $V_{GS}$  at a useful drain current) alongside the datasheet specifications.<sup>56</sup> See also Table 8.2 on page 516, which lists a nice selection of low-noise JFETs.

Advanced JFET topics are treated in Chapter 3x; here we discuss a few essential subjects – JFET operating regions (including the often neglected *subthreshold* region), JFET transconductance, and JFET capacitance.

#### 3.3.1 Drain current versus gate voltage

A persistent issue with JFET circuit design is *parameter spread*. This is nicely illustrated in Figures 3.51 and 3.52, where we've plotted measured  $I_D$  versus  $V_{GS}$  for six samples of the 2N5457 *n*-channel JFET (three from each of two manufacturers<sup>57</sup>), and three samples of the related 2N5458 (from the same 2N5457–59 family). In each case we chose the parts with the highest, lowest, and median values of measured  $I_{DSS}$  in a batch of ten parts. For these measurements we ventured into positive gate voltages (up nearly to a diode drop, the onset of gate conduction), well beyond the usual zero-bias limit; nothing terrible happens, but in general this is a practice that should be avoided.

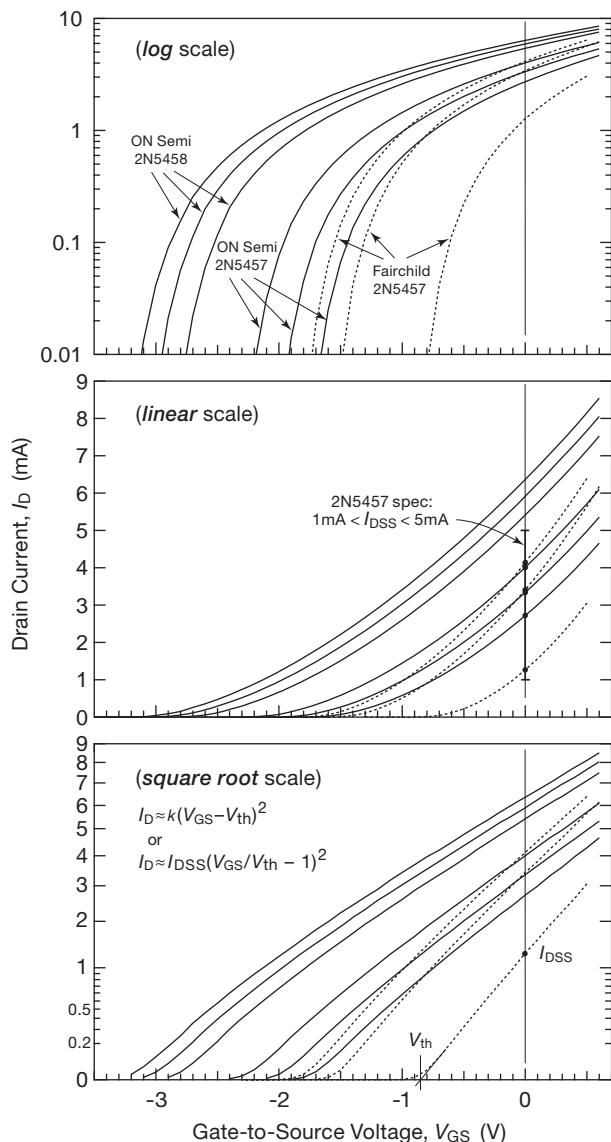
Let's look more closely at the data for these nine parts to understand different aspects of their performance and the impact on JFET circuit designs.

##### A. The quadratic region

The *linear* plot is what you usually see in datasheets and in textbooks. That scaling reveals well the quadratic behavior of drain current at a significant fraction of the zero-bias drain current ( $I_{DSS}$ ), where JFETs are most often operated [eq'n 3.4]. You can see also the variation of  $I_{DSS}$  among samples (the specified range for the half-dozen 2N5457 samples is indicated by a vertical bar), as well as the somewhat steeper slope for the Fairchild 2N5457 samples. The slope is, of course, just the transconductance gain,  $g_m = dI_D/dV_{GS}$ , increasing linearly with  $V_{GS}$  [eq'n 3.5] in this regime of where the drain current is quadratic in  $V_{GS} - V_{th}$ .

<sup>56</sup> See also Table 8.2 for noise parameters. These JFET tables represent several months' work – of comparing datasheet specifications, of checking distributors' stock, and of lab measurements. We found it revealing and rewarding, and we hope it is of use to the reader.

<sup>57</sup> We used dashed lines for the Fairchild parts, to untangle the crowd.



**Figure 3.51.** Drain current versus gate voltage for nine parts from the 2N5457–9 n-channel JFET family, operating with  $V_{DS}=5\text{V}$ . The same data is plotted on linear, log, and square-root axes. The ON Semi and Fairchild parts both meet the  $I_{DSS}$  spec (middle panel) despite their substantially different curves. Note that the measurements extend beyond  $I_{DSS}$ , to  $V_{GS} + 0.6\text{V}$ .

#### Finding $V_{th}$ : square-root plot

Look next at the *square-root* plot, which gives the quadratic region a chance to stretch out its limbs. The extrapolation to zero drain current defines the threshold voltage  $V_{th}$ . At that voltage the current is not zero – it's just down near the top of the subthreshold region.

#### B. The subthreshold region

Finally, the *log* plot greatly expands the low-current region. The curves bend over toward a straight-line (thus exponential) behavior at the lowest currents – this is the subthreshold region, which is explored six orders of magnitude more deeply in Figure 3.52.

#### C. The deep subthreshold region

From the extended curves in Figure 3.52 we can see that JFETs continue to do their stuff down to *picoamp* drain currents, conforming accurately to an exponential drain-current law (analogous to the BJT's Ebers–Moll), which can be written as

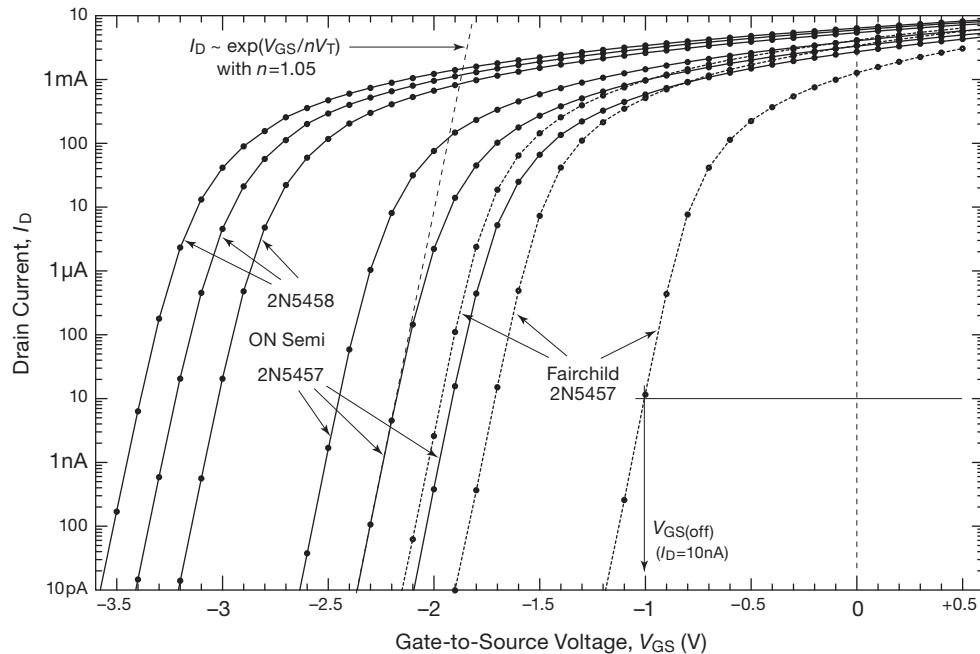
$$I_D = I_0 \exp(V_{GS}/nV_T), \quad (3.12)$$

with the same  $V_T=kT/q \approx 25\text{ mV}$  as in Ebers–Moll (but with an added fudge factor  $n$ ). The measured data in Figure 3.52 corresponds to a near-unity value for  $n$  ( $n = 1.05$ ). In other words, at very low drain currents, a JFET has nearly the same transconductance as a BJT operating at that same collector current (as confirmed in the measured transconductance data of Figure 3.54 on page 168).<sup>58</sup> Note that the Fairchild parts (in a gesture of egalitarian spirit) are no longer anomalous – unlike their behavior in the quadratic region (where they were the transconductance champs), at low currents they exhibit the same transconductance (slope) as the other parts.

#### 3.3.2 Drain current versus drain-source voltage: output conductance

Drain current (at constant  $V_{GS}$ ) *does* have some dependence upon drain-to-source voltage, in contrast to the idealized picture that you usually see (and to the promulgation of which we are guilty parties: see §3.1.4 and Figure 3.13). You can think of this effect as analogous to Early effect in BJTs (§2x.8), or, equivalently, as a finite output impedance  $r_o$  (or, more commonly, finite output conductance  $g_{os}=1/r_o$ ) seen at the drain terminal when the gate-source voltage is held constant. This effect limits the maximum gain of a grounded-source amplifier configured with

<sup>58</sup> But not much bandwidth! For instance, at  $I_D=10\text{pA}$  a 2N5457 would have a paltry gain-bandwidth product  $f_T$  of just 140 Hz ( $f_T=g_m/2\pi C_{iss}$ ; see “Bandwidth of the Cascode, BJT vs FET” in Chapter 3x). It's nice to know that JFETs (like BJTs) work well at very low currents, a useful characteristic for micropower and nanopower applications. But don't forget that device capacitances loom large at low currents, so current-starved designs are slower than circuits running at normal currents.



**Figure 3.52.** Log plot of measured drain currents (dots) versus gate voltage for the same JFETs as in Figure 3.51. JFETs work fine at 10 pA, far below the arbitrary 10 nA drain current that conventionally defines the gate-source cutoff voltage  $V_{GS(\text{off})}$ .

a current source as drain load (that's the  $G_{\max}$  parameter in Table 3.1 on page 141), and you need to take it into account if your gain approaches  $G_{\max}$ . It also further degrades the performance of a source follower (already degraded by low  $g_m R_L$ ; see §3.2.6) if  $G_{\max}$  is comparable to or less than  $g_m R_L$ . These consequences are discussed in detail in Chapter 3x (§3x.4); but they are important enough to be worthy of mention here.

### A. Degraded gain and linearity in the common-source amplifier

This “ $g_{os}$  effect” limits the maximum gain of the common-source amplifier, by effectively putting a resistance  $r_o = 1/g_{os}$  across the drain load impedance. For a simple resistive drain load  $R_D$  that reduces the gain from the ideal  $G = g_m R_D$  to  $G = g_m (R_D \parallel r_o)$ , or

$$G = g_m R_D \frac{1}{1 + g_{os} R_D} \quad (3.13)$$

A further unwelcome consequence is some nonlinearity, owing to the dependence of  $g_{os}$  upon drain voltage. See §3x.4 if you want to know more about this.

### B. Gain error in the source follower

The “ $g_{os}$  effect” also acts to reduce the gain of the source follower from the ideal  $G=1$ . This is most noticeable

with relatively light loads, where  $R_L \gg 1/g_m$  (and therefore where you would expect the voltage gain to be very close to unity). Taking this effect into account, the voltage gain of the simple JFET follower of Figure 3.40 becomes

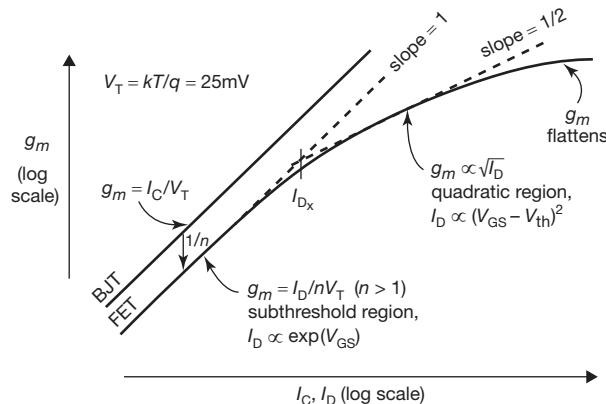
$$G = \frac{1}{1 + \frac{1}{g_m R_L} + \frac{1}{G_{\max}}} \quad (3.14)$$

where  $G_{\max}$  is the ratio of drain conductance to transconductance ( $G_{\max} = g_{os}/g_m$ ) at the operating voltage and current. (We defined  $G_{\max}$  because it's roughly constant with current<sup>59</sup> and therefore more useful than the current-dependent  $r_o$  and  $g_{os}$ ). Thus, for example, you need both  $g_m R_L$  and  $G_{\max}$  greater than 100 for <1% gain error – or you need to use some tricks, such as an active load, a cascode connection, or a “ $g_m$  enhancer.” Looking back to the tutorial source follower progression in Figure 3.43, that is the reason for the use of an active load and a cascode connection in circuit H. Table 3.1 lists measured values of  $G_{\max}$  for common JFETs. For further enlightenment read the exposé in §3x.4.

<sup>59</sup> Looking more closely, both  $g_m$  and  $g_{os}$  depend on drain current, approximately proportional to  $\sqrt{I_D}$  (§§3.3.3, 3x.2); their ratio  $G_{\max}$  is relatively flat with drain current, but approximately proportional to drain-source voltage (§§3.3.4, 3x.4).

### 3.3.3 Transconductance versus drain current

Recall that transconductance ( $g_m \equiv i_d/v_{gs}$ , the change of drain current with gate-to-source voltage) is the natural measure of FET gain. It is analogous to the Ebers-Moll view of bipolar transistors, where  $g_m = qI_C/kT = I_C/V_T$ . Generally speaking, at any given drain current an FET will have lower transconductance than a BJT. Starting at the lowest drain currents (recall §3.3.1 and Figure 3.52), the transconductance in the FET subthreshold region is proportional to drain current (like a BJT), but somewhat less than a BJT operating at the same current; see Figure 3.53. If you write  $g_m = I_D/nV_T$ , typical JFETs have  $n$  values in the range of 1.05 to 3, as can be seen in the measured data of Figure 3.54 (and similar expanded plots in Chapter 3x, see §3x.2).



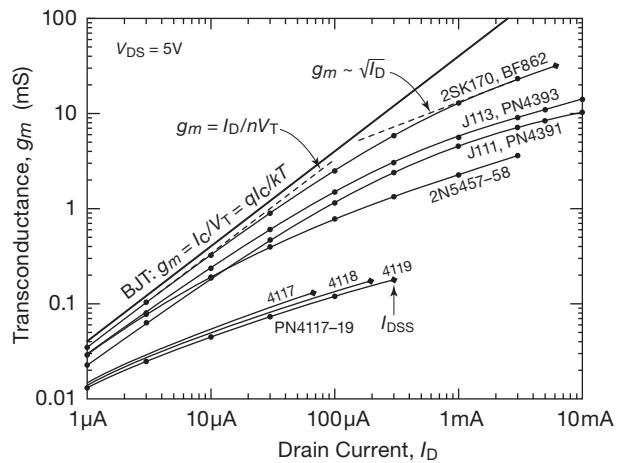
**Figure 3.53.** Transconductance versus drain current: in the subthreshold region an FET's transconductance is proportional to current (like a BJT); at normal currents where  $I_D \propto (V_{GS} - V_{th})^2$  it goes like  $\sqrt{I_D}$ ; then at still higher operating currents it flattens toward a maximum.

At higher currents, in the FET's “quadratic region” of drain current where  $I_D \propto (V_{GS} - V_{th})^2$ , the transconductance becomes proportional to the square root of drain current, falling away below the BJT's increasing ( $\propto I_C$ ) transconductance. At still higher currents the transconductance flattens further. For JFETs the story ends here; but for MOSFETs (where you can bring the gate voltage high enough to reach constant maximum drain current), the transconductance heads back down, for example with the LND150 whose  $I_D$  versus  $V_{GS}$  curve is plotted in Figure 3.9.

It's not hard to measure transconductance directly – see §3x.3 for circuit details of one simple method (where the secret sauce is a cascode stage to clamp the drain voltage). We used that circuit to measure the transconductance of

some sixty JFETs, at drain currents going from  $1\ \mu\text{A}$  to  $30\ \text{mA}$  in half-decade steps.

Figure 3.54 shows measured transconductance versus drain current, for a few representative JFETs.<sup>60</sup> At higher currents the transconductance is approximately proportional to the square root of drain current (thus proportional to  $V_{GS} - V_{th}$ ). And, even though these curves go down “only” to  $1\ \mu\text{A}$ , most of the JFETs (the four upper curves) are already well into the sub-threshold region (with transconductance proportional to drain current) down at that current.



**Figure 3.54.** Measured low-frequency transconductance for some representative  $n$ -channel JFETs. See also the analogous plot in §3x.2.

So, what's going on with the PN4117–19? Well, these are very small JFETs (you can tell by their very low  $I_{DSS}$  values), so even at a drain current of  $1\ \mu\text{A}$  the current *density* is high enough to bring them over into their quadratic region. Put another way, in the preceding discussion we should have been referring to “current density” (rather than simple drain current) as the measure of whether a given JFET is in its subthreshold region or its quadratic region.

Note that this holds also for MOSFETs: a large power MOSFET (with drain current ratings in the hundreds of amps) will be running in its subthreshold region if used in a linear circuit application.<sup>61</sup> As we'll see in §3.6.3, this has an important effect on thermal stability when multiple MOSFETs are connected in parallel in power applications, because in the subthreshold region MOSFETs have a *positive* thermal coefficient of drain current with temperature;

<sup>60</sup> See the rich thicket of measured curves in the analogous plot in §3x.2.

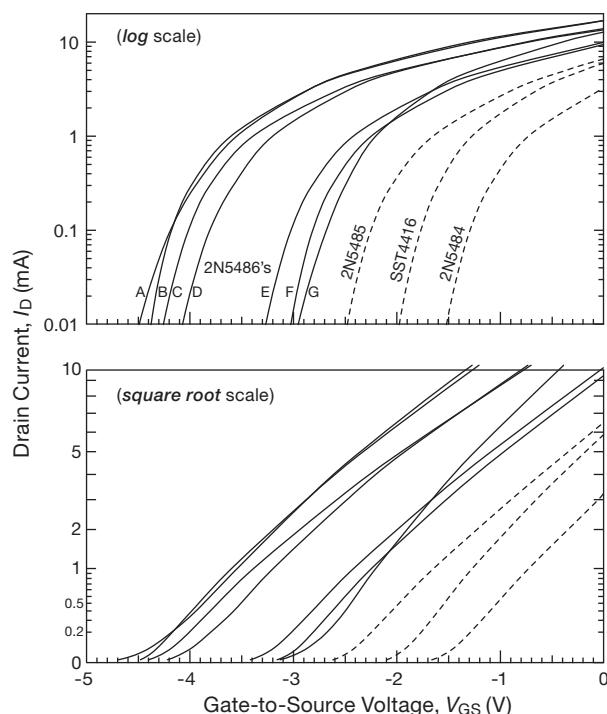
<sup>61</sup> SPICE-lovers beware: the simulation models for power MOSFETs are nearly useless in the subthreshold region.

this can cause serious mischief when not fully appreciated by the circuit designer.

### A. Transconductance within a JFET family

From Figure 3.54 and the preceding discussion it would appear that you cannot really predict a JFET's transconductance at any given operating current; that graph shows a 50:1 variation of  $g_m$  at a given current. Worse still, the data of Figure 3.51 suggests that you cannot even predict with any reasonable certainty the operating current of a given JFET part type for a given gate bias (or vice-versa).

As it turns out, the situation is not so bleak. Within a family (or extended family) of similar JFETs the transconductance depends primarily (and predictably) upon drain current, even though the corresponding gate voltage may be all over the map. Take a look at Figure 3.55, which shows measured drain current curves for a varied selection of seven 2N5486 JFETs (manufactured over a period of



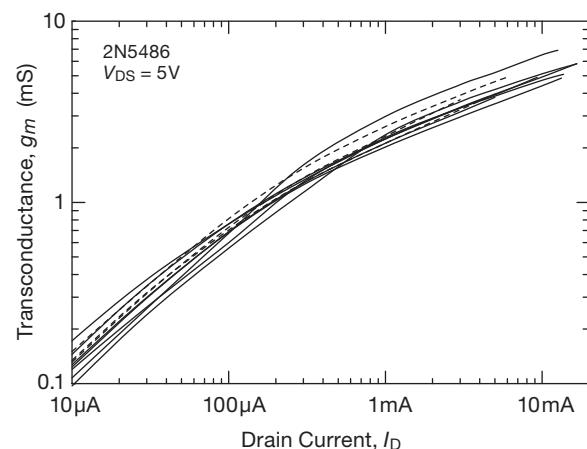
**Figure 3.55.** Measured drain current versus gate-source voltage for a collection of seven 2N5486 JFETs of varying vintages and manufacturers (along with three smaller JFET types, indicated with dashed lines), plotted on both log and square-root axes, illustrating the wide spread of threshold voltages  $V_{th}$  (the straight-line extrapolation of the latter's curves). Compare the large spread seen here with Figure 3.56, where the same set of JFETs exhibit only minor variations in transconductance at any given drain current.

35 years<sup>62</sup>), along with a sample of lower-current siblings (2N5484 and 2N5485) and a related cousin (SST416, also featured for RF applications, and with similar low capacitance and comparable range of  $I_{DSS}$ ).

These measured curves exhibit a wide variation of behavior for this selection of parts (there's a 3:1 spread of  $V_{GS(off)}$ , and a 5:1 spread of  $I_{DSS}$ ), suggesting a similar unpredictability of transconductance gain.<sup>63</sup> But that suggestion would be misleading. When you measure their transconductances, you get the curves of Figure 3.56: within this deliberately varied selection of ten related parts there's at most a  $\pm 20\%$  peak variation of transconductance at a given drain current.

The moral: within a JFET family of *similar part types*, the drain current (at whatever gate voltage is needed to get that current) is reasonably predictive of the transconductance. And the practical consequence is that a JFET amplifier circuit will have reasonably predictable gain if it is arranged with bias feedback to set the drain current to a desired value (which should be chosen no greater than the specified minimum  $I_{DSS}$ , unless you are willing to sort your incoming parts).

But a piece of advice: for better gain predictability in JFET amplifiers it's a good idea (a) to use some source degeneration (§3x.4), preferably in combination with a



**Figure 3.56.** Measured transconductance versus drain current for the same set of JFETs as in Figure 3.55.

<sup>62</sup> A, Intersil, date code 7328 (28th week of 1973); B, Central Semiconductor, contemporary; C and E, Fairchild, date code BF44; D, Vishay SST5486, contemporary; F, Vishay, date code 0536; G, Motorola, 1990 vintage.

<sup>63</sup> There are some helpful correlations, though: the gate threshold voltage  $V_{th}$  is predictive of the zero-bias drain current  $I_{DSS}$ .

transconductance-boosting enhancer circuit, or (b) to use overall feedback with enough minimum loop gain (with worst-case  $g_m$ ) to ensure the gain accuracy you need (as in Figures 3.31, 3.34, and 3.37).

### 3.3.4 Transconductance versus drain voltage

The transconductance of a FET at a given drain current is relatively independent of drain *voltage*, except for drain-to-source voltages below a volt or two. This contrasts with its strong dependence on drain *current*. See the discussion in Chapter 3x (§3x.2, with measured data).

### 3.3.5 JFET capacitance

Just as with bipolar transistors (and, later, MOSFETs), the capacitance seen between the terminals of a JFET depends on the (reverse) bias; this is often called a “nonlinear capacitance” and decreases markedly with increasing reverse-bias. Figure 3.57 shows datasheet plots of feedback and input capacitances for two common *n*-channel JFETs. These values – just a few picofarads – are typical of small-signal JFETs and quite a bit smaller than you see in power MOSFETs (compare with Figure 3.100 on page 197). JFETs are usually symmetrical, but owing to the larger reverse bias, the gate-drain capacitance is smaller than the gate-source capacitance.<sup>64</sup> That’s a good thing, of course, because the Miller effect multiplies the effect of  $C_{rss}$  in a common-source amplifier without cascode.

Manufacturers generally provide plots of capacitance versus  $V_{GS}$ , and sometimes also versus  $V_{DS}$  or  $V_{DG}$ . But you get only one set of plots for a family of parts, even though their  $I_{DSS}$  specs may range over a factor of ten or more; one is left wondering if their capacitances are correlated with their  $I_{DSS}$  values, thus suggesting that capacitance data should be considered only a rough guide to reality.

### 3.3.6 Why JFET (versus MOSFET) amplifiers?

We’ve devoted considerable space in this chapter to *junction* FETs (JFETs), a topic that often receives minimal discussion in standard references, in which the emphasis is on insulated-gate FETs (i.e., MOSFETs). We’ll soon shift the focus toward the latter, for good reasons: small integrated MOSFETs are dominant in (a) low-voltage and low-power analog circuits (op-amps, portable electronics, RF circuits,

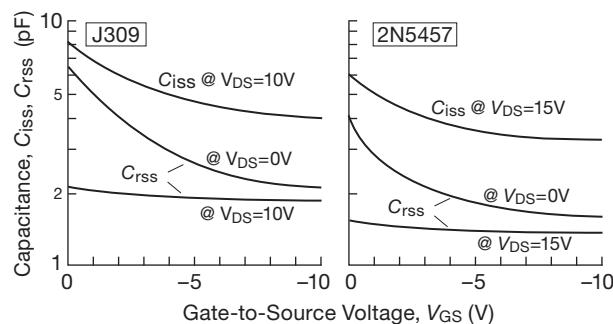


Figure 3.57. Input and feedback capacitance of JFETs.

and the like); (b) analog switching; (c) logic circuits, microprocessors, and memory; and, in the form of discrete packaged power transistors, (d) power switching and linear power applications. These are all hot topics in contemporary electronics, and MOSFETs are the dominant species of transistor on the planet, by a huge margin.

Before moving on, though, it is worth emphasizing that JFETs are the device of choice for analog circuits that demand a combination of high input impedance, low noise, and good accuracy; this in addition to their niche applications as two-terminal current sources, voltage-variable resistors, and constant- $R_{ON}$  analog switches.

The previous examples of amplifiers, followers, and current sources illustrate many of the JFET’s virtues. Going beyond this first look at JFETs, Chapter 3x includes some advanced material, and Chapter 8 discusses the important topic of noise, complete with measured data. To get a preview of the latter, look at Figure 3.58, where we’ve plotted the measured noise voltage density of three favorite “jellybeans”: the 2N3904 *npn* bipolar transistor, the 2N5457

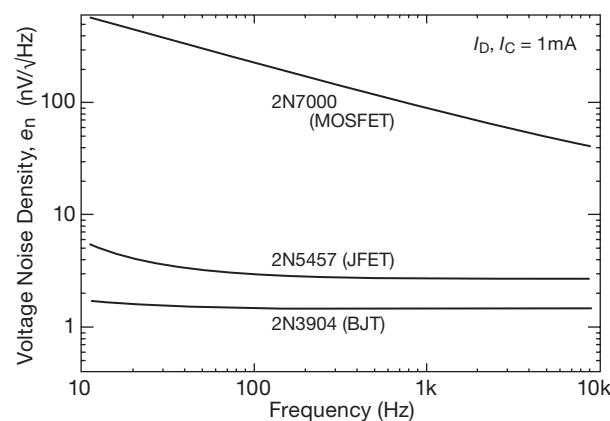


Figure 3.58. Noise voltage of three popular jellybeans, illustrating the poor low-frequency noise properties of MOSFETs.

<sup>64</sup> By “gate” and “source” we mean the pins your circuit uses for those functions, not the actual labels on the pins.

JFET and the 2N7000 MOSFET. These parts make no pretense of premium membership; they cost a pittance,<sup>65</sup> and they're not aimed at low-noise applications. But they illustrate an important trend: MOSFETs are inherently noisy at low frequencies, by as much as 40 dB relative to their BJT and JFET cousins.<sup>66</sup> Don't use them for low-level audio applications – but power MOSFETs can be used for the muscular output stage of an audio amplifier.

### 3.4 FET switches

The two examples of FET circuits that we gave at the beginning of the chapter were both *switches*: a logic-switching application and a linear signal-switching circuit. These are among the most important FET applications, and they exploit the FET's unique characteristics: high gate impedance and bipolar resistive conduction clear down to zero volts. In practice you usually use MOSFET integrated circuits (rather than discrete transistors<sup>67</sup>) in all digital logic and linear-switch applications, and it is only in power switching applications that you ordinarily resort to discrete FETs. Even so, it is essential (and fun!) to understand the workings of these chips; otherwise you're almost guaranteed to fall prey to some mysterious circuit pathology.

#### 3.4.1 FET analog switches

A common use of FETs, particularly MOSFETs, is as analog switches. Their combination of low ON-resistance (all the way to zero volts), extremely high OFF resistance, low leakage currents, and low capacitance makes them ideal as voltage-controlled switch elements for analog signals. An ideal analog, or linear, switch behaves like a perfect mechanical switch: in the ON state it passes a signal through to a load without attenuation or nonlinearity; in the OFF state it is an open circuit. It should have negligible capacitance to ground and negligible coupling from the signal applied to the control input.

Let's look at an example (Figure 3.59).  $Q_1$  is an *n*-channel enhancement-mode MOSFET, and it is nonconducting when the gate is grounded or negative. In that state the drain-source resistance ( $R_{\text{off}}$ ) is typically more

<sup>65</sup> Roughly \$0.02, \$0.05, and \$0.04, respectively, in 1000-piece quantities.

<sup>66</sup> According to John Willison, this phenomenon may be associated with intermittent trapping and release of charges on the insulated gate.

<sup>67</sup> It's hardly your choice – you are *forced* to, because discrete small-signal MOSFETs are a vanishing breed (perhaps qualifying for protected species status).

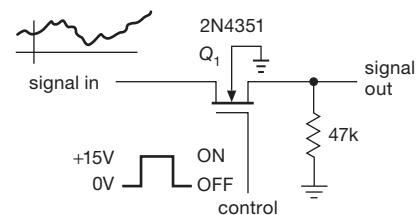


Figure 3.59. nMOS analog switch, with body terminal and diode shown.

than 10,000 M, and no signal gets through (though at high frequencies there will be some coupling through drain-source capacitance; more on this later). Bringing the gate to +15 V puts the drain-source channel into conduction, typically 20–200  $\Omega$  in FETs intended for use as analog switches. The gate signal level is not at all critical, as long as it is sufficiently more positive than the largest signal (to maintain  $R_{\text{ON}}$  low), and it could be provided from digital logic circuitry (perhaps using a FET or BJT to generate a full-supply swing) or even from an op-amp running from a +15 V supply. Swinging the gate negative (as from a bipolarity op-amp output) doesn't hurt, and in fact has the added advantage of allowing the switching of analog signals of either polarity, as will be described later. Note that the FET switch is a bidirectional device; signals can go either way through it. Ordinary mechanical switches work that way, too, so it should be easy to understand.

The circuit as shown will work for positive signals up to about 10 V; for larger signals the gate drive is insufficient to hold the FET in conduction ( $R_{\text{ON}}$  begins to rise), and negative signals would cause the FET to turn on with the gate grounded (it would also forward-bias the channel-body junction). If you want to switch signals that are of both polarities (e.g., signals in the range -10 V to +10 V), you can use the same circuit, but with the gate driven from -15 V (OFF) to +15 V (ON); the body terminal should then be tied to -15 V.

With any FET switch it is important to provide a load resistance in the range of 1k to 100k in order to reduce capacitive feedthrough of the input signal that would otherwise occur during the OFF state. The value of the load resistance is a compromise: low values reduce feedthrough, but they begin to attenuate the input signal because of the voltage divider formed by  $R_{\text{ON}}$  and the load. Because  $R_{\text{ON}}$  varies over the input signal swing (from changing  $V_{\text{GS}}$ ), this attenuation also produces some undesirable nonlinearity. Excessively low load resistance appears at the switch input, of course, loading the signal source as well. Several possible solutions to this problem (multiple-stage switches,  $R_{\text{ON}}$

cancellation) are shown in §3.4.2 and in the discussion of rail-to-rail op-amps in Chapter 4x. An attractive alternative is to use a second FET switch section to connect the output to ground when the series FET is off, thus effectively forming a single-pole double throw (SPDT) switch (more on this in the next section).

### A. CMOS linear switches

Frequently it is necessary to switch signals that may go nearly to the supply voltages. In that case the simple *n*-channel switch circuit just described won't work, because the gate is not forward-biased at the peak of the signal swing. The solution is to use paralleled complementary MOSFET (CMOS) switches (Figure 3.60). The triangular symbol is a digital inverter, which we'll discuss shortly; it inverts a HIGH input to a LOW output, and vice versa. When the control input is high,  $Q_1$  is held ON for signals from ground to within a few volts of  $+V_{DD}$  (where  $R_{ON}$  starts increasing dramatically).  $Q_2$  is likewise held ON (by its grounded gate) for signals from  $+V_{DD}$  to within a few volts of ground (where its  $R_{ON}$  increases dramatically). Thus signals anywhere between  $+V_{DD}$  and ground are passed through with low series resistance (Figure 3.61). Bringing the control signal to ground turns off both FETs, providing an open circuit. The result is an analog switch for signals all the way from ground to  $+V_{DD}$ . This is the basic construction of the classic 4066 CMOS "transmission gate." It is bidirectional, like the switches described earlier: either terminal can be the input.

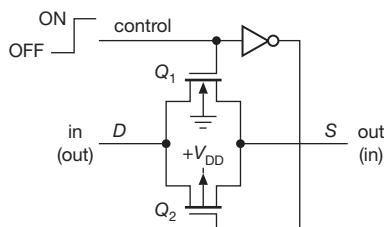


Figure 3.60. CMOS analog switch.

There is a variety of integrated circuit CMOS analog switches available, with different operating voltage ranges and with various switch configurations (e.g., several independent sections with several poles each). Going back to the prehistoric, there's the classic CD4066 "analog transmission gate," which belongs to the original 4000 series of digital CMOS logic and that acts as an analog switch for signals between ground and a single positive supply.<sup>68</sup>

<sup>68</sup> In that role it's happy to switch digital signals as well, thus its family membership.

More commonly, though, you'll choose a dedicated analog switch IC, for example a member of the industry-standard DG211 family. These parts (and their many variations; see Table 3.3 on page 176) are particularly convenient to use: they accept logic-level ( $0\text{V} = \text{LOW}$ ,  $>2.4\text{V} = \text{HIGH}$ ) control signals, they will handle analog signals to  $\pm 15\text{V}$  (compared with only  $\pm 7.5\text{V}$  for the 4000 series), they come in a variety of configurations, and they have relatively low ON resistance ( $25\Omega$  or less for some members of these families, and down to a fraction of an ohm for lower voltage switches). Analog Devices, Intersil, Maxim, Vishay–Siliconix, and other manufacturers offer a nice range of analog switches.

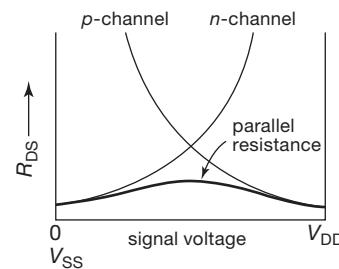


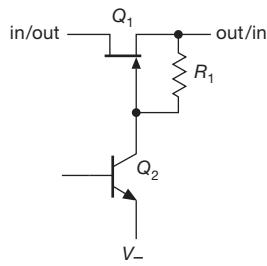
Figure 3.61. ON-resistance of the CMOS analog switch.

### B. JFET analog switches

Although most available analog switches are built with a pair of paralleled MOSFETs of complementary polarities – the CMOS architecture just described – it is possible to build analog switches with JFETs, and there are some advantages.

The basic circuit (Figure 3.62) uses a single *n*-channel JFET,  $Q_1$ , as the analog switch. Its conduction is controlled by a transistor switch,  $Q_2$ , which pulls the gate down to a large negative voltage ( $-15\text{V}$ , say) to cut off conduction in the JFET (switch OFF). Shutting off  $Q_2$  lets the gate float to the source voltage, putting the (depletion-mode) JFET into full conduction (switch ON). The gate resistor  $R_1$  is made deliberately large so that the output signal line does not see significant loading in the OFF state; its value is a compromise, because a larger resistance incurs a longer turn-on delay. With a low-impedance signal source (e.g., the output of an op-amp) it may be preferable to put the resistor on the input side (i.e., feed the signal in from the right).

Because there is only a single *n*-channel JFET, this switch cannot accept input signals all the way to the negative supply: signal voltages that come closer to the negative

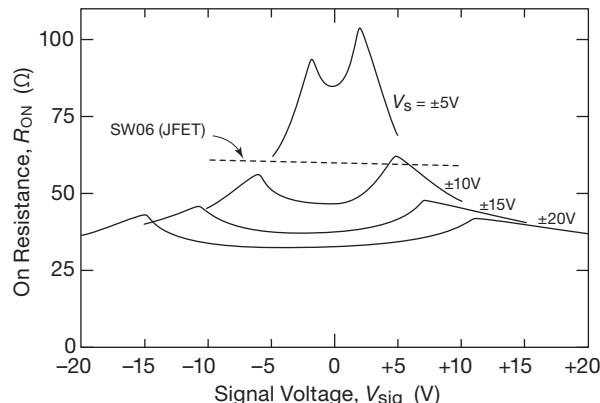


**Figure 3.62.** JFET analog switch.

supply than the FET's  $V_{GS(\text{off})}$  will bring it back into conduction.<sup>69</sup> There is no such limitation at the positive end.

A nice feature of this JFET analog switch is the constancy of  $R_{\text{ON}}$  with signal level: because the gate stays at the source voltage, there is *no* variation of  $R_{\text{ON}}$  with signal voltage; the JFET does not even know that the signal is varying! This nice characteristic is on display in Figure 3.63, which compares a plot of  $R_{\text{ON}}$  versus  $V_{\text{sig}}$  for a JFET analog switch (SW06) with that of the DG211 CMOS switch.<sup>70</sup>

In practice it's inconvenient to have to drive the switch control with a signal near the negative supply rail (and dou-

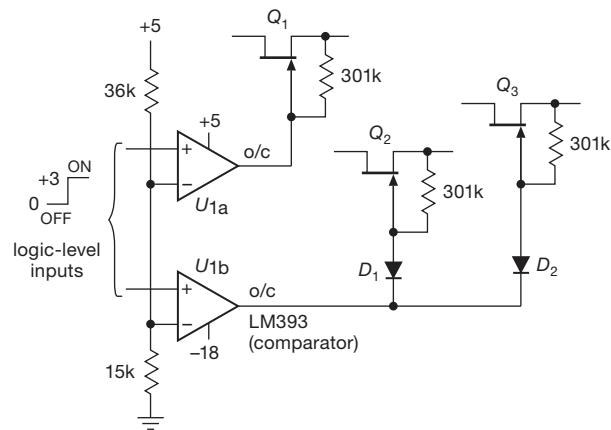


**Figure 3.63.** Contrasted with a CMOS analog switch like the DG211 (solid curves, shown for different supply voltages), the JFET analog switch maintains admirably flat  $R_{\text{ON}}$  over signal level.

<sup>69</sup> In fact, you've got to stay a bit further away from the actual  $V_{GS(\text{off})}$ : recall (§3.1.3) that it's defined as the gate-source voltage that results in a small (but nonzero) drain current, usually  $I_D = 10 \text{ nA}$ .

<sup>70</sup> The slight variation of  $R_{\text{ON}}$  is caused by substrate effects: the SW06 is an integrated circuit, built on a silicon substrate; so the JFET and associated components have some inkling of the absolute signal level. If even that small variation is unacceptable, you can rig up a discrete implementation (as in Figure 3.62); you'll see such circuitry used in some of Agilent's accurate digital multimeters.

bly so for CMOS switches). Instead, you would probably use a level-shifting circuit so a logic-level input that goes between 0 V and +3 V, say, would activate the switch. Figure 3.64 shows a simple way to do that, using comparators with “open-collector” outputs (§12.3) to drive the gates of discrete JFET switches. For *integrated* analog switches, this sort of control signal circuitry is ordinarily built in. The SW06 JFET analog switch (and the nearly extinct DG180–189 family from Vishay–Siliconix) includes such drivers, along with some other elegant tricks.<sup>71</sup>



**Figure 3.64.** A comparator with open collector outputs, powered from +5 V and -18 V, converts a 0-to-3 V logic-level input into a JFET gate drive that swings to -18 V. This method is used in some of Agilent's digital multimeters (see “Designs by the Masters,” 13.8.6) to accommodate analog signals over a full  $\pm 12 \text{ V}$  range. The added diodes allow a control signal to drive more than one JFET switch.

JFET analog switches are inherently more rugged than their CMOS cousins, for which performance-degrading protection circuits are required for robustness against overvoltage faults. They do suffer from high charge injection, however (see §3.4.2E). Despite their nice features, integrated JFET switches and multiplexers (see next subsection) are almost extinct, with fine examples like the SW-01, SW-7510, and MUX-08 series from Precision Monolithics (now Analog Devices) gone forever (but happily the SW06 lives on!).

## C. Multiplexers

A nice application of FET analog switches is the “multiplexer” (or MUX), a circuit that allows you to select any of several inputs, as specified by a digital control signal.

<sup>71</sup> Such as an internal MOSFET switch that disconnects the gate resistor ( $R_1$  in Figure 3.62) when the switch is OFF, to eliminate circuit loading.

The analog signal present on the selected input will be passed through to the (single) output. Figure 3.65 shows the basic scheme. Each of the switches SW0 through SW3 is a CMOS analog switch. The “select logic” decodes the address and *enables* (jargon for “turns on”) the addressed switch only, disabling the remaining switches. Such a multiplexer is usually used in conjunction with digital circuitry that generates the appropriate addresses (lots more in Chapters 10 and 11). A typical situation might involve a data-acquisition instrument in which a number of analog input voltages must be sampled in turn, converted to digital quantities, and used as input to some computation.

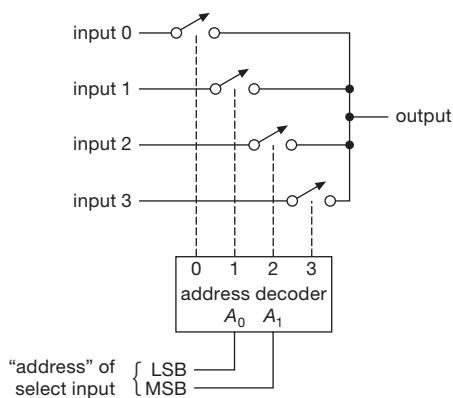


Figure 3.65. Analog multiplexer.

Because analog switches are bidirectional, an analog multiplexer such as this is also a “demultiplexer”: a signal can be fed into the “output” and will appear on the selected “input.” When we discuss digital circuitry in Chapters 10 and 11, you will see that an analog multiplexer such as this can also be used as a “digital multiplexer–demultiplexer,” because logic levels are, after all, nothing but voltages that happen to be interpreted as binary 1s and 0s.

Typical of analog multiplexers are the industry-standard DG408-09 and DG508-09 series (and their many improved versions), 8- or 16-input MUX circuits that accept logic-level address inputs and operate with analog voltages up to  $\pm 15$  V. The 4051-4053 devices in the CMOS digital family are analog multiplexers–demultiplexers with up to 8 inputs, but with 15 Vpp maximum signal levels; they have a  $V_{EE}$  pin (and internal level shifting) so that you can use them with bipolarity analog signals and unipolarity (logic-level) control signals. We are especially fond of the low-voltage '4053 series, with three SPDT switches. Our interest is evidently shared by others, with a large number of interesting parts available – see Table 13.7 (4053-style SPDT Switches) on page 917.

## D. Other analog switch applications

Voltage-controlled analog switches form essential building blocks for some of the op-amp circuits we’ll see in the next chapter – integrators, sample-and-hold circuits, and peak detectors. For example, with op-amps we will be able to build a “true” integrator (unlike the approximation to an integrator we saw in §1.4.4): a constant input generates a linear ramp output (not an exponential), etc. With such an integrator you must have a method to reset the output; a FET switch across the integrating capacitor does the trick. We won’t try to describe these applications here. Because op-amps form essential parts of the circuits, they fit naturally into the next chapter. Great things to look forward to!

### 3.4.2 Limitations of FET switches

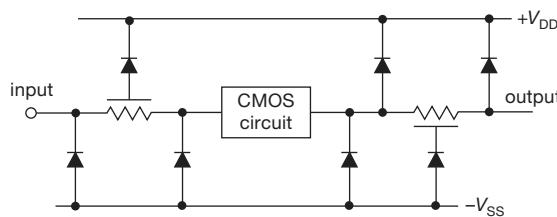
Analog switches aren’t perfect – they have nonzero resistance when ON, and nonzero leakage when OFF, as well as capacitive feedthrough and charge injection during changes in switch state. You can see some of the variety in Table 3.3 on page 176. Let’s take a look at these limitations.

#### A. Voltage range and latchup

Analog switches and multiplexers come in three broad voltage ranges: (a) “standard” parts (which might better be called “high-voltage”), which handle signals over the full traditional op-amp voltage range of  $\pm 15$  V; (b) reduced-voltage parts (“midvoltage”) that can handle  $\pm 7.5$  V (or 0 to  $+15$  V) signals; and (c) low-voltage parts intended for applications in which the signal swing does not exceed  $\pm 3$  V (or 0 to  $+6$  V). In all cases, analog switches operate properly (and with specified  $R_{ON}$ ) with input signals that go all the way to the positive and negative supply voltages (with the exception of JFET switches like the SW06, for which the operating signal voltage range does not reach the positive supply).

However, input signals *beyond* the supply rails are another story. All CMOS integrated circuits have some form of input protection circuit, because otherwise the gate insulation is easily destroyed (see §3.5.4H on handling precautions). The usual protection network is shown in Figure 3.66. Although it may use distributed diodes, the network is equivalent to clamping diodes to  $V_{SS}$  and to  $V_{DD}$ , combined with resistive current limiting. If you drive the inputs (or outputs) more than a diode drop beyond the supply voltages, the diode clamps go into conduction, making the inputs (or outputs) look like a low impedance to the respective supplies. Worse still, the chip can be driven into “SCR latchup,” a terrifying (and destructive)

condition we'll describe in more detail in §10.8. For now, all you need to know about it is that you don't want it! SCR latchup is triggered by input currents (through the protection network) of roughly 20 mA or more. Thus you must be careful not to drive the analog inputs as much as a diode drop beyond the rails.<sup>72</sup> This means, for instance, that for most parts you must be sure the power supply voltages are applied before any signals that have significant drive-current capability; alternatively, you can use series diodes in the supply lines, so that input signals applied before the dc power do not produce input current.



**Figure 3.66.** CMOS input–output protection networks. The series resistor at the output is often omitted.

The trouble with diode–resistor protection networks is that they compromise switch performance by increasing  $R_{ON}$ , shunt capacitance, and leakage. A different approach makes use of “dielectric isolation” to eliminate SCR latchup without the serious performance compromises inherent in traditional protection networks. Both methods result in a “protected” (or *fault-protected*) analog switch, in which you may safely overdrive the inputs without damage. Note, however, that the output does not follow the input beyond the rails.<sup>73</sup>

For example, the MAX4508 multiplexer adds fault protection to the standard DG508A 8-input analog multiplexer, making it tolerant to  $\pm 30$  V input swings; it has

<sup>72</sup> This prohibition goes for *digital* CMOS ICs as well as the analog switches we have been discussing.

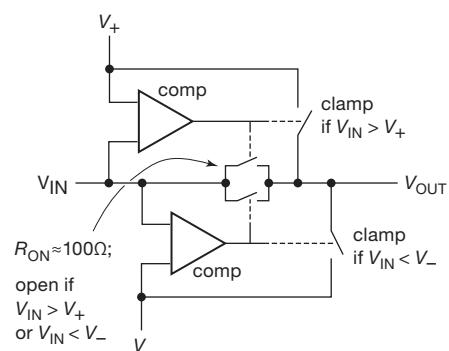
<sup>73</sup> With a few exotic exceptions, for example the MAX14778 “Dual  $\pm 25$  V Above- and Below-the-Rails 4:1 Analog Multiplexer.” This puppy, which runs on a single +3 to +5 V supply, is not only fault protected to  $\pm 25$  V, it operates properly with signal voltages over that full range! How do they do it?! Turns out they have included an on-chip “charge-pump” voltage converter (§3.4.3D) to power the guts. Even more remarkable, this device has a very low ON-resistance ( $1.5\ \Omega$ ) that is remarkably flat with signal voltage ( $0.003\ \Omega$ ) over the full  $\pm 25$  V range. And it can handle up to 300 mA of signal current. This device is intended for applications in which external large-swing signals need to be switched by a circuit with only a low-voltage single supply. Sadly for us experimenters, it comes only in the tiny TQFN (Thin Quad Flatpack, No leads) package, requiring a reflow oven for soldering onto a circuit board.

an  $R_{ON(\text{typ})}$  of  $300\ \Omega$ . The AD7510DI-series of “Protected Analog Switches” from Analog Devices uses dielectric isolation to achieve input signal fault protection to  $\pm 25$  V beyond the power supplies, while maintaining a respectable  $75\ \Omega R_{ON(\text{typ})}$  in the normal operating signal range. Watch out, though – fault protection is the exception in the analog switch arena, and most analog switch ICs are not forgiving!

Maxim offers a nice outboard solution that you can put in front of an unprotected switch (or any other analog component), in the form of its multichannel “signal-line protector” ICs (the three- and eight-channel MAX4506–07).<sup>74</sup> These accept input signal swings to  $\pm 36$  V (whether powered or unpowered), are free of latchup regardless of power sequencing, and they pass through signals that are properly within the supply limits while clamping their outputs to the supply rails (which can be  $\pm 8$  V to  $\pm 18$  V split supplies, or a +9 to +36 V single supply) when there's an input beyond the rails. They even have the grace to open-circuit the input when thus overdriven – see Figure 3.67. The price you pay (above their literal \$4 and \$6 cost, respectively) is an ON-resistance in the range of  $50$ – $100\ \Omega$  (depending as usual on total supply voltage) and an input capacitance of  $20\ pF$  (thus a rolloff at  $\sim 100$  MHz).

## B. ON-resistance

CMOS switches operated from a relatively high supply voltage ( $\pm 15$  V, say) will have low  $R_{ON}$  over the entire signal swing, because one or the other of the transmission FETs will have a forward gate bias of at least half the



**Figure 3.67.** Maxim's MAX4506–7 fault-protected “Signal-Line Protectors” prevent signal swings beyond the rails, both clamping the output and disconnecting the input when overdriven. The series analog switches are normally ON, but both are switched OFF if  $V_{IN}$  goes beyond either rail, controlled by logic shown in the more complete MAX4508 datasheet.

<sup>74</sup> Analog Devices offers a similar single-channel part, the ADG465, in a convenient SOT23-6 package.

Table 3.3 Analog Switches

Part # <sup>a</sup>	Config <sup>b</sup>				Packs										Comments	
	SPST (NO)	SPST (NC)	SPDT <sup>u</sup>	MUX <sup>v</sup>	Supply		R <sub>on</sub>	Q <sub>inj</sub>	Cap	Logic <sup>d</sup>	Control <sup>e</sup>	SOT-23	DIP	SOIC	other	
					split (±V)	single (+V)	typ @ V <sub>Supply</sub> (Ω)	typ (V)	typ (pC)	typ, ON (pF)						
<i>high voltage</i>																
MAX4800-02	-	-	-	8:1	40 to 100	c	22	±100	600	36	V <sub>L</sub>	S	-	-	28	16.18
MAX326-27	4	4	-	-	5 to 18	10 to 30	1500	±15	2	6	T	P	-	16	16	6.76
MAX4508-09	-	-	84	4.5 to 20	9 to 36		300	±15	2	28, 22	T	P	-	16	16	-
MAX354-55	-	-	84	4.5 to 18	4.5 to 36		285	±15	80	28, 14	T	P	-	16	16	-
<b>DG508-09</b>	-	-	84	5 to 20	10 to 36		180	±15	2	18, 11	T	P	-	16	16	2.56
ADG1211-13	w	w	-	-	5 to 15	10 to ?	120	±15	0.3	2.6	T	P	-	-	-	4.63
ADG1221-23	x	x	-	-	5 to 16.5	5 to 16.5	120	±15	0.1	3	T	P	-	-	-	3.04
AD7510-12DI	4	4	2	-	5 to 15	-	75	±15	30	17	T	P	-	16	-	13.90
SW06	4	-	-	-	12 to 18	-	60	±15	-	15	T	P	-	16	16	4.30
DG441-42	4	4	-	-	4.5 to 22	5 to 24	50	±15	2	16	V <sub>L</sub>	P	-	16	16	2.05
<b>DG211-12</b>	4	4	-	-	4.5 to 22	5 to 22	45	±15	1	16	V <sub>L</sub>	P	-	16	16	2.08
<b>DG408-09</b>	-	-	84	5 to 20	5 to 30		40	±15	20	37, 25	T	P	-	16	16	3.06
<b>ADG417-19</b>	1	1	1	-	5 to 20	5 to 20	25	±15	3	30	V <sub>L</sub>	P	-	8	8	2.72
<b>MAX317-19</b>	1	1	1	-	4.5 to 20	10 to 30	20	±15	3	30	V <sub>L</sub>	P	-	8	8	3.34
DG411-13	w	w	-	-	4.5 to 20	10 to 30	17	±15	5	35	V <sub>L</sub>	P	-	16	16	2.70
<b>DG447-48</b>	1	1	-	-	4.5 to 20	7 to 36	13	±15	10	30	T	P	-	-	6	-
ADG5412-13	y	y	-	-	9 to 22	9 to 40	10	±15	240	60	T	P	-	-	-	5.66
DG467-68	1	1	-	-	4.5 to 20	7 to 36	5	±15	21	76	T	P	-	-	8	-
<i>mid-voltage</i>																
DG4051-53	-	-	-	842	2.5 to 5	2.7 to 12	66	±5	0.25	3.4	V+	P	-	-	16	16
<b>74HC4051-53</b>	-	-	-	842	2.5 to 5	2 to 10.5	40	±5	5, 25	V+	P	-	16	16	0.41	
MAX4541-44	x	x	1	-	-	2.7 to 12	30	+5	1	13, 20	T	P	8	8	8	1.33
<b>ISL5120-23</b>	x	x	1	-	-	2.7 to 12	19	+5	3	28	T	P	8, 6	-	8	-
<b>ISL43210</b>	-	-	1	-	-	2.7 to 12	19	+5	3	28	T	P	6 <sup>s</sup>	-	-	1.33
<b>ADG619-20<sup>k</sup></b>	-	-	1	-	2.7 to 5.5	2.7 to 5.5	7	+5	6	95	T	P	8	-	-	2.56
<i>low-voltage</i>																
ADG708-09	-	-	-	84	2.5	1.8 to 5.5	3	+5	3	96, 48	T	P	-	-	16	-
<b>ADG719</b>	-	-	1	-	-	1.8 to 5.5	2.5	+5	-	27	T	P	6	-	-	1.76
MAX4624-25 <sup>k</sup>	-	-	1	-	-	1.8 to 5.5	0.65	+5	65	100	T	P	6	-	-	2.10
ADG884	-	-	2	-	-	1.8 to 5.5	0.3	+5	125	300	T	P	-	-	-	2.42
ISL84467	2xDPDT	-	-	-	-	1.8 to 5.5	0.55	+3	126	102	V+	P	-	-	16	16
ISL84714	-	-	1	-	-	1.8 to 3.6	0.44	+3	20	100	V+	P	6	-	-	1.24
NLAS52231	-	-	2	-	-	1.65 to 4.5	0.38	+3	53	85	V+	P	-	-	8	0.90
ISL43L110-11	1	1	-	-	-	1.1 to 4.7	0.25	+3	72	160	V+	P	5 <sup>s</sup>	-	-	0.65
<i>T-switch, RF</i>																
MAX4565-67	y	y	2	-	2.7 to 6	2.7 to 12	46	±5	25	6	T	P	-	16, 20	16, 20	4.74
DG540-42	z	z	-	-	+15 & -3	-	30	nom	25	14	T	P	-	16	16	5.00
AD8170, 74	-	-	1	4:1	4 to 6	-	NA	-	NA	1.1	T	P	-	8, 14	-	5.11 <sup>n</sup>
<b>ADG918-19</b>	-	-	1	-	-	1.65 to 2.75	-	-	-	1.6	V+	P	-	-	8	2.52
<i>crosspoint</i>																
AD75019	16x16	-	-	-	4.5 to 12	9 to 25	150	±12	-	10 m	T	S	-	-	44	26.20
ADG2188	8x8	-	-	-	4.5 to 5.5	8 to 12	34	±5	3	9.5	V <sub>L</sub>	I	-	-	32	9.37
MAX4359-60	4x4, 8x8	-	-	-	4.5 to 5.5	-	x1 buf	-	NA	8	T	P, ser	-	40	24	44
MAX9675	16x16	-	-	-	2.5 to 5.5	-	x1,x2 buf	-	(g)	5	V <sub>L</sub>	S	-	-	100	24.14
MAX4550, 70	dual 4x2	-	-	-	2.7 to 5.5	2.7 to 5.5	43	±5	7	11°	V+	S or I	-	2	-	28
																6.39
																audio <sup>10</sup>

**Notes:** (a) listed within categories by decreasing R<sub>on</sub>; all are CMOS except SW06; parts in **bold italic** are widely used “jellybeans.” (b) numerals represent quantity of each switch type in a single package; letters refer to explanatory footnotes for successively numbered parts. (c) V<sub>neg</sub> at least -15V, V<sub>pos</sub> at least +40V, total no more than 200V. (d) T=TTL thresholds; V<sub>L</sub>=external logic threshold supply; V<sub>+</sub>=“CMOS” threshold, depends on positive analog supply voltage. (e) P=parallel logic inputs; I=I<sup>2</sup>C serial; S=SPI serial. (g) 50mV glitch. (h) 0.1dB to 14MHz, -95dB xtalk at 20KHz & R<sub>L</sub>=10k, THD+N=0.014% (R<sub>L</sub>=1k, f=1kHz). (k) second p/n is make-before-break. (m) min or max. (n) higher p/n is ~50% more. (o) switch OFF. (s) SC-70. (u) SPDT are break-before-make unless noted otherwise. (v) 84=8:1 & dual 4:1; 842=8:1, dual 4:1, & triple 2:1. all dual 4:1 have single 2-bit address. (w) 4xNC, 4xNO, 2 each. (x) 2xNO, 2xNC, 1 each. (y) 4xNO, 2 each. (z) 4xNO, 4xNO, 2 each.

**Comments:** (1) Supertex HV2203. (2) 1pA typ. (3) 3pA typ; I<sub>s</sub>=10µA typ. MAX308-09 is similar. (4) 8kV HBM ESD. (5) 0.4Ω@V<sub>S</sub>=+3V; 400mA; -3dB/50Ω@18MHz. (6) low R<sub>on</sub>, e.g. speaker switch. (7) -3dB@350MHz, -90dB xtalk@10MHz. (8) 250 MHz; ext gain set. (9) -3dB at 4GHz, -30dB xtalk at 4GHz. (10) clickless.

supply voltage. However, when operated with lower supply voltages, the switch's  $R_{ON}$  value will rise, the maximum occurring when the signal is about halfway between the supply voltage and ground (or halfway between the supplies, for dual-supply voltages). Figure 3.68 shows why. As  $V_{DD}$  is reduced, the FETs begin to have significantly higher ON-resistance (especially near  $V_{GS} = V_{DD}/2$ ), because enhancement-mode FETs can have a  $V_{GS(th)}$  of at least a few volts, and a gate-source voltage of as much as 5 to 10 volts may be required to achieve low  $R_{ON}$ . Not only will the parallel resistances of the two FETs rise for signal voltages between the supply voltage and ground, but also the peak resistance (at half  $V_{DD}$ ) will rise as  $V_{DD}$  is reduced, and for sufficiently low  $V_{DD}$  the switch will become an open circuit for signals near  $V_{DD}/2$ .

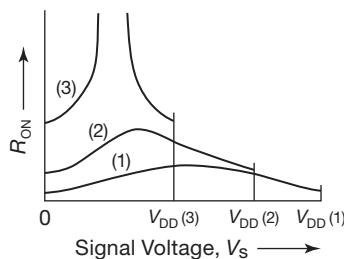


Figure 3.68. CMOS analog switch  $R_{ON}$  increases at low supply voltage.

There are various tricks used by the designers of analog switch ICs to keep  $R_{ON}$  low and approximately constant (for low distortion) over the signal swing. For example, the original 4016 analog switch used the simple circuit of Figure 3.60, producing  $R_{ON}$  curves that look like those in Figure 3.69. In the improved 4066 switch, the designers added a few extra FETs so that the  $n$ -channel body voltage follows the signal voltage, producing the  $R_{ON}$  curves of Figure 3.70. The “volcano” shape, with its depressed central

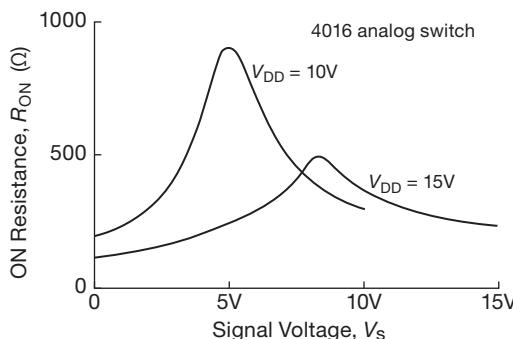


Figure 3.69. ON-resistance for 4016 CMOS switch.

$R_{ON}$ , replaces the “Everest” shape of the 4016. Improved switches, like the industry standard DG408-09, intended for serious analog applications, succeed even better, with low and flat  $R_{ON}$  curves that deviate less than 10% or so over the signal voltage range. This is often achieved at the expense of increased “charge transfer” (see the later section on *glitches*).

Looking through manufacturers' selection tables for analog switches, you'll find standard-voltage units with  $R_{ON}$  as low as several ohms and flatness to a few tenths of an ohm; low-voltage switches can be found with  $R_{ON}$  as low as  $0.25\ \Omega$ , and flatness of  $0.03\ \Omega$ . This static performance comes at a real cost though, namely, high capacitance and high charge injection (see discussion below, and Table 3.3 on the preceding page). If your application requires low distortion into moderate load impedances, the better approach may be to choose a switch with excellent “ON-resistance flatness” spec ( $R_{FLAT(ON)}$ ), and accept a higher overall  $R_{ON}$  with its lower capacitance.

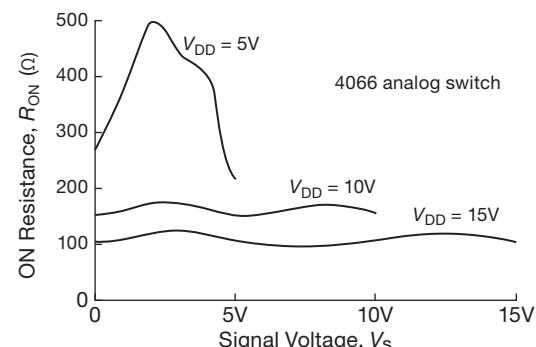
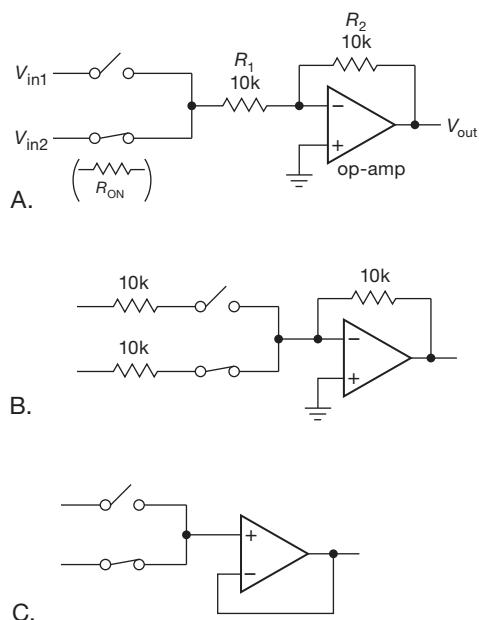


Figure 3.70. ON-resistance for the improved 4066 CMOS switch; note change of scale from previous figure.

Keep in mind, too, that in some cases you can finesse the problem altogether with a different choice of circuit configuration, as in Figure 3.71, which shows three approaches to a circuit that selects one of two input signals. Circuit A's gain is  $R_2/(R_1 + R_{ON})$ , so a variation of  $R_{ON}$  with signal amplitude produces changes of gain, and thus nonlinearity. Circuit B is better, because the switch output is held at ground by feedback around the op-amp; but the ON-resistance still reduces the gain somewhat, degrading the circuit's precision. Circuit C is blissfully unaware of  $R_{ON}$ , owing to the op-amp's very high input impedance; it's the most linear and precise of all.

This lesson can be applied to other circuit configurations as well. As an example, take a look at Figure 3.84 on page 183, where an analog multiplexer is used to select an amplifier's overall voltage gain. In the circuit of



**Figure 3.71.** Finessing  $R_{ON}$  variation in analog switches: three ways to select between a pair of input signals, with an op-amp to buffer the output.

Figure 3.84A the multiplexer's  $R_{ON}$  is in series with the selected resistor, and represents an error term (both in gain and nonlinearity); by contrast, in the circuit of Figure 3.84B the switch's  $R_{ON}$  is irrelevant, owing to the op-amp's essentially infinite ( $>10^{12}\Omega$ ) input impedance.

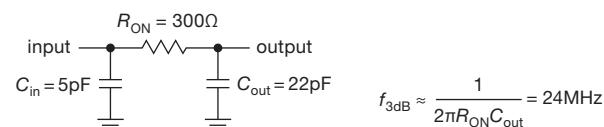
Another trick exploits the use of two identical (or closely-similar) JFET switches to largely cancel the effects of  $R_{ON}$ . See Chapter 4x's section “JFET linear switch with  $R_{ON}$  compensation” for an explanation and illustration of this elegant technique.

### C. Speed

High-voltage FET analog switches have ON-resistances  $R_{ON}$  generally in the range of 20 to 200  $\Omega$ .<sup>75</sup> In combination with substrate and stray capacitances, this resistance forms a lowpass filter that limits operating speeds to frequencies of the order of 10 MHz or less (Figure 3.72). FETs with lower  $R_{ON}$  tend to have larger capacitances (up to 50 pF or more), so no gain in speed results (unless the designer has made other design tradeoffs). Much of the rolloff is due to protection components – current-limiting series resistance and capacitance of shunt diodes.

However, low-voltage analog switches do better in terms of bandwidth (as is usually the case with smaller-geometry semiconductors): a contemporary  $\pm 2.5$  V analog switch like the popular ADG719 has  $2.5\ \Omega$  of ON-resistance,  $27\text{ pF}$  of capacitance, and  $400\text{ MHz}$  bandwidth. There is also a class of analog switches and multiplexers targeted specifically at video and RF applications. These include both passive (“unbuffered”) MUX/switches, and MUX/switches combined with an amplifier (“active” or “buffered”). Active MUX/switches operate on  $+5\text{ V}$  or  $\pm 5\text{ V}$  supplies and have fixed voltage gains of either  $\times 1$  or  $\times 2$  (the latter are intended for driving a  $50\ \Omega$  or  $75\ \Omega$  transmission line through a series matching resistor, which attenuates the output by a factor of 2); in some cases you set the gain with an external resistor pair. An example of the latter is the AD8174 4-input multiplexer, with a bandwidth of  $270\text{ MHz}$  at gains of  $+1$  or  $+2$  (at higher gains the bandwidth drops, e.g., to  $55\text{ MHz}$  at  $G=+10$ ).

For specialized applications you can get some really fast analog switches, for example the ADG918–19 listed in Table 3.3 on page 176, which is usable to 2 GHz (it's down 3 dB at 4 GHz). Parts like these are used in wireless applications, for example, to switch between two signal sources in “diversity reception” or to route gigahertz signals through a choice of filter paths. To reduce crosstalk, these wideband switches usually employ a T-switch topology (see Figure 3.77 in the next subsection).



**Figure 3.72.** The parasitic RC's of a CMOS switch limit the analog signal bandwidth.

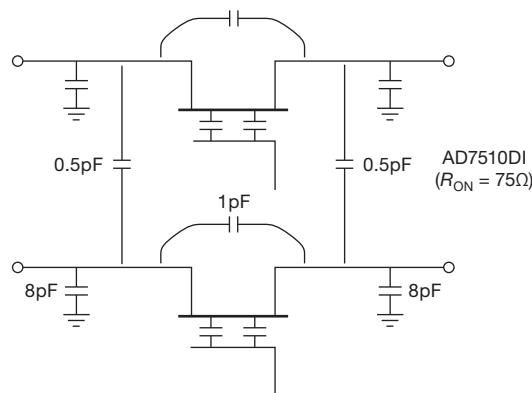
#### D. Capacitance

FET switches exhibit capacitance from input to output ( $C_{DS}$ ), from channel to ground ( $C_D$ ,  $C_S$ ), from gate to channel, and from one FET to another within one IC package ( $C_{DD}$ ,  $C_{SS}$ ); see Figure 3.73. Let's look at the effects.

$C_{DS}$ : capacitance from input to output

Capacitance from input to output causes signal coupling in an OFF switch, rising at high frequencies. Figure 3.74 shows the effect for the popular DG211 and DG411 series. Note the characterization with a  $50\Omega$  load, common in RF circuits, but much lower than normal for low-frequency signals, for which a typical load impedance is  $10\text{k}\Omega$  or more. Even with a  $50\Omega$  load, the feedthrough becomes

<sup>75</sup> As we remarked, you can get switches with lower  $R_{ON}$ , as low as  $0.25\ \Omega$ , at the expense of some combination of increased capacitance, increased charge injection, and reduced operating voltage range.

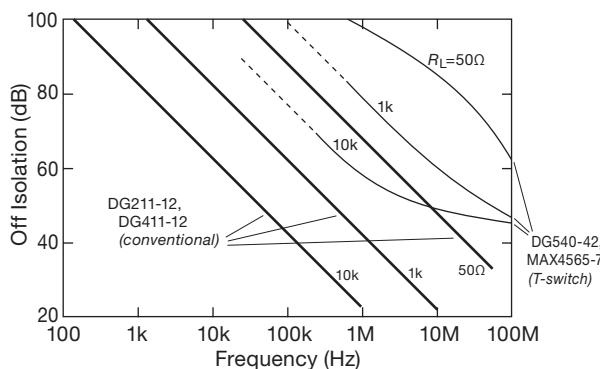


**Figure 3.73.** Capacitances between isolated sections of the AD7510 quad analog switch cause signal crosstalk.

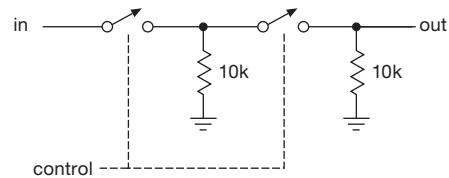
significant at high frequencies (1 pF has a reactance of 5k at 30 MHz, giving  $-40$  dB of feedthrough). And, of course, there is significant attenuation (and nonlinearity) driving a  $50\Omega$  load, because for these parts  $R_{ON}$  is typically  $45\Omega$  and  $17\Omega$ , respectively. With a  $10k$  load the feedthrough situation is much worse, of course.

**Exercise 3.5.** Calculate the feedthrough into  $10k$  at 1 MHz, assuming  $C_{DS} = 1$  pF.

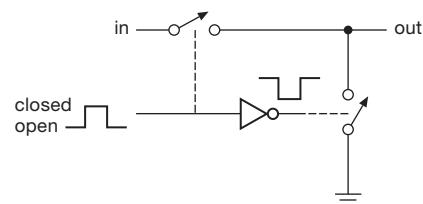
In most low-frequency applications capacitive feedthrough is not a problem. If it is, the best solution is to use a pair of cascaded switches (Figure 3.75) or, better still, a combination of series and shunt switches, enabled alternately (Figure 3.76). The series cascade doubles the attenuation (in decibels) at the expense of additional  $R_{ON}$ , whereas the series-shunt circuit (effectively an SPDT configuration) reduces feedthrough by dropping the effective load resistance to  $R_{ON}$  when the series switch



**Figure 3.74.** High-frequency feedthrough in analog switches. There is less feedthrough with a low load resistance and less still with a “T-switch” configuration.



**Figure 3.75.** Cascaded analog switches for reduced feedthrough.

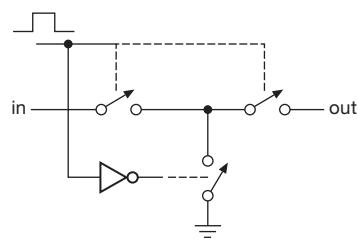


**Figure 3.76.** SPDT analog switch configuration for reduced feedthrough.

is off. Some commercial analog switches are built with a T-network of three switches (Figure 3.77) to achieve low feedthrough for signals going in either direction; from the outside you can't even tell that they've used this trick, except by noticing the excellent isolation specifications as in Figure 3.74 (unless, of course, they brag about it on the datasheet).

**Exercise 3.6.** Recalculate switch feedthrough into  $10k$  at 1 MHz, assuming  $C_{DS}=1$  pF and  $R_{ON}=50\Omega$ , for the configuration of Figure 3.76.

Most CMOS SPDT switches have controlled break-before-make (BBM) characteristics so that the signal sources are not momentarily connected during switching. In some cases, however, you need the reverse, i.e., make-before-break (MBB), for example in a gain-selecting feedback circuit like Figure 3.84B. To deal with this, some CMOS switches are available in both flavors, for example the ADG619 and ADG620 (BBM and MBB, respectively, as noted in Table 3.3).



**Figure 3.77.** A T-switch further reduces high-frequency feedthrough.

**$C_D, C_S$ : capacitance to ground**

Shunt capacitance to ground leads to the high-frequency rolloff mentioned earlier. The situation is worst with a high-impedance signal source, but even with a stiff source the switch's  $R_{ON}$  combines with the shunt capacitance at the output to make a lowpass filter. The following problem shows how it goes.

**Exercise 3.7.** An AD7510 (here chosen for its complete capacitance specifications, shown in Figure 3.73) is driven by a signal source of 10k, with a load impedance of 100k at the switch's output. Where is the high-frequency  $-3\text{ dB}$  point? Now repeat the calculation, assuming a perfectly stiff signal source and a switch  $R_{ON}$  of 75 ohms.

**Capacitance from gate to channel**

Capacitance from the controlling gate to the channel causes a different effect, namely, the coupling of nasty little transients into your signal when the switch is turned on or off. This subject is worth some serious discussion, so we'll defer it to the next section on glitches.

 **$C_{DD}, C_{SS}$ : capacitance between switches**

If you package several switches on a single piece of silicon the size of a grain of rice, it shouldn't surprise you if there is some coupling between channels ("crosstalk"). The culprit, of course, is cross-channel capacitance. The effect increases with frequency and with signal impedance in the channel to which the signal is coupled. Here's a chance to work it out for yourself.

**Exercise 3.8.** Calculate the coupling, in decibels, between a pair of channels with  $C_{DD} = C_{SS} = 0.5\text{ pF}$  (Figure 3.73) for the source and load impedances of the previous exercise. Assume that the interfering signal is 1 MHz. In each case calculate the coupling for (a) OFF switch to OFF switch, (b) OFF switch to ON switch, (c) ON switch to OFF switch, and (d) ON switch to ON switch.

It should be obvious from this example why most broadband RF circuits use low signal impedances, usually  $50\Omega$ . If crosstalk is a serious problem, don't put more than one signal on one chip.

**E. Glitches and charge injection**

During turn-on and turn-off transients, FET analog switches can do nasty things. The control signal being applied to the gate(s) can couple capacitively to the channel(s), putting ugly transients on your signal. The situation is most serious if the signal is at high impedance levels. Multiplexers can show similar behavior during transitions of the input address as well as a momentary connection between inputs if the turn-off delay exceeds the turn-on delay

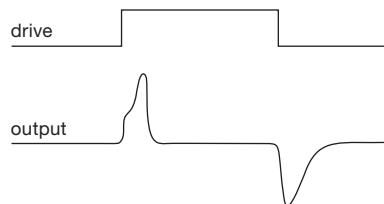
(i.e., MBB). A related bad habit is the propensity of some switches (e.g., the 4066) to short the input to ground momentarily during changes of state.

Let's look at this problem in a bit more detail. Figure 3.78 shows a typical waveform you might see at the output of an *n*-channel MOSFET analog switch circuit, similar to Figure 3.59, with an input signal level of zero volts and an output load consisting of 10k in parallel with 20 pF, realistic values for an analog switch circuit. The handsome transients are caused by charge transferred to the channel, through the gate-channel capacitance, at the transitions of the gate. The gate makes a sudden step from one supply voltage to the other, in this case between  $\pm 15\text{ V}$  supplies, transferring a slug of charge

$$Q = C_{GC}[V_G(\text{finish}) - V_G(\text{start})].$$

Here,  $C_{GC}$  is the gate-channel capacitance, typically around 5 pF. Note that the amount of charge transferred to the channel ("charge injection") depends only on the total voltage change at the gate, not on its rise time. Slowing down the gate signal gives rise to a smaller-amplitude glitch of longer duration, with the same total area under its graph. Lowpass filtering of the switch's output signal has the same effect. Such measures may help if the peak amplitude of the glitch must be kept small, but in general they are ineffective in eliminating gate feedthrough. In some cases the gate-channel capacitance may be predictable enough for you to cancel the spikes by coupling an inverted version of the gate signal through a small adjustable capacitor.

The gate-channel capacitance is distributed over the length of the channel, which means that some of the charge is coupled back to the switch's input. As a result, the size of the output glitch depends on the signal-source impedance and is smallest when the switch is driven by a voltage source. Of course, reducing the size of the load impedance will reduce the size of the glitch, but this also loads the source and introduces error and nonlinearity because of finite  $R_{ON}$ . Finally, all other things being equal, a switch with smaller gate-channel capacitance will introduce smaller

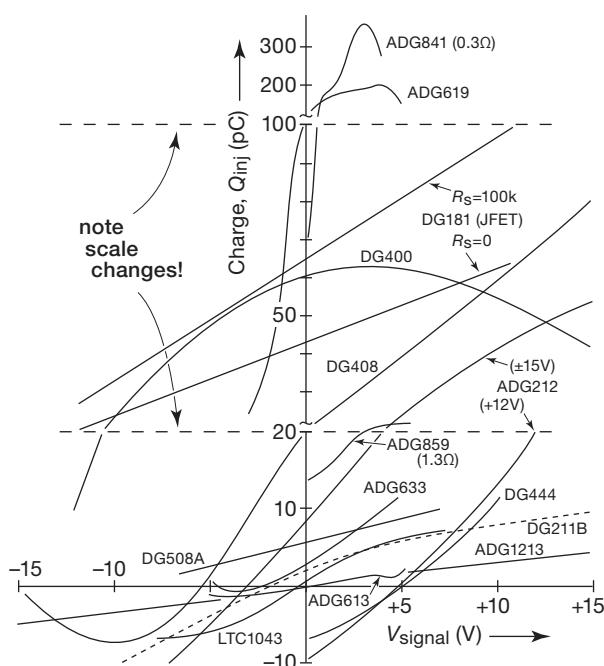


**Figure 3.78.** Charge-transfer glitches, on a greatly expanded scale.

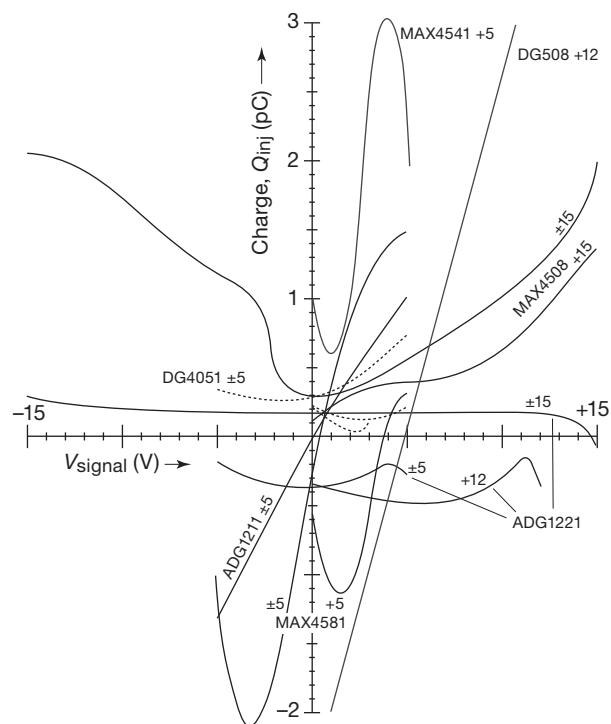
switching transients, although you pay a price in the form of increased  $R_{ON}$ .

Figure 3.79 shows an interesting comparison of gate-induced charge transfers for a collection of analog switches, including JFETs. For the CMOS switches the internal gate signals are making a full rail-to-rail swing (e.g.,  $\Delta V = 30\text{ V}$  for switches running from  $\pm 15\text{ V}$ ); for the JFET switch the gate swings from  $-15\text{ V}$  to the signal voltage. The JFET switch shows a strong dependence of glitch size on signal, because the gate swing is proportional to the level of the signal above  $-15\text{ V}$ . Well-balanced CMOS switches have relatively low feedthrough because the charge contributions of the complementary MOSFETs tend to cancel out (one gate is rising while the other is falling). Just to give scale to these figures, it should be pointed out that  $30\text{ pC}$  corresponds to a  $3\text{ mV}$  step across a  $0.01\text{ }\mu\text{F}$  capacitor. That's a rather large filter capacitor, and you can see that this is a real problem, since a  $3\text{ mV}$  glitch is pretty large when dealing with low-level analog signals.

In Figure 3.80 we've plotted, on an expanded scale, the charge injection scene for a selection of analog switches that exhibit particularly low charge injections. Switches optimized for low charge injection will usually brag about it on the datasheet's headline. For example, the Analog



**Figure 3.79.** Charge transfer for various FET linear switches as a function of signal voltage, taken from the respective datasheets.



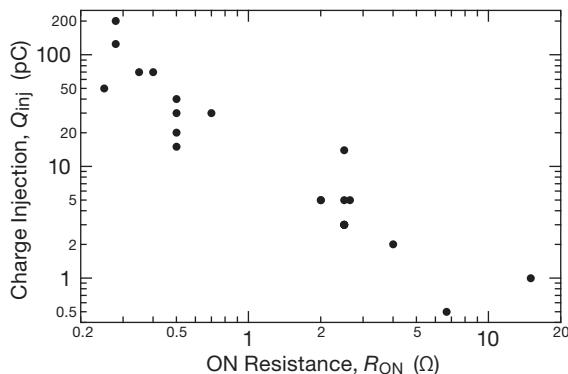
**Figure 3.80.** Need an analog switch with low charge injection? Here are some candidates, plotted on a greatly expanded scale. The three dotted curves are for the DG4051 with  $\pm 5V$ ,  $+5V$ , and  $+3V$  supplies. Check the datasheet for analogous plots for the DG4053 triple SPDT switch.

Devices ADG1221-series datasheet shouts, in bold letters, “**Low Capacitance, Low Charge Injection, ±15 V/+12 V iCMOS® Dual SPST Switches**”; quite a mouthful, but quite a switch!

As might be expected, switches with lower ON-resistance generally exhibit greater charge injection. Figure 3.81 shows this trend, in a scatterplot of datasheet values of  $Q_{inj}$  versus  $R_{ON}$  for the low-voltage CMOS analog switches currently offered by Analog Devices.

## **F. Other switch limitations**

Some additional characteristics of analog switches that may or may not be important in any given application are switching time, settling time, BBM delay, channel leakage current (both ON and OFF), device quiescent current, input current during overvoltage,  $R_{ON}$  matching among multiple channels, and temperature coefficient of  $R_{ON}$ . We'll show unusual restraint by ending the discussion at this point, leaving the reader to look into these details when the circuit application demands it.



**Figure 3.81.** Scatterplot of specified charge injection versus ON-resistance for Analog Devices' low-voltage analog switches, illustrating the  $R_{ON}$  versus  $Q_{inj}$  tradeoff.

### 3.4.3 Some FET analog switch examples

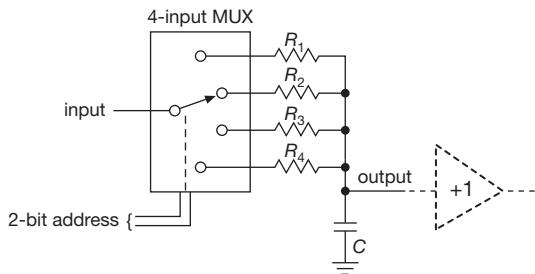
As we indicated earlier, many of the natural applications of FET analog switches are in op-amp circuits, which we treat in the next chapter. In this section we show a few switch applications that do not require op-amps, to give a feeling for the sorts of circuits you can use them in.

#### A. Switchable RC lowpass filter

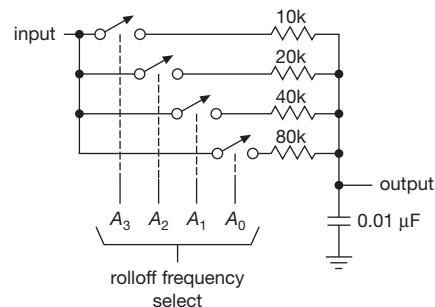
Figure 3.82 shows how you could make a simple *RC* low-pass filter with selectable 3 dB points. We've used a multiplexer to select one of four preset resistors, using a 2-bit (digital) address. We chose to put the switch at the input, rather than after the resistors, because there is less charge injection at a point of lower signal impedance. Another possibility, of course, is to use FET switches to select the *capacitor*. To generate a very wide range of time constants you might have to do that, but the switch's finite  $R_{ON}$  would limit attenuation at high frequencies, to a maximum of  $R_{ON}/R_{\text{series}}$ . We've also indicated a unity-gain buffer, following the filter, since the output impedance is high. You'll see how to make "perfect" followers (precise gain, high  $Z_{in}$ , low  $Z_{out}$ , and no  $V_{BE}$  offsets, etc.) in the next chapter. Of course, if the amplifier that follows the filter has high input impedance, you don't need the buffer.

Figure 3.83 shows a simple variation in which we've used four independent switches, rather than a 4-input multiplexer. With the resistors scaled as shown, you can generate 16 equally spaced 3 dB frequencies by turning on binary combinations of the switches.<sup>76</sup>

<sup>76</sup> An easy way to see that the 3 dB frequencies are integral multiples of the lowest setting is to rewrite  $f_{3dB}$  in terms of the conductance of the parallel resistance  $R_p$  of the selected resistors:



**Figure 3.82.** Analog-MUX selectable *RC* lowpass filter.



**Figure 3.83.** *RC* lowpass filter with choice of 15 equally spaced cutoff frequencies.

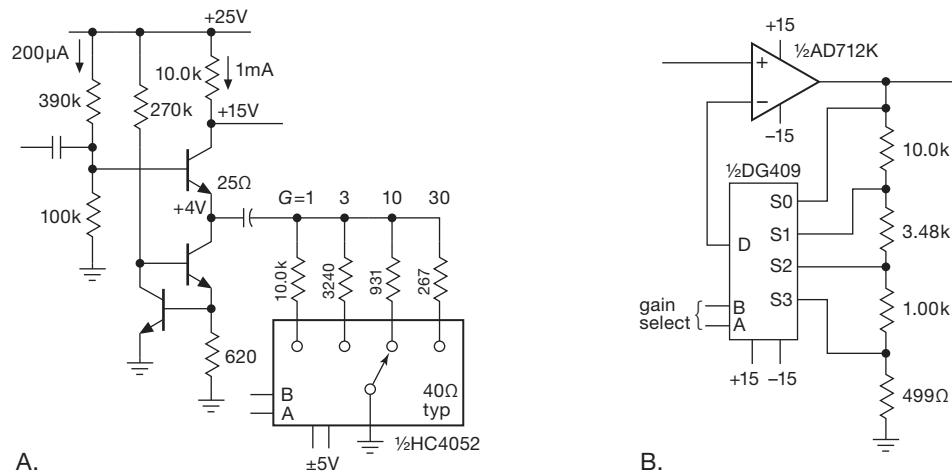
**Exercise 3.9.** What are the 3 dB frequencies for this circuit? Estimate the gain-switching glitch amplitude, assuming a charge injection specification of 20 pC, distributed equally to the input and output switch terminals, and a signal source of low impedance.

#### B. Switchable gain amplifier

Figure 3.84 shows how you can apply the same idea of switching resistors to produce an amplifier of selectable gain. Although this idea is a natural for op-amps, we can use it with the emitter-degenerated amplifier. We used a constant-current sink as emitter load to permit gains much less than unity. We then used the multiplexer to select one of four emitter resistors. Note the blocking capacitor, needed to keep the quiescent current independent of gain.

In Circuit A the switch's  $R_{ON}$  value is part of the gain equation. By contrast, in Circuit B the switch selects a voltage-divider tap and presents it to a high impedance op-amp input, so the switch's  $R_{ON}$  doesn't affect the gain accuracy. Other (more complex) examples of this approach are found in Figures 5.59, 5.62, and 5.80.

$f_{3dB}=1/2\pi R_p C = G_p / 2\pi C$ . Then it's easy, because the conductance of resistors in parallel is the sum of their individual conductances. So for this circuit  $f_{3dB}=nG_{80k}/2\pi C=199n\text{ Hz}$ , where  $G_{80k}=12.5\mu\text{S}$ ,  $C=10\text{nF}$ , and  $n$  is the integer [1..15] represented by the selected switches  $A_n$ .

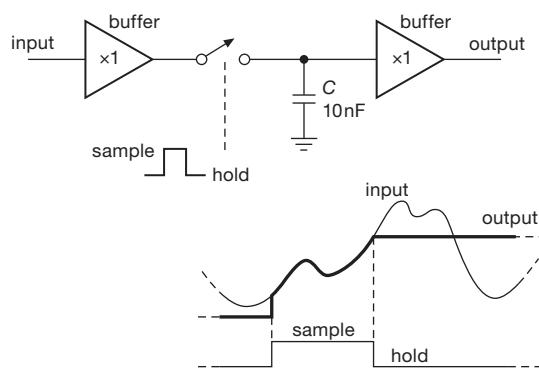


**Figure 3.84.** A. An analog multiplexer selects appropriate emitter degeneration resistors to achieve decade-switchable gain. B. A similar technique, but with the versatile “op-amp” building block (the hero of Chapter 4).

### C. Sample-and-hold

Figure 3.85 shows how to make a “sample-and-hold” (S/H) circuit, which comes in handy when you want to convert an analog signal to a stream of digital quantities (“analog-to-digital conversion”) – you’ve got to hold each analog level steady while you figure out how big it is. The circuit is simple: a unity-gain input buffer generates a low-impedance copy of the input signal, forcing it across a small capacitor. To hold the analog level at any moment, you simply open the switch. The high input impedance of the second buffer (which should have FET input transistors to keep input current near zero) prevents loading of the capacitor, so it holds its voltage until the FET switch is again closed.

**Exercise 3.10.** The input buffer must supply current to keep the capacitor following a varying signal. Calculate the buffer’s peak



**Figure 3.85.** Sample-and-hold.

output current when the circuit is driven by an input sinewave of 1 V amplitude at 10 kHz.

You can do considerably better by closing a feedback loop around the S/H circuit; take a look at §4.5.2. Better still, buy a complete integrated circuit S/H (e.g., the AD783 has an internal hold capacitor, settles to 0.01% in 0.25 $\mu$ s, and droops less than 0.02 $\mu$ V/ $\mu$ s) – let someone else do the hard work!

### D. Flying-capacitor voltage converter

Here’s a nice way (Figure 3.86) to generate a needed negative power-supply voltage in a circuit that is powered by a single positive supply. The pair of FET switches on the left connects  $C_1$  across the positive supply, charging it to  $V_{in}$ , while the switches on the right are kept open.<sup>77</sup> Then the input switches are opened, and the switches on the right are closed, connecting charged  $C_1$  across the output, transferring some of its charge onto  $C_2$ . The switches are diabolically arranged so that  $C_1$  gets turned upside down, generating a *negative* output! This particular circuit, often referred to as a *charge-pump dc-dc converter*, originated as the Intersil 7660 voltage converter chip, and is widely available in improved variants, including voltage-doubling versions and regulated versions. You find them, also, as built-in portions of larger integrated circuits that require dual supply

<sup>77</sup> The device labeled “inverter” turns a logic HIGH voltage into a logic LOW voltage, and vice versa. We’ll show you how to make one in the next section (and we’ll really get you up to speed on them in Chapters 10–14!).

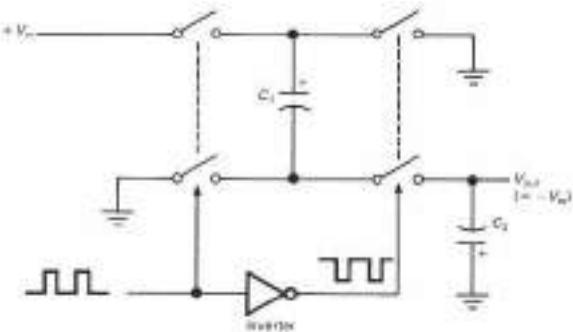


Figure 3.86. Flying-capacitor voltage inverter.

voltages, for example RS-232C serial-port drivers. We'll visit these devices in more detail in §9.6.3.

### E. Digital potentiometer

It's nice to be able to turn a pot electrically – for example, to adjust the volume control of a television set by the remote control "clicker." This kind of application is common, and the semiconductor industry has responded with a variety of electrically settable pots, known variously as *EERPOT*, *E<sup>2</sup>POT*, or just plain *digital pot*. A digital pot consists of a long resistor chain, with an array of FET switches that connects the selected tap to the output pin (Figure 3.87); the tap is selected by a digital input (Chapters 10 and onward).<sup>78</sup> Digital pots come in single, dual, and multiple units; many have "nonvolatile" memory, to retain the position of the pot after power has been turned off. Some have nonlinear taps, for example for audio volume controls, for which it's best to have equal step sizes in decibels (that is, each step produces the same *fractional* increase in voltage-divider ratio). Note that, whatever the configuration, the switch's  $R_{ON}$  appears as a series resistance at the output ("wiper") pin.

As an example, the Analog Devices website lists some 50 digital potentiometers, with from 32 to 1024 steps (256-step models seem to be the most popular), and with one to six channels (singles and duals seem most popular); they use a serial data connection (only two or three pins are needed, regardless of the length of the control data), and they average about \$1 (1000-piece quantity). The selection from Maxim/Dallas includes linear and log taper units (the term *taper* predates digital pots and refers to the resistance versus shaft rotation characteristic of a pot; the log taper is

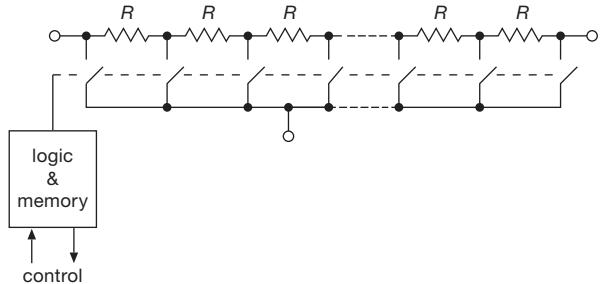


Figure 3.87. An IC "digital potentiometer." Internal digital logic turns on one of the  $n$  analog switches to select a tap along the chain of  $n-1$  fixed resistors.

for audio applications), again in single and multiple units, and with up to 1024 steps per unit. And at the bottom of Intersil's selection table, after all the usual digital pots, you'll even find digital *capacitors*!<sup>79</sup>

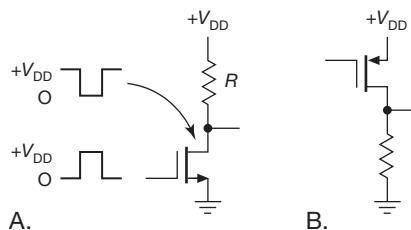
### 3.4.4 MOSFET logic switches

The *other* kinds of FET switch applications are *logic* and *power switching* circuits. The distinction is simple: in analog signal switching you use a FET as a series switch, passing or blocking a signal that has some range of analog voltage. The analog signal is usually a low-level signal at insignificant power levels. In logic switching, on the other hand, MOSFET switches open and close to generate full swings between the power supply voltages. The "signals" here are really digital, rather than analog – they swing between the power supply voltages, representing the two states HIGH and LOW. In-between voltages are not useful or desirable; in fact, they're not even *legal*! Finally, "power switching" refers to turning on or off the power to a load such as a lamp, relay coil, or motor winding; in these applications, both voltages and currents tend to be large. We'll take logic switching first.

Figure 3.88 shows the simplest kind of logic switching with MOSFETs: both circuits use a resistor as load and perform the logical function of *inversion* – a HIGH input generates a LOW output, and vice versa. The *n*-channel version pulls the output to ground when the gate goes HIGH, whereas the *p*-channel version pulls the resistor HIGH for grounded (LOW) input. Note that the MOSFETs in these circuits are used as common-source inverters rather than as source followers. In digital logic circuits like these, we are usually interested in the output voltage ("logic level") produced by a certain input voltage; the resistor serves merely

<sup>78</sup> There are two varieties here: one uses a serial digital data protocol, so that the desired tap position can be sent, as a number, from a controlling microprocessor; the other kind has UP and DOWN pins, with internal memory to hold the current tap position.

<sup>79</sup> The tapped-resistor trick is used in digital-to-analog converters (see §13.2.1), and many ADCs use digital capacitors (see §13.7).



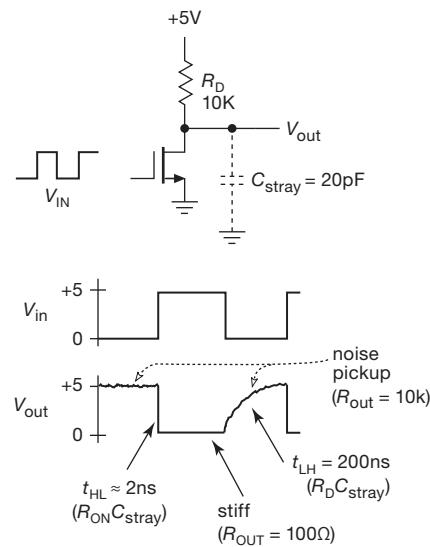
**Figure 3.88.** nMOS and pMOS logic inverters, with resistive “pullups”.

as a passive drain load, to make the output swing to the drain supply when the FET is off. If, on the other hand, we replace the resistor with a lightbulb, relay, print-head hammer, or some other hefty load, we’ve got a power-switching application (Figure 3.3). Although we’re using the same “inverter” circuit, in the power-switching application we’re interested instead in turning the load on and off.

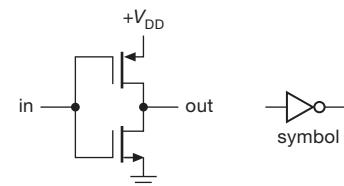
### A. CMOS inverter

The nMOS and pMOS inverters of the preceding circuits have the disadvantage of drawing current in the ON state and having relatively high output impedance in the OFF state. You can reduce the output impedance (by reducing  $R$ ), but only at the expense of increased dissipation, and vice versa. Except for current sources, of course, it’s never a good idea to have high output impedance. Even if the intended load is high impedance (another MOSFET gate, for example), you are inviting capacitive noise pickup problems, and you will suffer reduced switching speeds for the ON-to-OFF (“trailing”) edge (because of stray loading capacitance). In this case, for example, the nMOS inverter with a compromise value of drain resistor, say 10k, would produce the waveform shown in Figure 3.89.

The situation is reminiscent of the single-ended emitter follower in §2.4.1, in which quiescent power dissipation and power delivered to the load were involved in a similar compromise. The solution there – the push-pull configuration – is particularly well suited to MOSFET switching. Look at Figure 3.90, which you might think of as a push-pull switch: input at ground cuts off the bottom transistor and turns on the top transistor, pulling the output HIGH. A HIGH input ( $+V_{DD}$ ) does the reverse, pulling the output to ground. It’s an inverter with low output impedance in *both* states, and no quiescent current whatsoever. It’s called a CMOS (complementary MOS) inverter, and it is the basic structure of all digital CMOS logic, the logic family that has become universal in large and very-large scale integrated circuits (LSI, VLSI), and has largely replaced earlier logic families (with names like transistor-transistor logic,



**Figure 3.89.** High off-impedance in the nMOS inverter causes long rise times and susceptibility to capacitively-coupled noise.



**Figure 3.90.** CMOS logic inverter and circuit symbol.

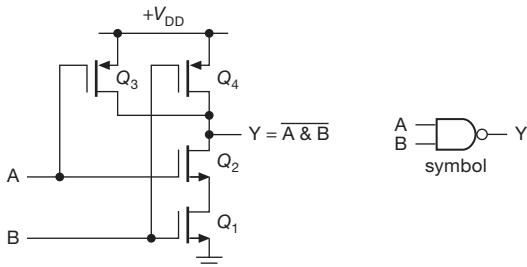
TTL) based on bipolar transistors. Note that the CMOS inverter is two complementary MOSFET switches *in series*, alternately enabled, whereas the CMOS analog switch (treated earlier in the chapter) is two complementary MOSFET switches *in parallel*, enabled simultaneously.

**Exercise 3.11.** The complementary MOS transistors in the CMOS inverter are both operating as common-source inverters, whereas the complementary bipolar transistors in the push-pull circuits of §2.4.1 (e.g., Figure 2.69) are (noninverting) emitter followers. Try drawing a “complementary BJT inverter” analogous to the CMOS inverter. Why won’t it work?

### B. CMOS gates

We’ll be seeing much more of digital CMOS in the chapters on digital logic and microprocessors (Chapters 10–14). For now, it should be evident that CMOS is a low-power logic family (with zero quiescent power) with high-impedance inputs, and with stiff outputs that swing the full supply range. Before leaving the subject, however, we can’t resist the temptation to show you one additional CMOS

circuit (Figure 3.91). This is a logic NAND gate, whose output goes LOW only if input A *and* input B are both HIGH. The operation is surprisingly easy to understand: if A and B are both HIGH, series nMOS switches  $Q_1$  and  $Q_2$  are both ON, pulling the output stiffly to ground; pMOS switches  $Q_3$  and  $Q_4$  cooperate by being OFF; thus, no current flows. However, if either A or B (or both) is LOW, the corresponding pMOS transistor is ON, pulling the output HIGH; since one (or both) of the series chain  $Q_1Q_2$  is OFF, no current flows.



**Figure 3.91.** CMOS NAND gate and circuit symbol.

This is called a NAND gate because it performs the logical AND function, but with inverted (“NOT”) output – it’s a NOT-AND, abbreviated NAND. Although gates and their variants are properly a subject for Chapter 10, you will enjoy trying your hand at the following problems.

**Exercise 3.12.** Draw a CMOS AND gate. *Hint:* AND = NOT NAND.

**Exercise 3.13.** Draw a NOR gate: the output is LOW if either A *or* B (or both) is HIGH.

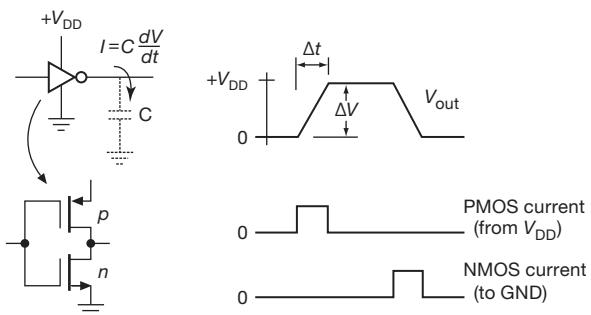
**Exercise 3.14.** You guessed it – draw a CMOS OR gate.

**Exercise 3.15.** Draw a 3-input CMOS NAND gate.

The CMOS digital logic we’ll be seeing later is constructed from combinations of these basic gates. The combination of very low-power dissipation and stiff rail-to-rail output swing makes CMOS logic the family of choice for most digital circuits, accounting for its popularity. Furthermore, for micropower circuits (such as wristwatches and small battery-powered instruments), it’s the only game in town.

Lest we leave the wrong impression, however, it’s worth noting that CMOS logic is not *zero* power. There are two mechanisms of current drain:

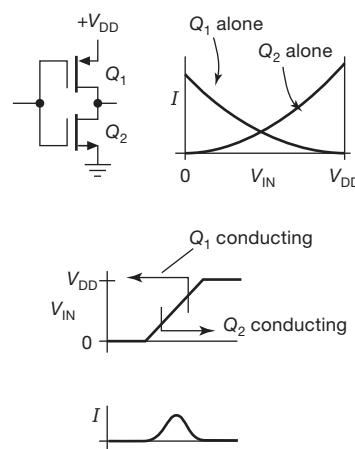
(a) During transitions, a CMOS output must supply a transient current  $I = CdV/dt$  to charge any capacitance it sees (Figure 3.92). You get load capacitance both from wiring (“stray” capacitance) and from the input capacitance of ad-



**Figure 3.92.** Capacitive charging current. The average supply current is proportional to switching rate, and equals  $CVf$ .

ditional logic that you are driving. In fact, because a complicated CMOS chip contains many internal gates, each driving some on-chip internal capacitance, there is some current drain in any CMOS circuit that is making transitions, even if the chip is not driving any external load. Not surprisingly, this “dynamic” current drain is proportional to the rate at which transitions take place.

(b) The second mechanism of CMOS current drain is shown in Figure 3.93: as the input jumps between the supply voltage and ground, there is a region where both MOSFETs are conducting, resulting in large current spikes from  $V_{DD}$  to ground. This is sometimes called “class-A current,” “shoot-through,” or “power-supply crowbarring.” You will see some consequences of this in Chapters 10–12. As long as we’re dumping on CMOS, we should mention that an additional disadvantage of CMOS (and, in fact, of



**Figure 3.93.** When the input gate voltage of a CMOS inverter is intermediate between  $V_{DD}$  and ground, both MOSFETs are partially conducting, causing “class-A conduction,” also known as “shoot-through” current.

all MOSFETs) is their vulnerability to damage from static electricity. We'll have more to say about this in §3.5.4H.

### 3.5 Power MOSFETs

MOSFETs work well as saturated switches, as we suggested with our simple circuit in §3.1.1B. Power MOSFETs are now available from many manufacturers, making the advantages of MOSFETs (high input impedance, easy paralleling, absence of "second breakdown") applicable to power circuits. Generally speaking, power MOSFETs are easier to use than conventional bipolar power transistors. However, there are some subtle effects to consider, and cavalier substitution of MOSFETs in switching applications can lead to prompt disaster. We've visited the scenes of such disasters and hope to avert their repetition. Read on for our handy guided tour.

FETs were feeble low-current devices, barely able to run more than a few tens of millamps, until the late 1970s, when the Japanese introduced "vertical-groove" MOS transistors. Power MOSFETs are now made by all the manufacturers of discrete semiconductors (e.g., Diodes-Inc, Fairchild, Intersil, IR, ON Semiconductor, Siliconix, Supertex, TI, Vishay, and Zetex, along with European companies like Amperex, Ferranti, Infineon, NXP, and ST, and many of the Japanese companies such as Renesas and Toshiba); they are called, variously, VMOS, TMOS, vertical DMOS, and HEXFET. Even in conventional transistor power packages such as TO-220, TO-247, and D-PAK they can handle surprisingly high voltages (up to 1500 V or more), and peak currents over 1000 amps (continuous currents of 200 A), with  $R_{ON}$  below 0.001 Ω. Small-power MOSFETs sell for much less than a dollar, and they're available in all the usual transistor packages. You can also get arrays (multiple MOSFETs) in standard multipin IC packages such as the traditional dual in-line package (DIP) and the smaller surface-mount varieties such as SOT-23, SOIC, and TSOP. Ironically, it is now discrete *low-level* MOSFETs that are hard to find, there being no shortage of power MOSFETs; Table 3.4a on page 188 lists a selection of small *n*-channel MOSFETs up to 250 V, Table 3.4b on pages 189–191 lists other *n*-channel MOSFET sizes and voltages, and Table 3.6 on page 210 has a nice selection of depletion-mode power MOSFETs. There are additional types listed in the MOSFET tables in Chapter 3x.

#### 3.5.1 High impedance, thermal stability

Two important advantages of the power MOSFET compared with the bipolar power transistor, are (a) its extremely high input impedance (essentially infinite at dc),

and (b) its inherent thermal stability. As simple as these might seem, there's more to say, and some important cautions.

##### A. Input impedance

First, the "infinite" input impedance holds only at dc because of substantial input capacitance, which can run to 1000–10,000 pF in typical power MOSFETs. In addition, for *switching* applications you have also to worry about feedback capacitance, i.e., drain-to-gate capacitance (also called *reverse transfer capacitance*,  $C_{rss}$ ), because the Miller effect (§2.4.5) boosts the effective value by the voltage gain. In §3.5.4 we'll discuss this further and display some waveforms showing how the Miller effect fights your efforts to bring about rapid switching. Jumping to the bottom line, you may have to supply several amperes of gate drive current to switch power loads in the tens of nanoseconds that MOSFETs can inherently achieve – hardly the characteristics of an infinite input-impedance device!

##### B. Thermal stability

Second, there are two mechanisms affecting thermal stability in MOSFETs, namely an increase in  $R_{ON}$  with increasing temperature, and, *at the higher end of the transistor's drain current only*, a decrease in drain current (at constant  $V_{GS}$ ) with increasing temperature; see Figure 3.14 and Figures 3.115, and 3.116 in §3.6.3. This latter effect is very important in power circuits and is worth understanding: the large junction area of a power transistor (whether BJT or FET) can be thought of as a large number of small junctions in parallel (Figure 3.94), all with the same applied voltages. In the case of a bipolar power transistor, the positive temperature coefficient of collector current at fixed  $V_{BE}$  (approximately +9%/°C, see §2.3) means that a local hot spot in the junction will have a higher current density, thus producing additional heating. At sufficiently high  $V_{CE}$  and  $I_C$ , this "current hogging" can cause local thermal runaway. As a result, bipolar power transistors are limited to a "safe operating area" (on a plot of collector current versus collector voltage) smaller than that allowed by transistor power dissipation alone. The important point here is that the *negative* temperature coefficient of MOS drain current, when operating at relatively high currents, prevents these junction hot spots entirely. MOSFETs also have no second breakdown, and their safe operating area (SOA) is limited only by power dissipation (see Figure 3.95, where we've compared the SOAs of an *npn* and an nMOS power transistor of the same  $I_{max}$ ,  $V_{max}$ , and  $P_{diss}$ ). This is one reason MOSFETs are favored in linear power applications such as audio power amplifiers.

**Table 3.4a MOSFETs — Small n-channel (to 250V), and p-channel (to 100V)**

small n-channel to 250V											p-channel to 100V										
nMOS type	Pkg <sup>p</sup>	S <sup>x</sup>	V <sub>DSS</sub>	P <sub>D</sub> <sup>C</sup>	I <sub>D</sub> <sup>y</sup>	R <sub>DS(on)</sub> @V <sub>GS</sub>	Q <sub>Gs</sub> <sup>s</sup>	C <sub>iss</sub>	Cost <sup>q</sup>	pMOS type	Pkg <sup>p</sup>	S <sup>x</sup>	V <sub>DSS</sub>	P <sub>D</sub> <sup>C</sup>	I <sub>D</sub> <sup>y</sup>	R <sub>DS(on)</sub> @V <sub>GS</sub>	Q <sub>Gs</sub> <sup>s</sup>	C <sub>iss</sub>	Cost <sup>q</sup>		
ZVN4424	TO-92	•	240	0.7	0.3	4.3Ω	2.5	8	110	0.85	FQT5P10	SOT-223	•	100	2	0.5	820	10	6.3	190	0.38
BSP89	SOT-223	•	240	1.5	0.4	2.8Ω	10	—	100	0.48	VP0106N3	TO-92	—	60	0.7	0.2	8Ω	5	0.5	45	0.55
ZVN1L20	TO-92	•	200	0.7	0.2	6Ω	3	2	55	0.53	BS250P	TO-92	•	45	0.7	0.2	9Ω	10	—	60	<b>0.61</b>
BS107A	TO-92	—	200	0.4	0.2	5Ω	10	—	60	0.31	ZVP2106A	TO-92	•	60	0.7	0.25	3Ω	10	1.8	100	0.61
FQT4N20L	SOT-223	•	200	2.2	0.7	1.0Ω	4.5	4	240	0.34	BSS84	SOT-23	•	50	0.3	0.13	3Ω	5	1	25	<b>0.26</b>
FQT7N10L	SOT-223	•	100	2	1.2	300	5	4.6	220	0.37	NDS0605	SOT-23	•	60	0.4	0.18	1.3Ω	4.5	0.8	79	<b>0.27</b>
ZXMN10A08E	SOT-23	•	100	1.1	0.6	200	10	7.8	500	0.57	FDV304P	SOT-23	•	25	0.4	0.3	1.2Ω	2.7	0.75	63	<b>0.27</b>
ZXMN10A08G	SOT-223	•	100	2	1.5	200	10	7.7	405	0.48	FDN358P	SOT-23	•	30	0.5	1.5	161	4.5	4	182	0.34
VN2222LL	TO-92	—	60	0.4	0.1	7.5Ω	10	—	<60	0.36	ZXMP4A16G	SOT-223	•	40	2	3	83	4.5	14	1000	0.93
VN10KN3	TO-92	—	60	0.7	0.2	6.6Ω	5	1.1	48	<b>0.18</b>	IRF7205	SO-8	—	30	2.5	3	60	10	27	870	<b>0.37</b>
RHU002N06T	SOT-323	•	60	0.2	0.1	2.8Ω	4	1	15	0.21	NTR4171P	SOT-23	•	30	0.5	1.5	60	4.5	16	720	<b>0.13</b>
2N7000	TO-92	•	60	0.4	0.2	2.5Ω	5	1	20	<b>0.17</b>	DMP4050	SO-8	•	40	1.6	3	55	4.5	6.9	670	0.54
2N7002	SOT-23	•	60	0.2	0.1	2.5Ω	4.5	0.9	20	<b>0.16</b>	Si4435DDY	SO-8	•	30	2.5	6	28	4.5	15	1350	<b>0.53</b>
2N7002W	SOT-323	•	60	0.28	0.1	2.5Ω	4.5	0.7	25	<b>0.15</b>	IRF7424	SO-8	•	30	2.5	7	20	4.5	75	4000	0.73
NDS7002A	SOT-23	•	60	0.36	0.20	1.3Ω	4.5	0.8	80	<b>0.27</b>	Si4463DY	SO-8	•	30	3	7 <sup>z</sup>	13	2.5	28	5800	1.03
Si1330EDL	SOT-323	•	60	0.18	0.15	1.4Ω	4.5	0.4	—	0.38	LP0701N3	TO-92	—	16.5	1	0.4	1.7Ω	3	1.6	120	0.82
BSS138	SOT-23	•	50	0.36	0.25	1.0Ω	4.5	0.95	27	<b>0.15</b>	ZXMP61P02F	SOT-23	•	20	0.6	0.5	550	2.7	1.8	150	0.29
ZVN2106A	TO-92	•	60	0.7	0.3	800	10	1.5	75	0.49	NDS332P	SOT-23	•	20	0.5	0.7	350	2.7	2.4	195	<b>0.30</b>
ZVN4306A	TO-92	•	60	0.85	1.0	320	5	3.5	350	1.11	IRLML6402	SOT-23	•	20	1.3	2.2	80	2.5	5	630	<b>0.33</b>
NDT3055	SOT-223	•	60	3	1.7	84	10	9	250	0.34	SI3443DV	TSOP-6	•	20	2	3.4	70	2.7	5.5	610	0.42
PHT8N06LT	SOT-223	•	55	8.3	5 <sup>z</sup>	65	5	11.2	500	0.75	Si2315BDS	SOT-23	•	20	0.57	2.1	71	1.8	8	715	<b>0.43</b>
IRF7470	SO-8	•	40	1.0	7	10	4.5	29	3400	0.76	FDS6575	SO-8	•	20	1.5	6 <sup>z</sup>	11	2.5	53	4950	1.19
2SK3018	SOT-323	•	30	0.2	0.05	5Ω	4	—	13	0.20	CSD25401	SON	•	20	2.8	8 <sup>z</sup>	14	2.5	5.5	1100	1.96
FDV303N	SOT-23	•	25	0.35	0.4	330	2.7	1.1	50	<b>0.23</b>	IRLML6401	SOT-23	•	12	1.3	2	125	1.8	6	830	0.19
IRLML2030	SOT-23	•	30	1.3	0.9	123	4.5	1	110	0.25	IRF7702	TSSOP-8	•	12	1.5	6	15	2.5	30	3500	0.90
NDS355AN	SOT-23	•	30	0.5	1.0	105	4.5	3.5	195	<b>0.29</b>	IRF7420	SO-8	•	12	2.5	8	15	2.5	24	3500	0.66
FDN337N	SOT-23	•	30	0.5	1.3 <sup>z</sup>	70	2.5	4.2	300	<b>0.16</b>	IRF7410G	SO-8	•	12	2.5	12	8	2.5	55	8700	<b>0.97</b>
FDT439N	SOT-223	•	30	1.3	4	55	4.5	10.7	500	0.60	IRF7210	SO-8	•	12	2.5	10	7	2.5	115	17200	<b>1.03</b>
NTR4170N	SOT-23	•	30	0.8	2	50	4.5	4.8	430	0.12	large										
PMV40UN	SOT-23	•	30	1.9	2 <sup>u</sup>	45	2.5	5.5	445	0.39	IRF5940	TO-220	—	100	150	19	120	10	40	1400	<b>2.20</b>
NDT451AN	SOT-223	•	30	3	5	42	4.5	11	720	0.72	IRF9540N	TO-220	—	*	140	20	110	*	60	1300	<b>1.28</b>
IRLML0030	SOT-23	•	30	1.3	2 <sup>u</sup>	33	4.5	2.6	380	0.18	IRFP9140	TO-247	—	100	180	19	120	10	38	1400	3.07
NTLJS4114N	WDFN	•	30	1.9	2.5 <sup>v</sup>	26	2.5	5	650	0.30	IRFP9140N	TO-247	—	*	140	18	111	*	60	1300	<b>1.42</b>
NTMS4800N	SO-8	•	30	0.75	5	20	4.5	7.7	940	0.22	IRF5210	TO-220	—	100	200	25	50	10	115	2700	1.93
IRF7807Z	SO-8	•	30	2.5	10	14.5	4.5	7.2	770	0.61	IIXTR90P10P	TO-247	—	100	190	55	20	10	120	5800	8.10
FDS6680A	SO-8	•	30	1.0	7	10	4.5	16	1600	0.68	IIXTK170P10	TO-264	—	100	890	130	10	10	240	12600	14.20
FDS8817NZ	SO-8	•	30	2.5	10	7	4.5	17	1800	0.75	IXTN170P10	SOT-227	—	*	170	*	*	*	*	*	19.37
NDS331N	SOT-23	•	20	0.5	0.8	150	2.7	2.2	160	<b>0.30</b>	SUD08P06	DPak	•	60	25	2 <sup>z</sup>	158	4.5	7	450	0.75
FDG327NZ	SC70-6	•	20	0.42	1.2	90	1.8	2.1	410	0.40	NTD2955G	DPak	•	60	55	2 <sup>z</sup>	155	10	14	500	0.45
IRLML2502	SOT-23	•	20	1.25	1.25	50	2.5	5	740	<b>0.30</b>	NTP2955	TO-220	—	*	62	11	*	*	*	*	<b>0.69</b>
Si2312CDS	SOT-23	•	20	0.8	2 <sup>u</sup>	35	1.8	3.8	870	0.28	MJE2955 <sup>b</sup>	TO-220	—	60	75	—	80 <sup>b</sup>	—	—	—	0.64
Si2312BDS	SOT-23	•	20	0.8	2 <sup>v</sup>	30	2.5	3.8	770	<b>0.40</b>	FQB21P06	D <sup>2</sup> Pak	•	60	53	2 <sup>z</sup>	140	10	13	420	0.61
IRF6201	SO-8	•	20	2.5	15	2.1	2.5	130	8600	1.06	FQP27P06	D <sup>2</sup> Pak	•	*	*	5 <sup>z</sup>	*	*	*	*	0.92
											IRF4905	TO-220	—	55	200	50	16	10	120	3400	<b>2.04</b>
											STB80PF55	D <sup>2</sup> Pak	•	55	2.4	7	16	10	190	5500	2.53
											STP80PF55	TO-220	—	*	300	55	*	*	*	*	1.78
											SUM55P06	D <sup>2</sup> Pak	•	60	125	7 <sup>z</sup>	15	10	76	3500	2.75
											IRF9234	TO-220	—	55	68	15	10	10	23	620	<b>0.97</b>
											SUP90P06	TO-220	—	60	250	90	9	4.5	90	9200	3.06
											SUP75P05	TO-220	—	55	250	75	8	10	140	8500	5.42
											IRFP064V	TO-247	—	60	250	80	5.5	10	175	6800	1.89
											IRF9204	TO-220	—	40	143	35	20	4.5	150	7700	1.47
											MTP50P03HDL	TO-220	—	30	125	30	20	5	74	3500	3.37
											FDD6637	DPak	•	35	57	7 <sup>z</sup>	14	4.5	25	2400	0.71
											SUP75P03	TO-220	—	30	187	75	5.5	10	140	9000	1.91
											IPB80P03P4L-04	D <sup>2</sup> Pak	•	30	137	16	4.7	4.5	60	8700	1.02

This table shows selected representative MOSFETs. The left nMOS column lists TO-92 and small surface-mount parts. The right column lists all pMOS parts up to 100 V. The lists are sorted by decreasing switch spec, R<sub>DS(on)</sub>. Ignore parts with inadequate V<sub>DSS</sub>, and evaluate parts with a good switch I<sub>D</sub> spec margins. Study candidate datasheets for viability. Amplifiers and linear regulators rely on the P<sub>D</sub> spec. But R<sub>θJC</sub>=125°C/P<sub>D</sub>, and the junction temp will be T<sub>J</sub>=T<sub>A</sub>+P<sub>D</sub>(R<sub>θJC</sub>+R<sub>θJA</sub>), where the latter term is your heatsink thermal path. Both R<sub>θ</sub> terms vary widely for different packages, and the P<sub>D</sub> spec is useful only in that context. You may find that high-voltage parts have lower R<sub>θJC</sub>.

Table 3.4b *n*-channel Power MOSFETs, 55V to 4500V<sup>a</sup> (page 1 of 3)

Part # <sup>k</sup>	Package <sup>p</sup>	Manufacturer <sup>n</sup> Surface mount	<i>I<sub>D</sub></i> ( <i>V<sub>GS</sub></i> =10V)						<i>R<sub>DS(on)</sub></i> <sup>e</sup> typ <sup>r</sup> (mΩ)	<i>V<sub>DS</sub></i> at (V)	Gate zener Superjunction <sup>k</sup>	Charge <i>Q<sub>G</sub></i> <sup>s</sup> typ (nC) <sup>t</sup>			Capacitance <sup>s,s2</sup> ( <i>V<sub>DS</sub></i> =25V <sup>k</sup> )				
			<i>V<sub>DSS</sub></i> 25°C (V)	<i>P<sub>Diss</sub></i> (W)	pulse	25°C <sup>b</sup> (A)	70°C <sup>c</sup> (A)	<i>R<sub>θJC</sub></i> (°C/W)				<i>Q<sub>GD</sub></i> <sup>s</sup> typ (nO) <sup>t</sup>	<i>C<sub>iss</sub></i> typ (pF)	<i>C<sub>oss</sub></i> typ (pF)	<i>C<sub>rss</sub></i> typ (pF)	Cost <sup>q</sup> US \$	year of intro <sup>y</sup>		
<b>2.5 to 4.5kV</b>																			
IXTT02N450	TO-268	Ix -	4500	113	0.6	0.2	0.14	1.1	480Ω	625Ω	10	- -	10.6	5.5	246	19	5.8	17.05	2013
IXTT1N450HV	TO-268HV	Ix -	4500	520	3	1	0.8	0.24	72Ω	85Ω	10	- -	40	20	1730	78	28	28.67	2013
IXTL2N450	i5-Pak	Ix -	4500	220	8	2	1.3	0.56	16Ω	20Ω	10	- -	180	83	6860	267	105	88.80	2013
IXTH02N250	TO-247	Ix -	2500	83	0.6	0.2	0.14	1.5	385Ω	450Ω	10	- -	7.4	5.3	116	8	3	9.52	2013
<b>1500V</b>																			
2SK1317	TO-3P	R -	1500	100	7	2.5	1.5	1.25	9000	12000	10	- -	- -	-	990	125	60	5.46	2004
STP3N150	TO-220	ST -	1500	140	10	2.5	1.6	0.89	6000	9000	10	- -	29.3	17	939	102	13.2	4.64	2008
STP4N150	TO-220	ST -	1500	160	12	4	2.5	0.78	5000	7000	10	- -	30	9	1300	120	12	4.15	2003
IXTH6N150	TO-247	Ix -	1500	540	24	6	4	0.23	2905	3500	10	- -	67	36	2230	170	64	6.07	2010
<b>1200V</b>																			
IXTY02N120	D-Pak	Ix -	1200	33	0.6	0.2	0.2	3.80	60Ω	75Ω	10	- -	4.7	3.2	104	8.6	1.9	2.92	2009
IXTP1N120P	TO-220	Ix -	1200	63	1.8	1	0.7	2	16Ω	20Ω	10	- -	17.6	10.6	550	25	5.4	2.46	2007
IXTP3N120	TO-220	Ix -	1200	200	12	3	2.5	0.62	3735	4500	10	- -	42	21	1100	110	40	4.65	2003
STP6N120K3	TO-220	ST -	1200	150	20	6	3.8	0.83	1950	2400	10	• -	39	23.5	1050	90	3	4.28	2010
IXFH16N120P	TO-247	Ix -	1200	680	35	16	8	0.19	850	950	10	- -	120	47	6900	390	48	11.02	2007
IXFX26N120P	TO-264	Ix -	1200	960	60	26	19	0.13	380	460	10	- -	225	96	16000	735	58	19.28	2007
IXFN32N120	SOT-227	Ix -	1200	780	128	32	20	0.16	290	350	10	- -	400	188	15900	1000	260	27.15	2001
CMF20120D <sup>x</sup>	TO-247	Cr -	1200	215	90	42	24	0.44	80	100	20	- -	91	43	1915	120	13	32.05	2011
<b>800-1000V</b>																			
IXTY01N100	D-Pak	Ix -	1000	25	0.4	0.1	0.1	5	60Ω	80Ω	10	- -	6.9	3	54	6.9	2	0.93	2004
IXTA05N100	D-Pak	Ix -	1000	40	3	0.75	0.5	3.1	15Ω	17Ω	10	- -	7.8	4.1	260	22	8	1.68	2006
IXTP1N100	TO-220	Ix -	1000	54	6	1.5	0.9	2.3	8300	11000	10	- -	14.5	7.5	400	37	13	0.93	2002
FQD2N100	D-Pak	F -	1000	50	6.4	1.6	1	2.5	7100	9000	10	- -	12	6.5	400	40	5	0.79	2001
STD2NK100Z	D-Pak	ST -	1000	70	7.4	1.85	1.16	1.8	6250	8500	10	• -	16	9	499	53	9	2.19	2006
STP2NK100Z	TO-220	ST -	1000	70	7.4	1.85	1.16	1.8	6250	8500	10	• -	16	9	499	53	9	2.16	2006
IRFBG20	TO-220	IR -	1000	54	5.6	1.4	0.86	2.3	8000	11000	10	- -	27	15	500	52	17	1.02	1993
IRFBG30	TO-220	IR -	1000	125	12	3.1	2	1	4000	5000	10	- -	50	30	980	140	50	1.31	1993
STP5NK100Z	TO-220	ST -	1000	125	14	3.5	2.2.	1	2700	3700	10	• -	42	22	1154	106	21	2.52	2005
FQA8N100	TO-3P	F -	1000	225	32	8	5	0.56	1200	1450	10	- -	53	23	2475	195	16	2.28	2005
IXTX24N100	TO-247	Ix -	1000	568	96	24	16	0.22	333	400	10	- -	267	142	8700	785	315	15.53	2009
FQP9N90C	TO-220	F -	900	205	32	8	2.8	1.85	1120	1400	10	- -	45	18	2100	175	14	1.73	2002
FQD1N80	DPak	F -	800	45	4	1	0.63	2.78	15000	20000	10	- -	5.5	3.3	150	20	2.7	0.22	1999
STQ1NK80	TO-92	ST -	800	3	5	0.25	0.16	40	13000	16000	10	• -	7.7	4.5	160	26	7	0.35	2004
STN1NK80	SOT-223	ST -	800	2.5	5	*	*	50	*	*	*	*	*	*	*	*	*	0.75	2004
STD1NK80	DPak	ST -	800	45	5	*	*	2.78	*	*	*	*	*	*	*	*	*	0.64	2004
SPP02N80C3	TO-220	Inf -	800	42	6	2	1.2	3	2400	2700	10	- -	12	6	290	130	6	1.04	2005
FQP7N80C	TO-220	F -	800	167	26	6.6	4.2	0.75	1570	1900	10	- -	27	11		10	1.25	2000	
SPP11N80C3	TO-220	Inf -	800	156	33	11	7.1	0.8	390	450	10	- -	50	25	1600	800	40	2.72	2003

**Notes:** (a) sorted by voltage groups, then by descending *R<sub>DS(on)</sub>*; parts with rated below 55V are not included; many MOSFETs have an alphanumeric part-number scheme, e.g., 10N60, where 10 is the continuous current capability, N means *n*-channel, and the last number is *V<sub>DSS</sub>*/10, so 60 means 600 volts; for these parts a prefix letter usually indicates the package type. (\*) same as row above. (b) the maximum continuous current with *T<sub>C</sub>* = 25°C (manifestly impossible), assuming *T<sub>J</sub>* = 150°C; in some cases the current is limited by the package (bonding wires). (c) *P<sub>Diss</sub>* for *T<sub>case</sub>*=25°C from datasheet. (e) at *T<sub>J</sub>* = 25°C. (k) parts in *italics* have super-junction technology, specs are at *V<sub>DS</sub>* = 50 or 100V. (m) max. (n) Cr = Cree, F = Fairchild, Inf = Infineon, IR = International Rectifier, Ix = IXYS, N = NXP, R = Renesas, ST = STMicroelectronics, Su = Supertex, To = Toshiba, V = Vishay. (p) package types (check datasheet to see which pkgs the mfgr offers for that part); plastic power types with three leads are (largest first): TO-264 with 0.215" lead spacing (to match TO-3 center hole and insulated side mounting notches), TO-247 (same spacing, optional center hole), TO-3P (same spacing and center hole, both to match TO-3), TO-220 (smaller with 0.2" tab with hole, 0.1" lead spacing, very popular); the I2-PAK (TO-262) is a "sawed-off TO-220" stand-up part with 0.1" pitch (three leads plus tab), while the D2-PAK (TO-263) is an SMT version (2 leads plus tab); the I2-PAK (TO-251) is a smaller version of the I2-PAK (i.e., stand-up, 3 leads plus tab, 0.09" pitch), with its corresponding D-PAK (TO-252) SMT version (2 leads plus tab); SOT-223 three leads 0.09" spacing with short tab; power SO-8P is similar to SOIC-8 with a metal "power pad"; LFPAK=SOT669 has SO-8 footprint replacing 4 pins with drain tab; the SOT-227 package measures 1x1.5 in, has four #4 captive-nut spade-lug lead connections, and a super-useful insulated metal heat-sink plate. (q) qty 100. (r) *R<sub>DS</sub>* typ for *T<sub>J</sub>* = 25°C; if hot, scale by 1.5 to 2x for low-voltage parts, or by 2.2 to 3.5x for high-voltage parts, see Chapter 3x. (s) total gate charge to *V<sub>GS</sub>*; gate switching loss = *Q<sub>G</sub>V<sub>GS</sub>f*. (s2) drain capacitive switching *f<sub>loss</sub>* = *C<sub>oss</sub>V<sub>DS</sub><sup>2</sup>f*. (t) typical. (v) at the *V<sub>GS</sub>* at which *R<sub>DS(on)</sub>* is specified. (w) newer -N version costs less, but has higher *R<sub>θJC</sub>*. (x) this is a silicon-carbide (SiC) rather than silicon MOSFET; these have lower capacitance but require higher gate voltages. (y) est'd year-of-introduction; see "A 30-year MOSFET Saga" in chapter 3x. (z) guideline conservative estimate, saturated switch at *V<sub>GS</sub>, T<sub>C</sub>* = 70°C.

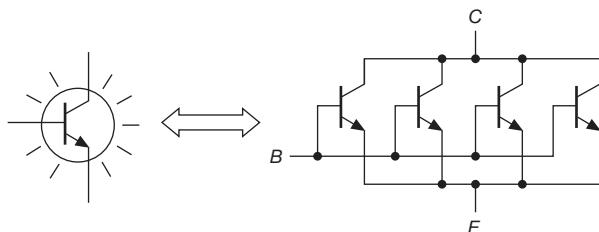
**Usage comments:** • when scanning the list for low *C<sub>oss</sub>*, or maximum power, or low *R<sub>on</sub>*, etc., you may find a better part for your spec, with much higher *V<sub>DSS</sub>* than required -- especially true for high-power capability, a linear servo or pass transistor. • 500V parts have been superceded by 600V parts by many manufacturers. • linear apps often need power dissipation capability, this is often better with older large-die parts; newer designs use narrow V-grooves to achieve low *R<sub>DS(on)</sub>*, and therefore require much less die area. There are good candidates in the table.

Table 3.4b *n*-channel Power MOSFETs, 55V to 4500V<sup>a</sup> (page 2 of 3)

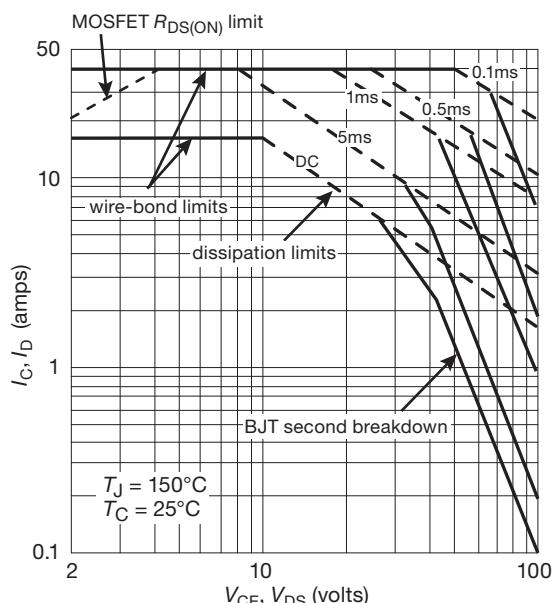
Part # <sup>k</sup>	Package <sup>p</sup>	Manufacturer <sup>r</sup> Surface mount	<i>I</i> <sub>D</sub> ( <i>V</i> <sub>GS</sub> =10V)										Gate zener Superjunction <sup>k</sup>	Charge						Capacitance <sup>s,s2</sup> ( <i>V</i> <sub>DS</sub> =25V <sup>k</sup> )						
			<i>V</i> <sub>DSS</sub> 25°C (V)	<i>P</i> <sub>Diss</sub> <sup>c</sup> (W)	<i>I</i> <sub>D</sub>		<i>R</i> <sub>θJC</sub> (°C/W)	<i>R</i> <sub>D(on)</sub> <sup>e</sup>		<i>V</i> <sub>DS</sub> at (V)				<i>Q</i> <sub>G</sub> <sup>v,s</sup> typ (nC) <sup>t</sup>		<i>Q</i> <sub>GD</sub> <sup>s</sup> typ (nC) <sup>t</sup>		<i>C</i> <sub>iss</sub> typ (pF)		<i>C</i> <sub>oss</sub> typ (pF)		<i>C</i> <sub>rss</sub> typ (pF)		Cost <sup>q</sup> US \$	year of intro <sup>y</sup>	
					pulse (A)	25°C <sup>b</sup> (A)		25°C (A)	70°C (A)					typ <sup>f</sup> (mΩ)	max (mΩ)											
<b>600V</b>																										
VN2460N3, N8	TO-92	Su •	600	1	0.5	0.16	0.12	125	25000	25000	5	-	-	5.5	4	120	10	5	0.84	2000						
FQN1N60C	TO-92	F -	600	3	1.2	0.3	0.18	50	9300	11500	10	-	-	4.8	2.7	130	19	3.5	0.31	2003						
FQD1N60C	DPak	F •	600	28	4	1	1.6	4.5	*	*	*	*	-	*	*	*	*	*	*	0.47	2003					
SPD01N60c3	DPak	IR •	600	11	1.6	0.8	0.5	11	5600	6000	10	-	-	2.2	0.9	100	40	2.5	0.60	2003						
IXTP2N60P	TO-220	Ix -	600	55	4	2	1.4	2.25	4200	5500	10	-	-	7	2.1	240	28	3.5	0.79	2005						
FQP2N60C	TO-220	F -	600	54	8	2	1.35	2.32	3600	4700	10	-	-	8.5	4.1	180	20	4.3	0.60	2002						
FQP3N60C	TO-220	F -	600	75	12	3	1.8	1.67	2800	3400	10	-	-	10.5	4.5	435	45	5	0.70	2005						
FQP5N60C	TO-220	F -	600	100	18	4.5	2.6	1.25	2000	2500	10	-	-	15	6.6	515	55	6.5	0.75	2003						
FDP5N60NZ	TO-220	F -	600	100	18	4.5	2.7	1.25	1650	2000	10	•	-	10	4	450	50	5	0.82	2011						
SPD03N60	DPak	Inf •	600	38	9.6	3.2	2	3.3	1280	1400	10	-	-	13	6	400	150	5	0.84	2002						
FDP7N60NZ	TO-220	F -	600	147	26	6.5	3.9	0.85	1050	1250	10	-	-	13	5.6	550	70	7	1.10	2010						
STP7NM60	TO-220	ST -	600	45	20	5	3	2.8	840	900	10	-	•	14	7.7	363	25	1.1	1.31	2008						
FQB7N60	D2Pak	F •	600	142	30	7.4	4.7	0.88	800	1000	10	-	-	29	14.5	1100	135	16	1.14	2000						
FDP10N60NZ	TO-220	F -	600	185	40	10	6	0.68	640	750	10	•	-	23	8	1110	130	10	1.49	2010						
STP8N65M5	TO-220	ST -	650	70	28	7	4.4	1.8	560	600	10	-	•	15	6	690	18	2	4.48	2009						
FCP7N60N	TO-220	F -	600	64	20	6.8	4.3	1.95	460	520	10	-	•	18	6	719	30	2.1	1.60	2009						
STP11N65M5	TO-220	ST -	650	85	36	9	5.6	1.5	430	480	10	-	•	17	8.5	644	18	2.5	1.67	2012						
FCP9N60N	TO-220	F -	600	83	27	9	5.7	1.5	330	385	10	-	•	22	7.1	930	35	2	2.37	2009						
TK10E60W	TO-220	To -	600	100	39	9.7	-	1.25	327	380	10	-	•	20	9.5	700	20	2.3	2.04	2012						
FQA19N60	TO-3P	F -	600	300	74	18.5	11.7	0.42	300	380	10	-	-	70	33	2800	350	35	2.71	2000						
SPP15N60CFD	TO-220	Inf -	600	156	33	13	8.4	0.80	280	330	10	-	-	63	38	1820	520	21	2.87	2006						
STP15NM60ND	TO-220	ST -	600	125	56	14	9	1	270	299	10	-	•	40	22	1250	65	5	3.42	2007						
IPP60R280C6	TO-220	Inf -	600	104	40	14	8.7	1.2	250	280	10	-	•	43	22	950	60	1.95	2009							
SPP15N60C3	TO-220	Inf -	600	156	45	15	9.4	0.8	250	280	10	-	-	63	29	1660	540	40	2.33	2002						
FCP13N60N	TO-220	F -	600	116	39	13	8.2	1.07	244	258	10	-	•	30	9.5	1325	50	3	2.41	2009						
SiHP15N60E	TO-220	V -	600	180	39	15	9.6	0.7	230	280	10	-	•	38	17	1350	70	5	1.77	2011						
STP16N65M5	TO-220	ST -	650	90	48	12	7.3	1.4	230	279	10	-	•	31	12	1250	30	3	2.96	2011						
IPP60R250CP	TO-220	Inf -	600	104	40	12	8	1.2	220	250	10	-	•	26	9	1200	54	2.56	2010							
STP22NM60N	TO-220	ST -	600	125	64	16	10	1	200	220	10	-	•	44	25	1330	84	4.6	3.82	2009						
IPP60R199CP	TO-220	Inf -	600	139	51	16	10	0.9	180	199	10	-	•	32	11	1520	72	2	2.88	2006						
IRFP27N60K	TO-247	IR -	600	500	110	27	18	0.29	180	220	10	-	-	105	50	4660	460	41	5.53	2002						
TK16E60W	TO-220	To -	600	130	63	16.8	11	0.96	160	190	10	-	•	38	16	1350	35	4	2.86	2012						
STP21N65M5	TO-220	ST -	650	125	68	17	10.7	1	150	179	10	-	•	50	23	1950	46	3	2.33	2010						
SiHP24N65E	TO-220	V -	700	250	70	24	16	0.5	120	145	10	-	-	81	37	2740	122	4	2.45	2009						
IXFH50N60P3	TO-247	Ix -	600	1040	125	50	35	0.12	118	145	10	-	-	94	23	6300	630	2.5	4.95	2011						
IPP60R099CP	TO-220	Inf -	600	255	93	31	19	0.5	90	99	10	-	•	60	20	2800	130	100	4.68	2005						
SPW47N60C3	TO-247	Inf -	600	415	141	47	30	0.3	60	70	10	-	-	252	121	9800	2200	145	5.00	2004						
FCH47N60N	TO-247	F -	600	368	141	47	30	0.34	52	62	10	-	•	115	34	5037	200	2.5	9.97	2010						
IXKN75N60C	SOT-227	Ix -	600	568	-	75	60	0.22	30	36	10	-	-	500	220	15000	6000	300	33.33	2003						
FCH76N60N	TO-247	F -	600	543	228	76	48	0.23	28	36	10	-	•	218	66	9310	370	3.3	18.80	2010						
<b>500V</b>																										
IRF820	TO-220	IR -	500	50	8	2.5	1.6	2.5	2500	3000	10	-	-	16	8.2	360	92	37	0.43	1980						
IRF830	TO-220	IR -	500	74	18	4.5	2.9	1.7	1200	1500	10	-	-	26	12	610	160	68	0.59	1980						
IRF840	TO-220	IR -	500	125	32	8	5.1	1	800	850	10	-	-	52	22	1300	310	120	0.70	1982						
FDP5N50NZ	TO-220	F -	500	78	18	4.5	2.7	4.1	1380	1500	10	•	-	9	4	330	50	4	0.95	2010						
STD5NK50Z	DPak	ST •	500	70	17.6	4.4	2.7	1.78	1220	1500	10	•	-	20	10	535	75	17	1.06	2002						
FDP7N50	TO-220	F -	500	89	28	7	4.2	1.4	760	900	10	-	-	12.8	5.8	720	95	9	1.07	2006						
FQP9N50	TO-220	F -	500	147	36	9	5.7	0.85	580	730	10	-	-	28	12.5	1100	160	20	0.72	2002						
IRFP450	TO-247	IR -	500	190	56	14	9	0.65	350	400	10	-	-	110	55	2600	720	340	2.12	1993						
IRFP460	TO-247	IR -	500	280	80	20	13	0.45	230	270	10	-	-	150	70	4200	870	350	2.58	1993						
STW20NK50Z	TO-247	ST -	500	190	68	20	12.6	0.66	230	270	10	•	-	85	42	2600	328	72	3.15	2002						
FDA28N50	TO-3P	F -	500	310	112	28	17	0.4	122	155	10	-	-	80	32	3866	576	42	3.81	2007						
FCP22N60N	TO-220	F -	500	205	66	22	13.8	0.61	140	165																

Table 3.4b *n*-channel Power MOSFETs, 55V to 4500V<sup>a</sup> (page 3 of 3)

Part # <sup>k</sup>	Package <sup>p</sup>	Manufacturer <sup>n</sup>	Surface mount				<i>I<sub>D</sub></i> ( <i>V<sub>GS</sub></i> =10V)			<i>R<sub>DS(on)</sub></i> <sup>e</sup>			Gate zener at <i>V<sub>DS</sub></i>	Superjunction <sup>k</sup>	Charge		Capacitance <sup>s,2</sup> ( <i>V<sub>DS</sub></i> =25V <sup>k</sup> )			Cost <sup>q</sup> US \$	year of intro <sup>y</sup>
			<i>V<sub>DSS</sub></i> 25°C (V)	<i>P<sub>diss</sub></i> (W)	pulse (A)	25°C <sup>b</sup> (A)	70°C <sup>c</sup> (A)	<i>R<sub>θJC</sub></i> (°C/W)	typ <sup>r</sup> (mΩ)	max (mΩ)	<i>V<sub>DS</sub></i> (V)	typ <sup>s</sup> (nC) <sup>t</sup>	typ <sup>s</sup> (nC) <sup>t</sup>	<i>C<sub>iss</sub></i> typ (pF)	<i>C<sub>oss</sub></i> typ (pF)	<i>C<sub>rss</sub></i> typ (pF)					
<u>200-400V</u>																					
IRF710	TO-220	IR -	400	36	6	2	1.2	3.5	3100	3600	10	-	-	5.7	2.2	170	34	6.3	0.38	1980	
IRF720	TO-220	IR -	400	50	13	3.3	2.1	2.5	1300	1800	10	-	-	15	7	410	120	47	0.28	1980	
IRF730	TO-220	IR -	400	74	22	5.5	3.5	1.7	740	1000	10	-	-	24	13	700	170	64	0.53	1980	
IRF740	TO-220	IR -	400	125	40	10	6.3	1.0	435	550	10	-	-	43	21	1400	330	120	0.67	1981	
STP7NK40Z	TO-220	ST -	400	70	22	5.4	3.4	1.78	850	1000	10	-	-	19	10	535	82	18	1.13	2002	
STP11NK40Z	TO-220	ST -	400	110	36	9	5.67	1.14	490	550	10	-	-	32	18.5	930	140	30	1.45	2003	
FQP17N40	TO-220	F -	400	170	64	16	10.1	0.74	210	10	-	-	-	45	21.7	1800	270	30	1.45	2006	
IRFP244	TO-220	V -	250	150	60	15	9.7	0.83	180	280	10	-	-	63.0	39	1400	320	73	2.99	1997	
FQP16N25	TO-220	F -	250	250	64	16	10	0.88	180	230	10	-	-	27	15	920	190	23	1.36	2000	
FQA30N40	TO-3P	F -	400	290	120	30	19	0.43	107	140	10	-	-	90	46	3400	580	60	3.53	1999	
FDP33N25	TO-220	F -	250	235	132	33	20	0.53	77	94	10	-	-	36.8	17	1640	30	39	1.23	2006	
FDP2710	TO-220	F -	250	403	50	31	0.48	36	42	10	-	-	78	18	5470	426	97	2.92	2007		
FDA69N25	TO-3P	F -	250	480	276	69	44	0.28	34	41	10	-	-	77	37	3570	780	84	2.50	2006	
IXTK120N25P	TO-264	Ix -	250	700	300	120	80	0.18	19	24	10	-	-	185	80	8000	1300	220	8.16	2004	
IRF610	TO-220	IR -	200	36	10	3.3	2.1	3.5	1250	1500	10	-	-	6.3	3.2	140	53	15	0.28	1980	
IRL620	TO-220	IR -	200	50	21	5.2	3.3	2.5	630	800	5	-	-	8.2	5.5	360	91	27	1.03	1993	
IRF620	TO-220	IR -	200	50	18	5.2	3.3	2.5	550	800	10	-	-	10.6	5	260	100	30	0.38	1980	
IRL630	TO-220	IR -	200	74	36	9	5.7	1.7	290	400	5	-	-	24	24	1100	220	70	1.08	1989	
IRF630	TO-220	IR -	200	74	36	9	5.7	1.7	220	400	10	-	-	27	14	800	240	76	0.46	1980	
IRL640	TO-220	IR -	200	125	68	17	11	1.0	125	180	5	-	-	42	24	1800	400	120	0.78	1992	
IRF640	TO-220	IR -	200	125	72	18	11	1.0	130	180	10	-	-	45	24	1300	430	130	0.59	1981	
PSMN102-200Y	LFPak	N •	200	113	65	21.5	13.6	1.1	86	102	10	-	-	31	10	1568	170	55	0.92	2008	
IRFP260N	TO-247	IR -	200	300	200	50	35	0.50	35	40	10	-	-	234	110	4057	603	161	2.15	2009	
IRFP4668	TO-247	IR -	200	520	520	130	92	0.29	8	9.7	10	-	-	161	52	10720	810	160	4.88	2008	
<u>55-100V</u>																					
IRF510	TO-220	IR -	100	43	20	5.6	4	3.5	410	540	10	-	-	5.2	2.2	180	81	15	0.35	1980	
IRF520	TO-220	IR -	100	60	37	9.2	6.5	2.5	200	270	10	-	-	10.3	3.9	360	150	34	0.38	1980	
IRF530	TO-220	IR -	100	88	56	14	10	1.7	100	160	10	-	-	16.2	7	670	250	60	0.44	1980	
IRF540	TO-220	IR -	100	130	110	28	20	1.0	50	77	10	-	-	47	17	1700	560	120	0.60	1980	
FQP33N10	TO-220	F -	100	127	132	33	23	1.18	40	52	10	-	-	38	18	1150	320	62	0.89	1995	
PSMN039-100YS	LFPak	N •	100	74	112	28	20	1	31	40	10	-	-	23	8	1847	86	64	0.57	2010	
FQP44N10	TO-220	F -	100	146	174	43	31	1	30	39	10	-	-	48	24	1400	425	85	1.02	2000	
SUP85N10	TO-220	V -	100	250	240	85	60	0.6	10	12	5	-	-	105	23	6550	665	265	4.88	2000	
HUF7565G3	TO-247	F -	100	515	1200	75	75	0.29	6.7	8	10	-	-	393	74	7585	2345	630	4.87	1998	
IRFB4110	TO-220	IR -	100	370	670	180	120	0.4	3.7	4.5	10	-	-	150	43	9620	670	250	2.84	2005	
IRFZ14	TO-220	IR -	60	43	40	10	7.2	3.5	135	200	10	-	-	9.7	4.7	300	160	29	0.44	1986	
IRFZ24	TO-220	IR -	60	60	68	12	17	2.5	68	100	10	-	-	19	8	640	360	79	0.53	1986	
IRFZ34	TO-220	IR -	60	88	120	30	21	1.7	42	50	10	-	-	30	15	1200	600	100	0.65	1986	
IRFZ44 <sup>w</sup>	TO-220	IR -	60	150	200	50	36	1.0	24	28	10	-	-	42	17	1900	920	170	0.77	1986	
IRLZ44N	TO-220	IR -	55	110	160	47	33	1.4	20	25	5	-	-	28	17	1700	400	150	0.47	1992	
NDP6060L	TO-220	F -	60	100	144	48	24	1.5	20	28	5	-	-	43	21	1630	460	150	1.63	1995	
IRL3705N	TO-220	IR -	55	170	310	89	63	0.9	11	12	5	-	-	95	49	3600	870	320	1.30	2004	
IRFP054N	TO-247	IR -	60	170	290	81	57	0.9	10	12	10	-	-	130	53	2900	880	330	1.36	1996	
IRL2505	TO-220	IR -	55	200	360	104	74	0.75	9	10	5	-	-	130	67	5000	1100	390	1.65	1996	
STP80NF55-08	TO-220	ST -	55	300	320	80	h	80	0.5	6.5	8	10	-	112	40	3740	830	265	2.57	2007	
IRF3205Z	TO-220	F -	55	170	440	110	78	0.9	4.9	6.5	10	-	-	76	30	3450	550	310	1.26	2001	
IRF1405	TO-220	IR -	55	330	680	169	118	0.45	4.6	5.3	10	-	-	170	62	5480	1210	280	1.76	2001	
FDP025N06	TO-220	F -	60	395	1060	265	120	0.38	1.9	2.5	10	-	-	174	50	11190	1610	750	3.41	2006	
FDP020N06B	TO-220	F -	60	333	1252	313	221	0.45	1.65	2.0	10	-	-	87	34	16100	3640	127	3.79	2011	



**Figure 3.94.** A large-junction-area transistor can be thought of as many paralleled small-area transistors.



**Figure 3.95.** Power MOSFETs do not suffer from second breakdown: comparing safe operating areas (SOAs) of a 160 W BJT (MJH6284) and MOSFET (RFP40N10).

This negative temperature coefficient (“tempco”) of  $I_D$  (at fixed  $V_{GS}$ ) has produced some bad advice in the amplifier community, in particular the statement that it’s always OK to connect a set of power MOSFETs in parallel without the current-equalizing “emitter-ballasting” resistors that are necessary with bipolar transistors.<sup>80</sup> You could do that, if the MOSFETs were to be operated in the high-current regime where you get the stabilizing negative tempco. But in practice you usually can’t operate up there anyway, because of power-dissipation limitations that we’ll see in §9.4.1A. And at the lower currents where the tempco is positive and destabilizing, one of a set of paralleled

<sup>80</sup> Because of their positive tempco of  $I_C$  at constant  $V_{BE}$ , see §§2.3 and 2x.2.

MOSFETs will tend to hog the current and suffer excessive power dissipation, often leading to early failure. The solution is to use a small source-ballasting resistor in each of the paralleled MOSFETs (which should be of the same type, and from the same manufacturer), chosen to drop about a volt at the operating current.

By contrast, you can parallel power MOSFETs in switching applications. That’s because the MOSFET is here operated in the ohmic region of low  $V_{DS}$  (characterized by approximately constant resistance  $R_{ON}$ , as opposed to the higher-voltage “current-saturation” region where the transistor is characterized by approximately constant  $I_D$ ): it is the positive tempco of  $R_{ON}$  that stabilizes the current sharing in paralleled power MOSFETs. No ballasting resistors are needed, or even desirable. We’ll say more about this in §3.6.3.

### 3.5.2 Power MOSFET switching parameters

Most power MOSFETs are enhancement type, available in both *n*- and *p*-channel polarities. Relevant parameters are the breakdown voltage  $V_{DSS}$  (ranging from 12 V to 4.5 kV for *n*-channel, and to 500 V for *p*-channel); the channel on-resistance  $R_{DS(on)}$  (as low as 0.8 mΩ); the current- and power-handling ability (as much as 1000 A and 1000 W); and the gate capacitances  $C_{rss}$  and  $C_{iss}$  (as much as 2000 pF and 20,000 pF, respectively).

We need to rain on this parade! These impressively high current and power ratings are usually specified at 25°C case temperature, allowing the junction temperature to rise to 175°C (while the impressively low  $R_{on}$  is specified at 25°C junction temperature!). Unless you’re resident at the South Pole, these are completely unrealistic conditions during continuous high-power switching.<sup>81</sup> See the discussion in §3.5.4D.

### 3.5.3 Power switching from logic levels

You often want to control a power MOSFET from the output of digital logic. Although there are logic families that generate swings of 10 V or more (the “legacy” 4000-series CMOS), the most common logic families (known generically as CMOS) use supply voltages of +5, +3.3, or +2.5 V, and generate output levels close to that voltage or to ground (HIGH and LOW, respectively).<sup>82</sup> Figure 3.96

<sup>81</sup> In their defense, datasheets do provide derating coefficients, but the derated current and power somehow are neglected when the front-page banner is typeset.

<sup>82</sup> The family known as TTL operates from +5 V, but its HIGH output can be as low as +2.4 V, a characteristic shared by some other +5 V parts.

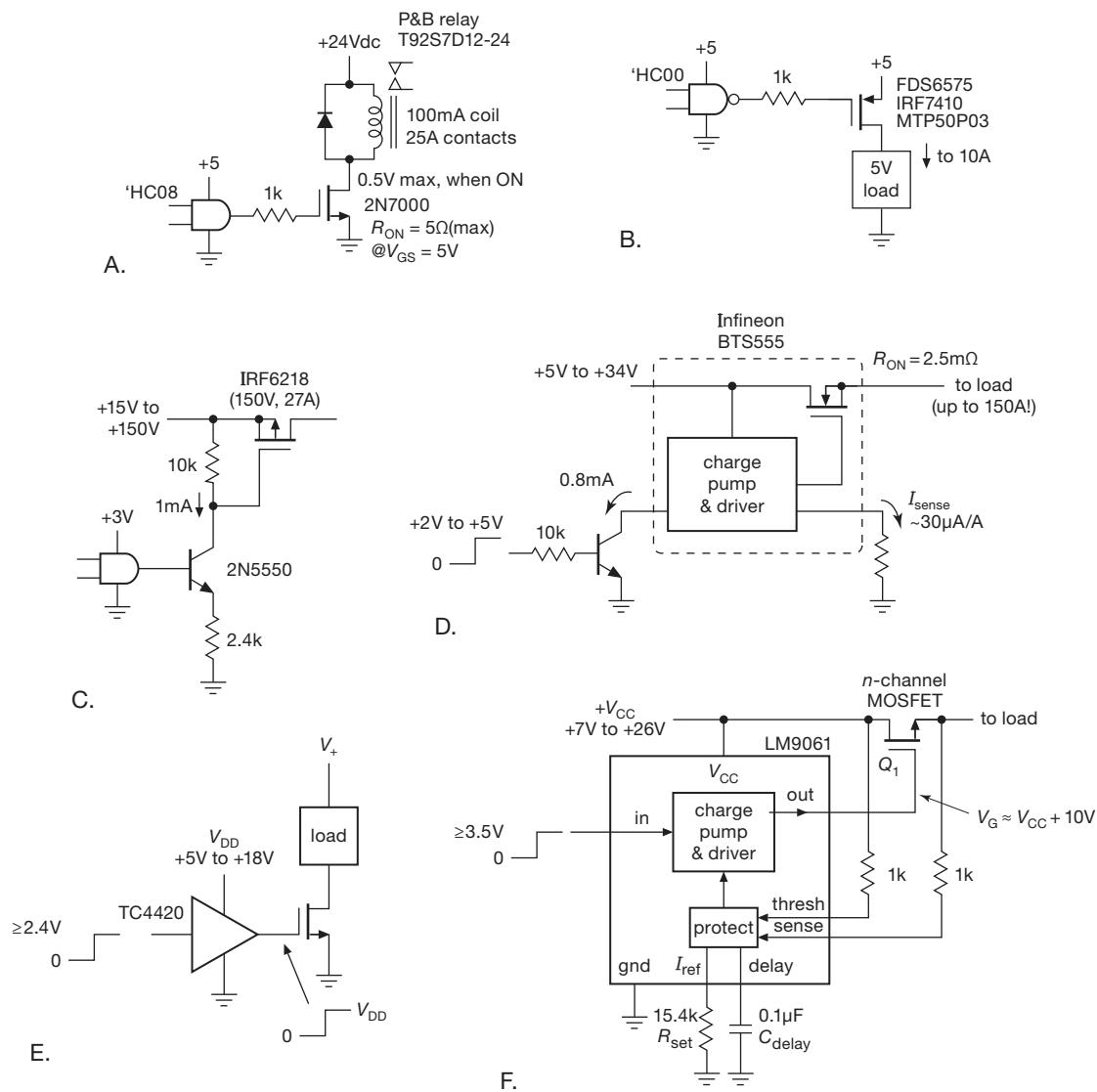


Figure 3.96. MOSFETs can switch power loads when driven from digital logic levels. See also Figure 3.106 on page 204.

shows how to switch loads from these logic families. In the first circuit (Figure 3.96A), the +5 V gate drive will fully turn on a garden-variety MOSFET, so we chose the 2N7000, an inexpensive transistor (\$0.04 in quantity!) that specifies  $R_{ON} < 5\Omega$  at  $V_{GS} = 4.5$  V. The diode protects against inductive spike (§1.6.7). The series gate resistor, though not essential, is a good idea, because MOSFET drain-gate capacitance can couple the load's inductive transients back to the delicate CMOS logic (more on this soon).

For variety, in the second circuit (Figure 3.96B) we've used the *p*-channel MTP50P03HDL, driving a load returned to ground. In a commonly used technique called

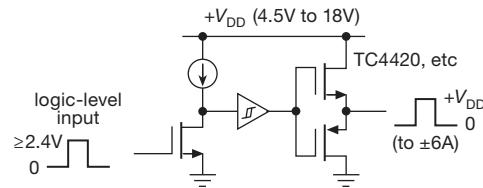
*power switching*, the “load” could be additional circuitry, powered up electrically on command. The '50P03 specifies a maximum  $R_{ON}$  of  $0.025\Omega$  at  $V_{GS} = -5$  V and can handle a 50 amp load current; for lower  $R_{ON}$  you could select the IRF7410 ( $0.007\Omega$ , 16 A, \$1.50); see Table 3.4a on page 188.

Lower voltage logic is increasingly popular in digital circuits. The switching configurations of Figures 3.96A & B can be used for lower voltages, but be sure to use MOSFETs with specified “logic thresholds.” For example the 20 V 16 A *n*-channel FDS6574A from Fairchild specifies a maximum  $R_{ON}$  of  $0.009\Omega$  at a paltry  $V_{GS}$  of

1.8 V, and its *p*-channel FDS6575 sibling specs 0.017  $\Omega$  maximum at a  $V_{GS}$  of  $-2.5$  V; they cost about \$1.25 in small quantities. When choosing low-threshold MOSFETs, watch out for misleading specifications. For example, the IRF7470 MOSFET specifies “ $V_{GS(th)}=2$  V(max),” which sounds good until you read the fine print (“at  $I_D=0.25$  mA”). It takes considerably more gate voltage than  $V_{GS(th)}$  to turn a MOSFET on fully (see Figure 3.115 on page 212). However, the circuit will work OK, because the IRF7470 further specifies “ $R_{ON}(\text{max}) = 30\text{ m}\Omega$  at  $V_{GS} = 2.8$  V.”

The next two circuits show another way to handle the lower drive voltages from low-voltage logic. In Figure 3.96C we’ve used a switched *npn* current sink to generate a “high-side” gate drive for a *p*-channel power-switching MOSFET. Note that if the current sink were replaced with a bipolar *switch*, the circuit would fail immediately for switched voltages greater than the gate-source breakdown voltage. In Figure 3.96D an integrated MOSFET and high-side driver (from Infineon’s “PROFET” line of smart high-side power switches) is used to switch truly prodigious currents – up to 165 A for this particular device.<sup>83</sup> They make it easy to drive from logic levels, by including internal voltage level translation circuitry, and a charge pump (§9.6.3) for the high-side gate drive. You can also get high-side driver ICs for use with an *external n*-channel MOSFET, for example the LM9061 shown in Figure 3.96F. This particular driver also has an internal charge pump to generate the gate voltage for the external nMOS power switch  $Q_1$ ; the gate currents are modest, so the switching speeds are relatively slow. This driver also includes a protection scheme for  $Q_1$  that senses  $V_{DS(\text{ON})}$ , shutting off drive if the MOSFET’s forward drop exceeds a threshold (set by  $R_{\text{set}}$ ), with a delay (to accommodate higher inrush currents) set by  $C_{\text{delay}}$ .

Finally, Figure 3.96E shows how you can sidestep this issue entirely and ensure healthy gate drive voltage *and* current, by using a “MOSFET gate driver” chip like the TC4420. It accepts logic-level input (threshold guaranteed less than +2.4 V), and produces a muscular full-swing output with internal push-pull MOSFETs of its own (Figure 3.97). It can source or sink several amperes of gate current, ensuring fast switching with the large capacitive loads that power MOSFETs present (see §3.5.4B). The trade-off here is cost (about \$1) and complexity. Table 3.8 on page 218 lists a selection of nice gate-driver ICs. We’ll see these again in detail in Chapter 9x (“High-side Cur-



**Figure 3.97.** A MOSFET driver like the TC4420 accepts logic-level input signals and generates a fast (~25 ns), high-current ( $\pm 6$  A), full-swing output. The TC4429 is similar, but with inverted output.

rent Sensing”) and in Chapter 12 (§12.4) in the context of controlling external loads from logic-level signals, and in §12.7 in connection with opto-isolated MOSFET drivers.

### A. Some more MOSFET switching examples

The more, the merrier: let’s look at a few more circuits that take advantage of the MOSFET’s combination of low  $R_{ON}$  and negligible gate current. We’ll keep the descriptions mercifully brief.

#### Pulse stretcher

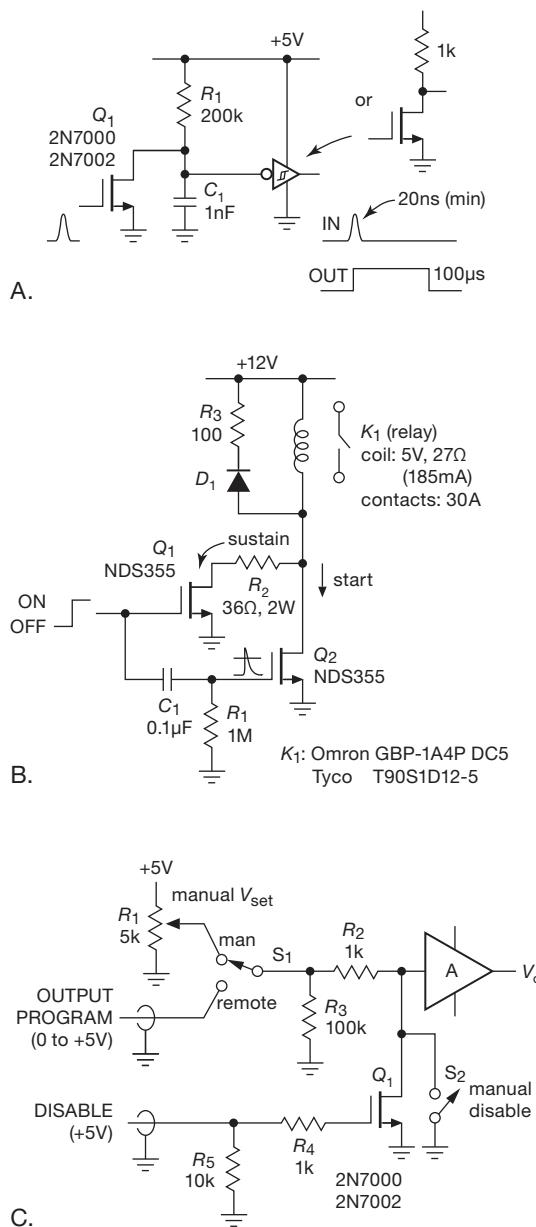
Figure 3.98A is drop-dead simple: MOSFET switch  $Q_1$ , driven with a short positive pulse, discharges capacitor  $C_1$ , making the output go to the positive rail (here +5 V); when the capacitor charges up to the switching threshold of the output inverter (which can be another nMOS transistor or a logic inverter) the output returns to zero. Note that the timing begins at the trailing edge of the input. See §2.2.2 for some fancier discrete circuits that generate an output pulse width insensitive to the input pulse duration. And in Chapter 7 we go into greater detail (§§7.2.1 and 7.2.2), including integrated timing circuits such as the monostable multivibrator (also called a “one-shot”).

#### Relay driver

An electromechanical relay (more detail in Chapter 1x) switches its contacts in response to an energizing current in the coil. The latter has some rated voltage that is guaranteed to switch the contacts and hold them in the energized position. For example, the relays specified in Figure 3.98B have a coil rating of +5Vdc, at which they draw 185 mA (i.e., a coil resistance of 27  $\Omega$ ).<sup>84</sup> In some sense the rated voltage is a compromise: enough to operate the relay reliably, but without excessive current. But you can cheat a bit and get faster closure if you overdrive the coil momentarily,

<sup>84</sup> There’s more: it specifies a “must operate” coil voltage of 3.75 V and a “must release” voltage of 0.5 V. Lots of specs, too, about the switched contacts: configuration, voltage and current ratings, endurance, and so on.

<sup>83</sup> One of our acquaintances uses the powerful BT555 switch in their “Jaws of Life”-type rescue device.



**Figure 3.98.** Useful – and simple – MOSFET applications: A. pulse stretcher; B. relay driver, with initial pulsed overdrive; C. programmable voltage supply with disable control.

as shown in the figure. Here  $Q_2$  applies 12 V for an initial  $\sim 0.1$  s, after which  $Q_1$  sustains the by-now firmly closed contacts with the rated 5 V across the coil. Diode  $D_1$  provides a conduction path for the inductive current at release, with series resistor  $R_3$  allowing  $\sim 20$  V during current decay for faster release (see the discussion in Chapter 1x to understand why this is a good thing).

### Control of a programmable supply

It's nice to control things remotely, with a brainy computer in charge. You might assemble (or buy) a voltage source that accepts a low-level analog input, in the manner of Figure 3.98C, in which the “A” symbol represents a dc amplifier that outputs a voltage  $V_{out}=AV_{in}$ , perhaps capable of substantial output current as well. But it's always good to provide a way to disable external control, so things don't go nuts when the computer crashes or is booting up (or gets taken over by an evildoer). The figure shows a simple way to implement a manual DISABLE control (which duplicates the external DISABLE input); might as well have a manual voltage mode as well, as shown.

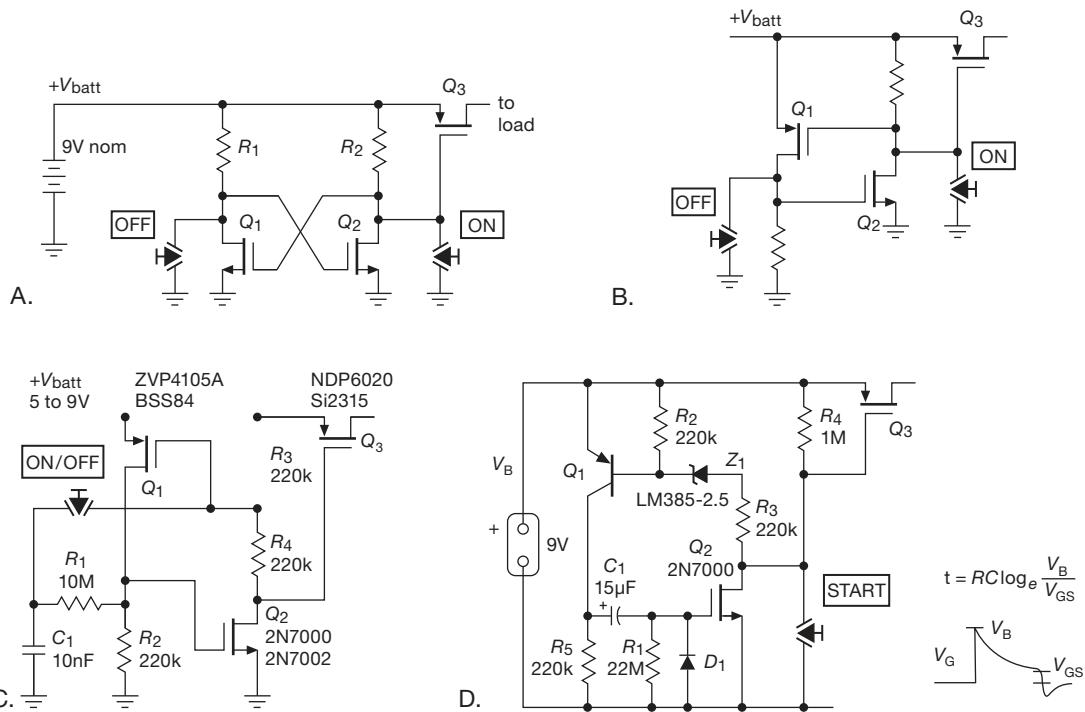
### On/Off battery control

It's convenient to power battery-operated instruments with a 9 volt battery: easy to get; provides plenty of voltage headroom; can be used to create a split supply, see §4.6.1B; but you've got to conserve the battery's vital juices, which run out after about 500 mAh.

Figure 3.99 shows some ways to implement power switching with MOSFETs. Circuit A is the classic flip-flop (two buttons: SET, and RESET; it's called an “SR flip-flop”). The OFF button shuts off  $Q_2$ , whose drain goes HIGH, holding  $Q_1$  on and simultaneously holding pass transistor  $Q_3$  off; you can easily make yourself believe that the ON button does the opposite (and you would be right). Two buttons are OK (though we can do better, stay tuned), but this circuit has the disadvantage of drawing current in either state. You could minimize the standby current by using a 10 MΩ resistor for  $R_1$ , say; then it draws 1 μA when off, which calculates out to a 50-year battery lifetime – far longer than its  $\sim 5$  year shelf life.

But there's a better way. Look at Circuit B, where the complementary pair flip-flop  $Q_1Q_2$  draws no current (other than nanoamp-scale leakage) in the OFF state. The next step is Circuit C (also zero power when OFF), which achieves the minimum button count, in which a single pushbutton acts as an ON/OFF “toggle.” This circuit is a bit tricky, because you have to juggle several time constants appropriately.<sup>85</sup> But the basic concept is simple, and elegant: charge a capacitor from the inverted output of the flip-flop's control input, then momentarily connect the charged capacitor to the control input to make it toggle.

<sup>85</sup> The capacitor's charging time to the new state  $\tau_C=R_1C_1$  should be  $\sim 100$  ms to allow for switch bounce;  $R_3C_1$  should be much shorter, and the discharge time constant of output transistor's gate capacitance  $(R_3+R_4)C_g$  should be faster still. Here we chose 100 ms, 2 ms, and  $\sim 0.4$  ms, respectively.



**Figure 3.99.** ON/OFF battery power control with MOSFETs (no fancy integrated circuits allowed!): A. Classic flip-flop enables *p*-channel series switch; separate ON and OFF buttons. B. Ditto, but zero power when off. C. Single button toggles power ON/OFF; zero power when off. For each MOSFET two choices are listed: the upper is a TO-92 through-hole type, the lower a SOT23 surface-mount type. D. “Five minutes of power” (approximately); zero power when off. The *p*-channel MOSFET pass transistors can be as large as you need, see Table 3.4a.

Finally, for the forgetful we suggest something like Circuit D, a one-button control that turns off automatically (to zero power) after roughly five minutes. Here  $Q_1Q_2$  form a complementary flip-flop, held in the ON state by ac coupling through  $C_1$ ; once ON the latter discharges through  $R_1$ , with a time constant of 330 s. This is an approximation to the actual shutoff time interval, which is set in detail by the ratio of  $Q_2$ ’s gate switching voltage to the actual battery voltage.<sup>86</sup> There’s a bit of complexity here, in the form of the zener  $Z_1$ , which was forced upon us by the need to keep  $Q_3$  fully ON during the critical interval when the flip-flop is deciding to switch OFF. We chose a BJT (rather than a MOSFET) for  $Q_1$  because of its well-defined turn-on voltage; even so, there are the usual headaches here, caused by the uncertainty in gate threshold voltages of  $Q_2$  and  $Q_3$ .

In all four of the circuits in Figure 3.99 you can use as large a *p*-channel pass transistor as needed; see Table 3.4a on page 188 for suggestions. But remember that

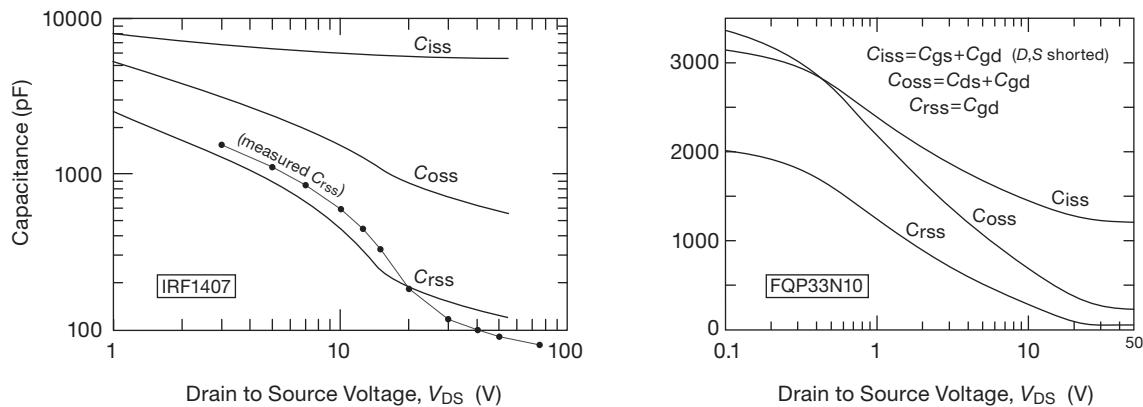
large MOSFETs have high input capacitance  $C_{iss}$ , which slows down switching. As an example, if you were to use an SUP75P05 *p*-channel MOSFET in circuit D, you’d benefit from a very low 8 mΩ on-resistance, but you would have to contend with its substantial 8500 pF value of  $C_{iss}$  (thus nearly 10 ms gate turn-off time constant  $R_4C_{iss}$ ). In these circuits that would not matter. Note that a transistor like the SUP75P05 can switch 50 A, and dissipates about one watt when passing 10 A with a full 10 V gate drive (recall  $R_{ON}$  rises with temperature, at 75°C it’s about 10 mΩ); at most a small 2 W clip-on heatsink is needed (see §9.4.1).<sup>87</sup>

### 3.5.4 Power switching cautions

As nice as MOSFETs are, designing circuits with them is not entirely simple, owing to numerous details that can be devil you. We simply summarize some of the important issues here and delve further into power switching in Chapter 3x and Chapter 9.

<sup>86</sup> A 9 V alkaline battery starts life at about 9.4 V, and reaches old age at 6 V (1V/cell), or very old age at 5.4 V (0.9V/cell).

<sup>87</sup> If you want to work backward from your available heatsinking capability, use  $I = \sqrt{P}/R_{ON}$ .



**Figure 3.100.** Interelectrode capacitances in two power MOSFETs, from graphical data in their respective datasheets. The feedback capacitance  $C_{rss}$ , though smaller than the input capacitance  $C_{iss}$ , is effectively multiplied by the Miller effect and typically dominates in switching applications.

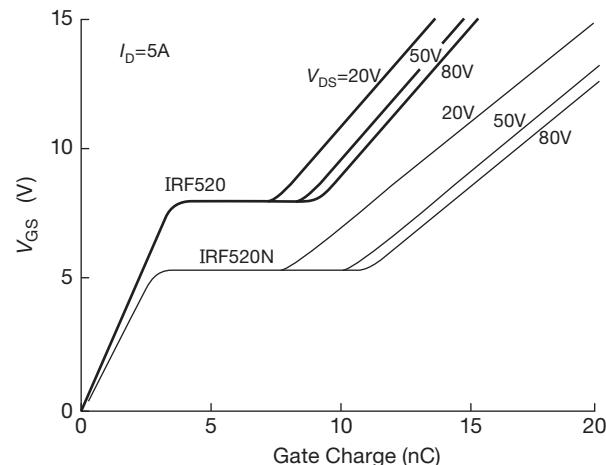
### A. MOSFET gate capacitance

Power MOSFETs have essentially infinite input *resistance*, but they have plenty of input capacitance and also feedback capacitance, such that fast switching may require literally amperes of gate drive current.<sup>88</sup> Although you may not care about speed in many applications, you still have to worry, because low gate drive current causes dramatically higher power dissipation (from the  $VI\Delta t$  product during extended switching transitions); it may also permit oscillations during the slow transition. The various interelectrode capacitances are *nonlinear* and increase with decreasing voltage, as shown in Figure 3.100. The capacitance from gate to ground (called  $C_{iss}$ ) requires an input current of  $i = C_{iss} dV_{GS}/dt$ , and the (smaller) feedback capacitance (called  $C_{rss}$ ) produces an input current  $i = C_{rss} dV_{DG}/dt$ . The latter usually dominates in a common-source switch, because  $\Delta V_{DG}$  is usually much larger than the  $\Delta V_{GS}$  gate drive, effectively multiplying the feedback capacitance by the voltage gain (Miller effect). A nice way to look at this is in terms of gate *charge*, next.

### B. Gate charge

In a common-source switch, charging of the gate-source and gate-drain capacitances require an input gate driving current whenever the gate voltage is changing. Additionally, during drain voltage transitions the Miller effect contributes additional gate current. These effects are often plotted as a graph of “gate charge versus gate-source voltage,” as in Figure 3.101.

The initial slope is the charging of  $C_{iss}$ . The horizontal portion begins at the turn-on voltage, where the rapidly



**Figure 3.101.** Gate charge versus  $V_{GS}$ . The newer small-geometry (“shrunk-die”) IRF520N has a lower threshold voltage, but comparable gate charge. Note in all cases the larger capacitance (reduced slope of  $V_{GS}$  versus  $Q_g$ ) to the right of the “Miller shelf,” the result of larger interelectrode capacitances at low  $V_{DS}$  (Figure 3.100).

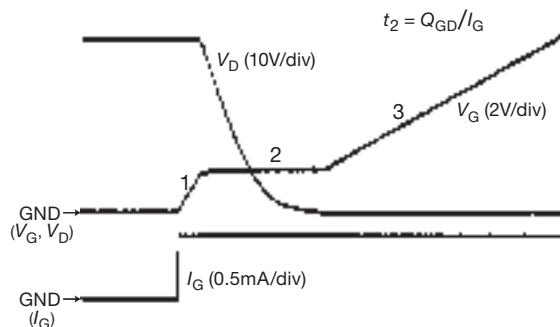
falling drain forces the gate driver to supply additional charge to  $C_{rss}$  (Miller effect). If the feedback capacitance were independent of voltage, the length of the horizontal portion would be proportional to initial drain voltage, after which the curve would continue upward at the original slope. In fact, the “nonlinear” feedback capacitance  $C_{rss}$  rises rapidly at low voltage (Figure 3.100), which means that most of the Miller effect occurs during the low-voltage

<sup>88</sup> Newer-generation MOSFETs generally exhibit somewhat lower capacitances, but their smaller size allows less power dissipation, so for high-

power applications you may be forced to a larger part, thus giving up the capacitance advantage.

portion of the drain waveform.<sup>89</sup> This explains the change in slope of the gate charge curve, as well as the fact that the length of the horizontal portion is almost independent of initial drain voltage.<sup>90</sup>

We hooked up a common-source MOSFET switch and drove the gate with a constant-current drive step, producing the “textbook” traces in Figure 3.102 (hey, this is a textbook, right?). With constant gate drive current, the horizontal (time) axis is proportional to gate charge, in this case 3 nC/div. Here you can see clearly the three regions of gate activity: in region 1 the gate is charging up to the threshold voltage; in region 2 the gate voltage is clamped at the voltage that produces drain currents from 0 to 40 mA (40 V positive rail, 1k load resistor); after the drain is brought to ground the gate resumes its upward voltage ramp, but with reduced slope (owing to increased input capacitance at zero drain voltage).



**Figure 3.102.** Gate charge. Waveforms of an IRLZ34N *n*-channel MOSFET, wired as a common-source switch (1k load to +40 V), with 0.75 mA gate drive. The horizontal scale of 4  $\mu$ s/div therefore corresponds to 3 nC/div of gate charge.

Notice also that the drain voltage trace is curved, caused by the increasing drain-gate capacitance as it heads toward ground: with constant applied input  $I_G$ , the increasing  $C_{rss}$  mandates decreasing  $dV_D/dt$  (to keep the product, i.e., the feedback current, equal to the input current).

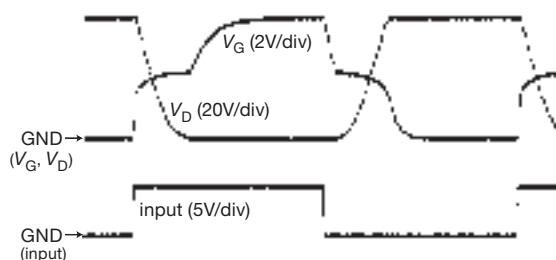
The Miller effect and gate charge in MOSFET switches

<sup>89</sup> This effect can be rather abrupt in power MOSFETs, as seen in the  $C_{rss}$  versus  $V_{DS}$  plot for the IRF1407 (Figure 3.100). In fact, our measured data shows even steeper behavior than that of the datasheet’s plot. As explained in Chapter 3x, this is due to the effective formation of a cascode within the MOSFET, wherein a depletion-mode JFET acts to clamp the drain of the active MOSFET, isolating the latter’s gate and thus greatly reducing the feedback capacitance.

<sup>90</sup> The height ( $V_{GS}$ ) of the horizontal portion depends modestly on drain current; see the figures in §3x.12.

are no laughing matter – they seriously limit switching speed, and you may need to supply hundreds of millamps, or even amps, to get fast transition times in a hefty power switch. For example, the robust IRF1405 featured earlier has  $Q_g \sim 100$  nC; so, to switch it on in 10 ns requires  $I = Q_g/t = 10$  amperes!<sup>91</sup>

On a more modest scale, imagine driving a humble 2N7000 switch from a 0 → 5 V square wave, perhaps from the output of some digital logic. Figure 3.103 shows what happens if you drive the gate through a 10k resistor. Miller is the killer here, causing ~2  $\mu$ s transition times in a transistor which the datasheet says can switch 200 times faster (10 ns). Of course, the datasheet also says  $R_{GEN}=25\Omega$ ... and they mean it!



**Figure 3.103.** A 2N7000 MOSFET switch (1k load to +50 V), whose gate is driven by a 5 V (logic-level) voltage step through a 10k series resistor. The Miller effect stretches the switching time to ~2  $\mu$ s. Horizontal scale: 2  $\mu$ s/div.

There’s plenty more to say about gate charge in MOSFETs: dependence on load current, the shape of the “Miller plateau,” variations among MOSFET types, and measurement techniques. This is discussed in detail in the advanced material in Chapter 3x.

### C. MOSFET drain capacitances

In addition to the gate-to-ground capacitance  $C_{iss}$ , MOSFETs also have a gate-drain feedback capacitance  $C_{dg}$  (usually called  $C_{rss}$ ), and an output capacitance (called  $C_{oss}$ ) that is the combined capacitances from drain to gate  $C_{dg}$  and drain to source  $C_{ds}$ . As just seen, the effect of the feedback capacitance  $C_{rss}$  is evident in the gate-charge waveforms in Figure 3.102. The output capacitance is important, too: it is the capacitance that must be charged and discharged

<sup>91</sup> Quite often it is the output *transition* time that you care about; that is, the time spent just in region 2 (apart from the delay time in region 1, or the gate overcharging time in region 3); for this reason MOSFET datasheets separately specify  $Q_{gd}$ , the gate-to-drain “Miller” charge. For the IRF1405, for example,  $Q_{gd} = 62$  nC, thus requiring a gate input current of 6.2 A to bring about a 10 ns transition time.

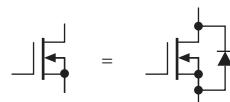
each switching cycle, which, if not reactively recycles, consumes power  $P = C_{oss} f V_{DD}^2$ , which can become significant at high switching frequencies. See §§9.7.2B for details.

### D. Current and power ratings

MOSFET datasheets specify a maximum continuous drain current, but this is done assuming an unrealistic 25°C case temperature. It is calculated from  $I_{D(\max)}^2 R_{DS(ON)} = P_{\max}$ , substituting a maximum power (see §9.4)  $P_{\max} R_{\Theta JC} = \Delta T_{JC} = 150^\circ\text{C}$ , where they have assumed  $T_{J(\max)} = 175^\circ\text{C}$  (thus a 150°C  $\Delta T_{JC}$ ), and they use the value of  $R_{DS(ON)}$  (max) at 175°C from an  $R_{DS}$  tempco plot (e.g., see Figure 3.116). That is,  $I_{D(\max)} = \sqrt{\Delta T_{JC} / R_{\Theta JC} R_{ON}}$ . Some datasheets list power and drain current at a more realistic 75°C or 100°C case temperature. This is better, but you don't really want to run your MOSFET junction at 175°C, so we recommend using a still lower maximum  $I_D$  rating for continuous current, and corresponding dissipated power.

### E. Body diode

With rare exceptions<sup>92</sup> power MOSFETs have the body connected to the source terminal. Because the body forms a diode with the channel, this means that there is an effective diode between drain and source (Figure 3.104) (some manufacturers even draw the diode explicitly in their MOSFET symbol so that you won't forget). This means that you cannot use power MOSFETs bidirectionally, or at least not with more than a diode drop of reverse drain-source voltage. For example, you couldn't use a power MOSFET to zero an integrator whose output swings both sides of ground, and you couldn't use a power MOSFET as an analog switch for bipolarity signals. This problem does not occur with *integrated circuit* MOSFETs (analog switches, for example), where the body is connected to the most negative power-supply terminal.



**Figure 3.104.** Power MOSFETs connect body to source, forming a drain-source diode.

The MOSFET's body diode exhibits the same reverse-recovery effect as ordinary discrete diodes. If biased into forward conduction, it will require some duration of reversed current flow to remove the stored charge, ending with a sharp "snap-off." This can cause curious misbehavior, analogous to the rectifier snap-off transients discussed

<sup>92</sup> Such as the 2N4351 and the SD210 series of lateral MOSFETs. These are available from Linear Systems, Fremont, CA.

in the colorfully-named section "Transformer + rectifier + capacitor = giant spikes!" in Chapter 9x.<sup>93</sup>

### F. Gate-source breakdown

Another trap for the unwary is the fact that gate-source breakdown voltages ( $\pm 20\text{ V}$  is a common figure) are lower than drain-source breakdown voltages (which range from 20 V to more than 1000 V). This doesn't matter if you're driving the gate from the small swings of digital logic, but you get into trouble immediately if you think you can use the drain swings of one MOSFET to drive the gate of another.

### G. Gate protection

As we discuss next, all MOSFET devices are extremely susceptible to gate oxide breakdown, caused by electrostatic discharge. Unlike JFETs or other junction devices, in which the junction avalanche current can safely discharge the overvoltage, MOSFETs are damaged irreversibly by a single instance of gate breakdown. For this reason it is a good idea to use a series gate resistor of 1k or so (assuming that speed is not an issue), particularly when the gate signal comes from another circuit board. This greatly reduces the chances of damage; it also prevents circuit loading if the gate is damaged, because the most common symptom of a damaged MOSFET is substantial dc gate current.<sup>94</sup> You can get additional protection by using a pair of clamp diodes (to  $V_+$  and to ground), or a single clamp zener to ground, downstream of the gate resistor (which can then be of much lower resistance, or omitted altogether); but note that a zener clamp adds some input capacitance.<sup>95</sup> It's also a good idea to avoid floating (unconnected) MOSFET

<sup>93</sup> This can be a serious problem in certain types of switching circuits, where inductive currents continue to flow after termination of switch conduction. It can be addressed by adding an external diode between drain and source, which will be in parallel with the MOSFET's intrinsic drain-source diode. You can use a Schottky diode for voltages below roughly 60 V, but at higher voltages even Schottky diodes have too much voltage drop and thus fail to take over from the FET's intrinsic diode. To deal with this there are power MOSFETs available that include special soft-recovery diodes in their design. These diodes have low reverse-recovery-charge  $Q_{rr}$ , and hence faster recovery time  $t_{rr}$ , so they don't get stuffed with as much charge from this persistent inductive current. They may also have a slower snapoff, further reducing the spike energy.

<sup>94</sup> A MOSFET with damaged gate may exhibit drain conduction when it should be in a non-conducting state: leakage current from drain to (damaged) gate brings the drain down to a voltage that produces the  $V_{GS}$  corresponding to the drain current.

<sup>95</sup> Power MOSFETs used to incorporate internal zener protection, but now it's rare: the zener itself became a dominant failure mechanism! MOSFETs with internal gate zener diodes are marked in the "Gate zener" column of Table 3.4b on page 189.

gates, which are susceptible to damage when floating (there is then no *circuit* path for static discharge, which otherwise provides a measure of safety). This can happen unexpectedly if the gate is driven from another circuit board. A good practice is to connect a pull-down resistor (say 100k to 1M) from gate to source of any MOSFETs whose gates are driven from an off-card signal source. This also ensures that the MOSFET is in the off state when disconnected or unpowered.

## H. MOSFET handling precautions

The MOSFET gate is insulated by a layer of glass ( $\text{SiO}_2$ ) a hundred nanometers thick (less than a wavelength of light). As a result, it has very high resistance, and no resistive or junction-like path that can discharge static electricity as it is building up. In a classic situation you have a MOSFET (or MOSFET integrated circuit) in your hand. You walk over to your circuit, stick the device into its socket, and turn on the power, only to discover that the FET is dead. You killed it! You should have grabbed onto the circuit board with your other hand before inserting the device. This would have discharged your static voltage, which in winter can reach thousands of volts.<sup>96</sup> MOS devices don't take kindly to "carpet shock," which is officially called *electrostatic discharge* (ESD). For purposes of static electricity, you can be approximated by the "human body model" (HBM), which is 100 pF in series with 1.5k;<sup>97</sup> in winter your capacitance may charge to 10 kV or more with a bit of shuffling about on a fluffy rug, and even a simple arm motion with shirt or sweater can generate a few kilovolts. Here are some scary-looking numbers:

Typical Electrostatic Voltages<sup>a</sup>

Action	Electrostatic Voltage	
	10%–20% humidity	65%–90% humidity
(V)	(V)	
Walk on carpet	35,000	1,500
Walk on vinyl floor	12,000	250
Work at bench	6,000	100
Handle vinyl envelope	7,000	600
Pick up poly bag	20,000	1,200
Shift position on foam chair	18,000	1,500

(a) adapted from Motorola Power MOSFET Data Book.

<sup>96</sup> "Smokey, my friend, you are entering a world of pain."

<sup>97</sup> A bit simplistic, though. The HBM, charged to 2.5 kV, peaks at 1.7 A, with a time constant of 150 ns. There are other models, for example the "machine" model (several cycles of 12 kHz, up to 6 A), or the "charge device model" (CDM), which recognizes that a portion of a charged object with less series resistance can discharge directly into the circuit with 6 A pulses 2 ns wide. See also §12.1.5.

Although any semiconductor device can be clobbered by a healthy spark, MOS devices are particularly susceptible because the energy stored in the gate-channel capacitance, when it has been brought up to breakdown voltage, is sufficient to blow a hole through the delicate gate oxide insulation. (If the spark comes from your finger, your additional 100 pF only adds to the injury.) Figure 3.105 (from a series of ESD tests on a power MOSFET<sup>98</sup>) shows the sort of mess this can make. Calling this "gate breakdown" gives the wrong idea; the colorful term "gate *rupture*" might be closer to the mark!

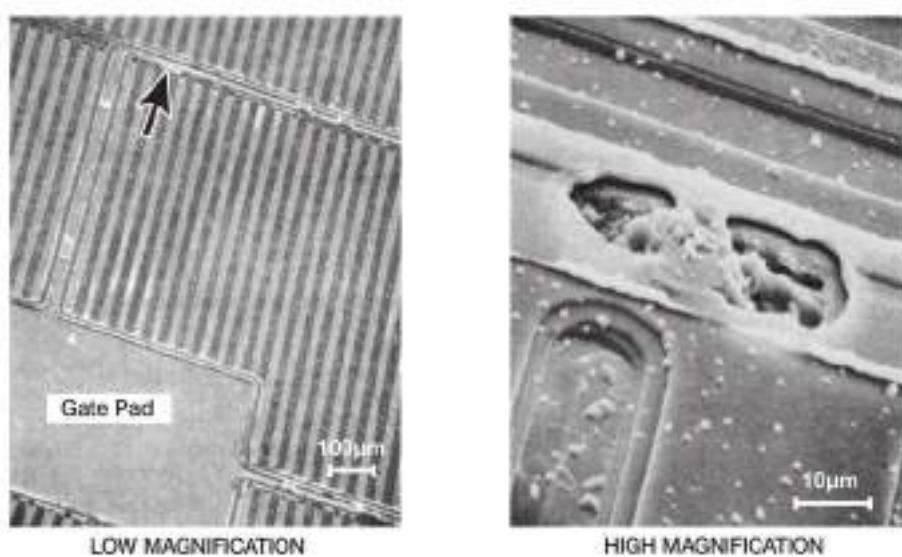
The electronics industry takes ESD very seriously. It is probably the leading cause of nonfunctional semiconductors in instruments fresh off the assembly line. Books are published on the subject, and you can take courses on it.<sup>99</sup> MOS devices, as well as other susceptible semiconductors<sup>100</sup> should be shipped in conductive foam or bags, and you have to be careful about voltages on soldering irons, etc., during fabrication. It is best to ground soldering irons, table tops, etc., and use conductive wrist straps. In addition, you can get "antistatic" carpets, upholstery, and even clothing (e.g., antistatic smocks containing 2% stainless steel fiber). A good antistatic workstation includes humidity control, air ionizers (to make the air slightly conductive, which keeps things from charging up), and educated workers. In spite of all this, failure rates increase dramatically in winter.

Once a semiconductor device is safely soldered into its circuit, the chances for damage are greatly reduced. In addition, most small-geometry MOS devices (e.g., CMOS digital integrated circuits, but not power MOSFETs) have protection diodes in the input gate circuits. Although the internal protection networks of resistors and clamping diodes (or sometimes zeners) compromise performance somewhat, it is often worthwhile to choose those devices because of the greatly reduced risk of damage by static electricity. In the case of unprotected devices, for example power MOSFETs, small-geometry (low current) devices tend to be the most troublesome, because their low input capacitance is easily brought to high voltage when it comes in contact with a charged 100 pF human. Our personal experience with the small-geometry VN13 MOSFET

<sup>98</sup> The MOSFET is an MTM6M60, for which  $C_{iss}=1100 \text{ pF}$ . This forms a capacitive divider with the HBM's 100 pF, attenuating the 1 kV 150 ns spike to about 80 V. But that's still way over the part's 20 V maximum  $V_{GS}$  rating.

<sup>99</sup> Of course, we academics love to give courses on just about anything.

<sup>100</sup> Which includes just about everything: small-geometry RF bipolar transistors are very delicate; and you can zap a plain ol' BJT if you hit it hard enough.



**Figure 3.105.** Scanning electron micrograph of a 6 amp MOSFET destroyed by 1 kV charge on “human body equivalent” (1.5k in series with 100 pF) applied to its gate. (Courtesy of Motorola, Inc.)

was been so dismal in this regard that we stopped using it in production instruments.

It is hard to overstate the problem of gate damage caused by breakdown in MOSFETs. Chip designers realized the seriousness of the problem, and routinely rate the “ESD tolerance” of their devices. Typically MOS ICs survive 2 kV, applied with the HBM of a charged 100 pF capacitor in series with a 1.5k resistor, and the datasheet will say so. Devices that might be exposed to external impulses (e.g., interface and line drivers) are sometimes rated to 15 kV (for example, Maxim RS-422/485 and RS-232 interface chips with an “E” suffix, and many of the analogous parts from other manufacturers).

See also the discussion of input protection in Chapter 12 (§12.1.5).

### I. MOSFETs in parallel

Sometimes you need to use several power transistors in parallel, either to handle greater currents or to be able to dissipate more power, or both. As we discussed earlier, bipolar transistors, owing to their +9%/°C tempco of collector current with temperature, need emitter-ballasting resistors to ensure that the current is distributed equally among the participating transistors. For MOSFETs, as we mentioned in §3.5.1, the situation is different: sometimes you can connect them in parallel without any resistors (e.g., as saturated switches), and sometimes you can’t (as linear power devices<sup>101</sup>). There’s also the related topic of thermal run-

away. These are important topics, and deserve the more extended discussion in §3.6.3.

### 3.5.5 MOSFETs versus BJTs as high-current switches

Power MOSFETs are attractive alternatives to conventional power BJTs most of the time. They’re comparable in price, simpler to drive, and they don’t suffer from second breakdown and consequently reduced safe-operating-area (SOA) constraints (Figure 3.95).

Keep in mind that, for small values of drain voltage, an ON MOSFET behaves like a small resistance ( $R_{ON}$ ), rather than exhibiting the finite saturation voltage ( $V_{CE(sat)}$ ) of its bipolar transistor cousin. This can be an advantage, because the “saturation voltage” goes clear to zero for small drain currents. There is a general perception that MOSFETs don’t saturate as well at high currents, but our research shows this to be largely false. In the following table we’ve chosen comparable pairs (*n*p*n* versus *n*-channel MOSFET), for which we’ve looked up the specified  $V_{CE(sat)}$  or  $R_{DS(ON)}$ . The *low-current* MOSFET is comparable to its “small-signal” *n*p*n* cousin, but in the range of 6–10 A and 0–100 V, the MOSFET does better. Note particularly that very large base currents are needed to bring the bipolar power transistor into good saturation – 10% or more of the collector current (thus as much as 1 A!) – compared with the (zero-current) 10 volt bias at which MOSFETs are usually specified. Note also that

<sup>101</sup> Exception: “lateral” power MOSFETs, such as the 2SK1058.

high-voltage MOSFETs (say,  $BV_{DS} > 200V$ ) tend to have larger  $R_{DS(ON)}$ , with larger temperature coefficients, than the lower-voltage units; here IGBTs excel over MOSFETs above 300 to 400 volts. We've listed capacitances in the table, because power MOSFETs traditionally had more capacitance than BJTs of the same rated current. In some applications (particularly if switching speed is important) you might want to consider the product of capacitance and saturation voltage as a figure of merit.

parameter	<i>n</i> -channel FQP9N25	<i>p</i> -channel FQP9P25
$V_{max}$	250 V	250 V
$I_{max}$	9.4 A	9.4 A
$R_{ON(max)}$	0.42 Ω	0.62 Ω
$C_{rss(typ)}$	15 pF	27 pF
$C_{iss(typ)}$	540 pF	910 pF
$Q_g(typ)$	15.5 nC	29 nC
$T_{JC(max)}$	1.39 °C/W	1.04 °C/W
Price (qty 1k)	\$0.74	\$0.97

### BJT-MOSFET-IGBT Comparison<sup>a</sup>

class	part#	$V_{sat}$		$C_r^b$ (pF)	price
		25°C	125°C		
60 V, 0.5 A	2N4401 <sup>N</sup>	0.75	0.8	8	\$0.06
	2N7000 <sup>V</sup>	0.6	0.95	25	\$0.09
60 V, 6 A	TIP42A <sup>N</sup>	1.5	1.7	50	\$0.63
	IRFZ34E <sup>V</sup>	0.25	0.43	50	\$1.03
100 V, 10 A	TIP142 <sup>D</sup>	3.0	3.8	low	\$1.11
	IRF540N <sup>V</sup>	0.44	1.0	40	\$0.98
400 V, 10 A	2N6547 <sup>N</sup>	1.5	2.5	125	\$2.89
	FQA30N40 <sup>V</sup>	1.4	3.2	60	\$3.85
600 V, 10 A	STGP10NC60 <sup>I</sup>	1.75	1.65	12	\$0.86

(a)  $I_B=I_C/10$ ,  $V_{GS}=10$  V, except  $I_B=I_C/250$  for Darlington.

(b)  $C_{ob}$  or  $C_{rss}$ .

(D) Darlington. (I) IGBT. (N) *npn* BJT. (V) vertical nMOS.

Remember that power MOSFETs can be used as BJT substitutes for linear power circuits, for example audio amplifiers and voltage regulators (we'll treat the latter in Chapter 9). Power MOSFETs are also available as *p*-channel devices, although there tends to be a much greater variety available among the (better-performing) *n*-channel devices. The available *p*-channel MOSFETs go only to 500 V (or occasionally 600 V), and generally cost more for comparable performance in some parameters ( $V_{DS(max)}$  and  $I_{D(max)}$ , say), with reduced performance in other parameters (capacitance,  $R_{ON}$ ). Here, for example, are specifications for a pair of complementary MOSFETs from Fairchild, matched in voltage and current ratings, and packaged in the same TO-220 power package.

Note that the *p*-channel device, having been fabricated with a larger area to achieve comparable  $I_{D(max)}$ , winds up with inferior (i.e., larger) capacitance, gate charge,  $R_{ON}$ , and pricing. It is also slower and has lower transconductance, according to the datasheet. Paradoxically, the *p*-channel device has improved thermal conductivity (see §9.4.1A), presumably resulting from the larger required chip size.

### 3.5.6 Some power MOSFET circuit examples

Enough theory! Let's look at a few circuit examples with power MOSFETs.

#### A. Some basic power switches

Figure 3.106 shows six ways to use a MOSFET to control the dc power to some subcircuit that you want to turn on and off. If you have a battery-operated instrument that needs to make some measurements occasionally, you might use Circuit A to switch the power-hungry microprocessor off except during those intermittent measurements. Here we've used a pMOS switch, turned on by a 1.5 V logic swing to ground; the particular part shown is specified for low gate voltage, in particular  $R_{ON}=17$  mΩ (max) at  $V_{GS}=-1.5$  V. The "1.5 V logic" is micropower CMOS digital circuitry, kept running even when the microprocessor is shut off (remember, CMOS logic has zero static dissipation).

An important point: you have to worry about proper switch operation at lower voltages, if the "1.5 V supply" is in fact an alkaline battery, with an end-of-life voltage of ~1.0 V. In that case you may be better off using a *pnp* transistor – see the discussion in "Low-voltage switching: MOS versus BJT" in Chapter 3x.

In the second circuit (B), we're switching dc power to a load that needs +12 V at considerable current; maybe it's a radio transmitter, or whatever. Because we have only a 3.3 V logic swing available, we've used a small *npn* current sink to generate an 8 V negative-going swing (relative to +12 V) to drive the pMOS gate. Note the high-value

collector resistor, perfectly adequate here because the pMOS gate draws no dc current (even a beefy 10 A brute), and we don't need high switching speed in an application like this.

The third circuit (C) is an elaboration of circuit B, with short-circuit current limiting courtesy of the *pnp* transistor. That's always a good idea in power-supply design – for example, it's impressively easy to slip with the oscilloscope probe. In this case, the current limiting also prevents momentary short-circuiting of the +12 V supply by the initially uncharged bypass capacitor. See if you can figure out how the current limiting circuit works.

**Exercise 3.16.** How does the current-limiting circuit work? How much load current does it allow?

An interesting detail: in Circuits B and C we could have hooked up the driver transistor as a *switch* (instead of a current source), omitting the emitter resistor and adding a current-limiting base resistor of 100k or so. But that circuit would create problems if you attempted to operate the circuit from a higher supply voltage, owing to the limited gate breakdown voltages of MOSFETs ( $\pm 20$  V or less). It would also defeat the current-limit scheme of Circuit C. You could fix those problems by adding a resistor directly in series with the collector, tailoring its value for correct gate drive; but the current-source scheme we've used solves these problems automatically, and it can be used to switch 24 or 48 V with no component changes.

**Exercise 3.17.** You have a dc source using full-wave rectified 120 Vac. Design a 155 to 175 V version of Figure 3.106C to pulse 0.5 A into a flash string consisting of 38 white LEDs in series. Explain your choice for  $R_1$  and  $R_2$  and the ratio  $R_2/R_1$ . Select  $Q_1$  and  $Q_2$  and evaluate their power dissipation. Use Table 2.1 on page 74, along with the MOSFET tables in this chapter. *Extra credit:* evaluate  $Q_2$ 's worst-case heating with a 10 ms maximum flash length (hint: use the datasheet's plot of "Transient Thermal Impedance").

There is still a problem (uh, an "issue"?") with Circuit C, namely the large power dissipation in pass transistor  $Q_2$  under fault conditions such as a shorted output. The brute-force approach (which we've adopted more often than we care to admit) is to use a husky MOSFET with enough heatsinking to handle  $P=V_{IN}I_{lim}$ ; that works OK for modest voltages and currents. Better to add foldback current limiting, as in Figure 12.45C on page 823. But ideally we'd want something like a pass transistor with internal thermal limiting.<sup>102</sup> That's one benefit of devices like that in Circuit E.

A popular alternative, at least for low voltage switching, is the use of a low  $R_{ON}$  analog switch (recall Table 3.3 on page 176), as in Circuit D. The switch listed there operates with power-supply voltages from 1.1 V to 4.5 V, with a worst-case  $R_{ON}$  that is plenty good enough to power loads up to 100 mA or so. It may seem strange to be using an analog switch, designed with complementary *n*-channel and *p*-channel MOSFETs for good signal properties over the full rail-to-rail range, as a simple positive-voltage power switch; but these things are inexpensive, and they take care of the logic interfacing and other details for you, so why not?

In Circuit E we've shown the interesting alternative of an *n*-channel MOSFET switch, for which you need to generate a gate drive that is more positive than the input supply voltage, preferably by a healthy 10 volts or so. You can get "high-side driver" ICs for this job, in varieties that cope with speed and voltage tradeoffs (for example the LM9061 in Figure 3.96; see also §§3.5.3, 12.4.2, 12.4.4, and Table 12.5 on page 826). Here we've gone a step further, using a high-side driver that includes the power MOSFET as well. It gets its gate drive signal with an internal oscillator and charge-pump converter (of the sort we saw in §3.4.3D). This particular device is intended for low-voltage operation, and includes internal current limiting and overtemperature protection.

Why bother with all this, when a *p*-channel MOSFET is easier to drive? Although the use of an nMOS switch with high-side drive does add complexity, it benefits from the better characteristics and much wider variety of *n*-channel MOSFETs; it is generally the preferred scheme.

Finally, Circuit F shows how to switch a negative supply rail to a load; it's analogous to Circuit B, but with an *n*-channel switch and a grounded-base *pnp* transistor to convert a positive logic level into a sourcing current that creates a 10 V gate swing across  $R_2$ . You can (and probably should) add current limiting in the manner of Circuit C.

## B. Floating power switches

Sometimes you need to switch a voltage (and its load) that's "floating" far from ground. For example, you might want to test the pulsed power capability of a resistor while sensing the current at the low side; or you might want to make millisecond-scale pulsed measurements of a transistor to circumvent heating effects; or you might want a general-purpose floating two-terminal switch that can handle ac or dc. In such situations you can't use the basic ground-referenced schemes of Figure 3.106. Figure 3.107 shows two straightforward approaches, both using an optoisolator (§12.7) to convey the switching command from its ground-referenced home to the floating switch circuit.

<sup>102</sup> If instead we were switching the *low* side (with an *n*MOS switch), we could use a protected MOSFET; see Table 12.4 on page 825.

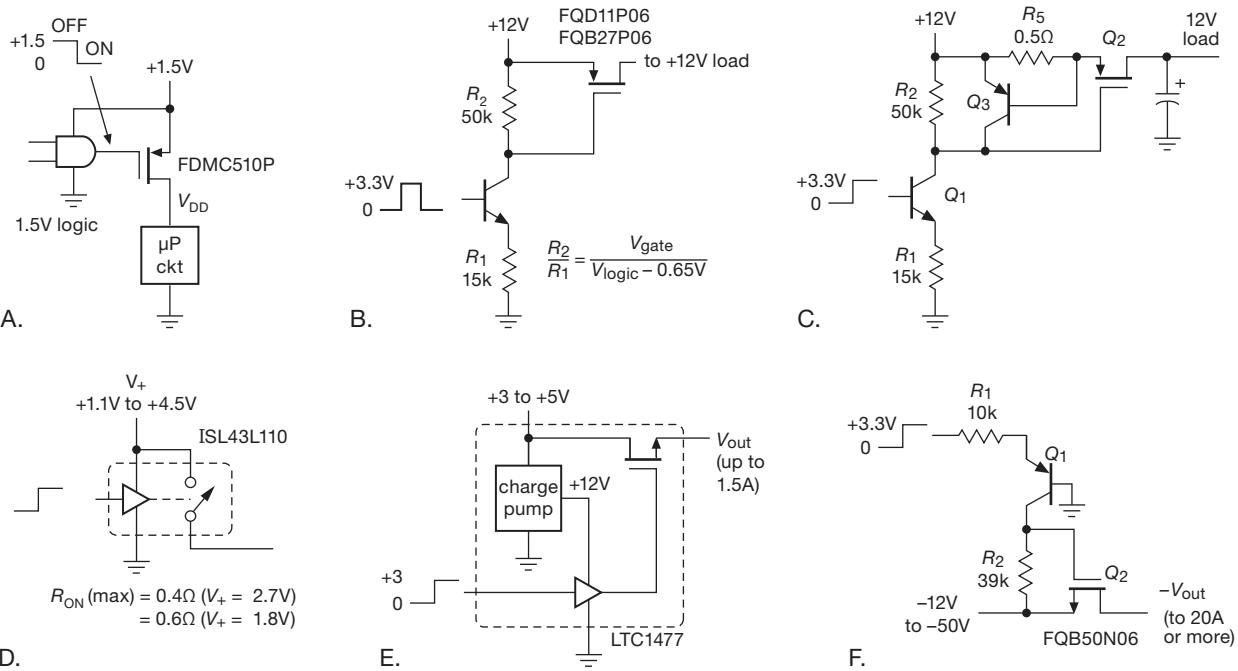


Figure 3.106. dc power switching with MOSFETs.

In Circuit A the gates of a pair of series-connected *n*-channel power MOSFETs are driven by a push–pull BJT follower that receives its base drive signal from a self-generating (“photovoltaic”) optocoupler,  $U_2$ . The latter uses a series-connected photovoltaic stack to generate a floating  $\sim 8$  V signal in response to a 10 mA LED input drive current (see Figure 12.91, and discussion in §1x.7), with some internal circuitry to enhance the turn-off time. The gate driver pair  $Q_1Q_2$  could be omitted at the expense of greater switching time (see below). These drivers reduce the effective load capacitance of the MOSFETs by a factor of beta, so that the resulting switching times (with typical power MOSFETs for  $Q_3$  and  $Q_4$ ) are limited by the optocoupler’s intrinsic speed, of order 200  $\mu$ s.

Of course the gate drivers  $Q_1$  and  $Q_2$  need a floating voltage source, here provided by a second inexpensive photovoltaic generator  $U_1$ , which need not be fast (assuming the circuit is not operated at a rapid switching rate) since it serves only to keep  $C_1$  charged to  $\sim 8$  V. You can substitute a floating 9 V battery for  $U_1$ : it can provide lots more current than  $U_1$ ’s feeble  $\sim 20\ \mu$ A output, but of course you have to replace it from time to time (an alkaline 9 V “1604”-style battery is good for about 500 mAh, and has a shelf life of 5 years or so). This circuit can switch either polarity – when ON, the series MOSFETs sum to  $2R_{\text{ON}}$  (the body diodes conduct only during ON–OFF transitions, or

at very high currents). Note that this circuit is an “unprotected” switch – there’s no provision for current or power limiting of the output transistors.

Circuit B addresses this vulnerability and leverages the benefits of the BTS555 integrated protected switch. Here we’ve taken the simple approach of a floating 9 V battery to supply its internal circuitry’s operating power (15  $\mu$ A typical when off, 1 mA on). This thing is protected against pretty much anything bad that you can throw at it. Its switching speed is comparable to that of Circuit A (typically 300  $\mu$ s on, 100  $\mu$ s off), and it’s good for lots of current (100 A or more); but it’s limited to 34 V across its switch terminals. See §12.4.4 for more details, and Table 12.5 on page 826 for additional component suggestions.

Returning to Circuit A in Figure 3.107, what sort of performance can you get with readily available *n*-channel MOSFETs? Here’s a selection of candidates (see Table 3.5 on page 206), culled from the many thousands of possibilities,<sup>103</sup> spanning the full range of voltages.<sup>104</sup>

Several trends are clearly evident:

<sup>103</sup> A search today for *n*-channel discrete MOSFETs finds 20,330 types at Digi-Key, 11,662 at Mouser, and 4,607 at Newark. A bit of an overestimate, because different package options are listed separately – but you get the point.

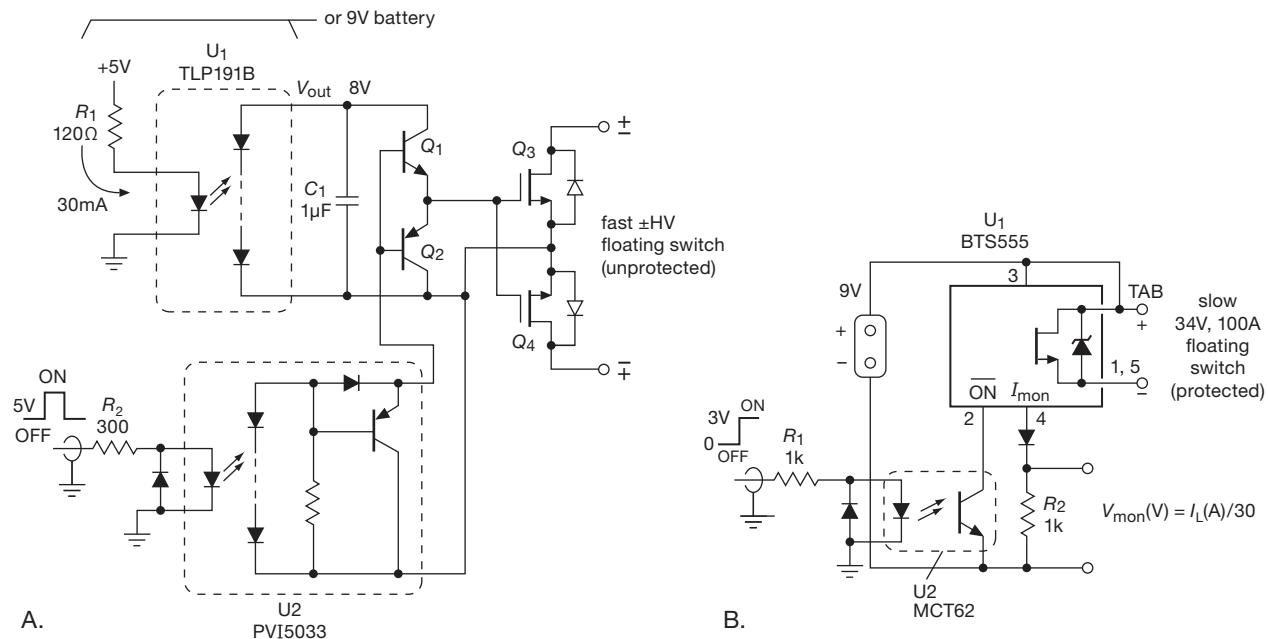


Figure 3.107. Floating MOSFET power switches: A. bipolarity, unprotected; B. unipolarity, limited voltage, protected.

- (a) There's a severe tradeoff of  $R_{ON}$  versus voltage rating – for the listed types the ON-resistance spans a  $\sim 100,000:1$  range over the  $\sim 100:1$  voltage range.
- (b) You also pay a price, literally, for very high-voltage parts; for example, the listed 4.5 kV part costs \$22.
- (c) Higher current parts have greater output capacitance (which is what you see across the switch terminals when OFF), even here where we've carefully selected parts to minimize the  $R_{ON}C_{oss}$  figure-of-merit tradeoff. They also have larger input capacitance and gate charge, which are relevant to switching speed.
- (d) And, important data is *missing!* You need to go to the datasheets for important information such as thermal resistance, pulse current and pulse energy specifications, gate charge, and the like. The data presented here are at most advisory, and you need to use detailed specifications in the circuit context to predict actual performance. For example, the “maximum pulsed current” specs generally apply to pulse lengths somewhat shorter than this circuit can produce; and the  $R_{ON}$  spec assumes 10 V of gate drive, also somewhat greater than we have here.

Let's finish this example by estimating the switching speed of the circuit of Figure 3.107A. Imagine we want 600 V capability and choose the middle-of-the-road FCP22N60N, a MOSFET that delivers a good combination

of ON-resistance and capacitance at modest cost (about \$5, qty 100). For switching speed the relevant parameter here is gate charge ( $Q_{GS}+Q_{GD}$ ), approximately 25 nC according to tabulated and graphical data. That must be supplied by the isolating driver  $U_2$ , boosted by the current gain  $\beta$  of  $Q_1$  and  $Q_2$ . From  $U_2$ 's datasheet we can estimate the output sourcing current (from its “Typical Response Time” plot) as approximately 3 μA. If for the moment we imagine that  $Q_1$  and  $Q_2$  were omitted, with  $U_2$  driving the MOSFET gates directly, the turn-on time would be  $t \approx Q_{gate}/I_{U2}$ , or 8.3 ms. Now magically restore the BJT drivers, and the estimated switching time drops by a factor of beta; for typical  $\beta \sim 200$  it becomes  $\sim 40 \mu s$ .<sup>105</sup>

Not so fast! Look again at the datasheet for  $U_2$ , you'll find that the turn-on time bottoms out at about 100 μs; likewise, its intrinsic turn-off time is about 350 μs, even with small load capacitance. Those numbers dominate the performance of Circuit A, for nearly every MOSFET listed, assuming of course that the BJT driver  $Q_1Q_2$  is included. If you can tolerate slower switching, you can simplify things by omitting the drivers and their floating power source.

If you need faster switching, there are many integrated high-side driver chips that can do the job, for example the series of “high-voltage gate-driver ICs” from Interna-

<sup>104</sup> See the MOSFET tables in this chapter and in Chapter 3x for additional data on these and other power MOSFETs.

<sup>105</sup> If you look at the MOSFET drivers in Table 3.8 on page 218, you'll see that the ZXGD3002–04 are simply a pair of very-high-gain *npn* and *pnp* BJTs, in SOT23-6 packages, perfect for  $Q_1$  and  $Q_2$ .

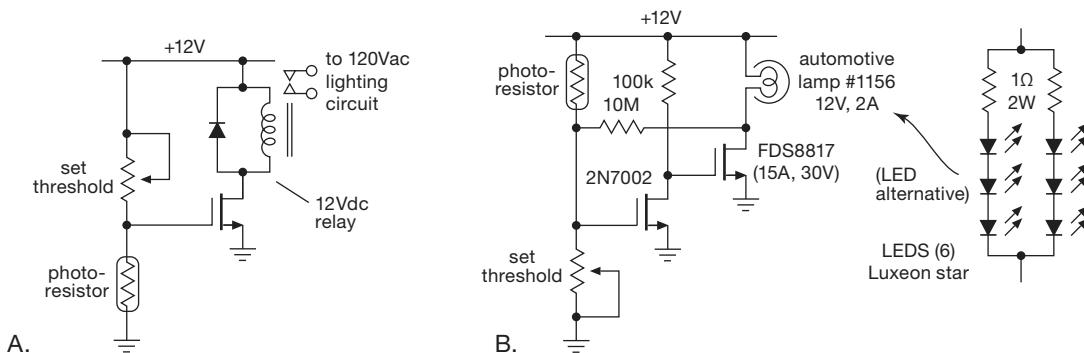


Figure 3.108. Ambient-light-controlled power switch.

Table 3.5 MOSFET switch candidates<sup>a</sup>

Part #	<i>BV<sub>DS</sub></i> (V)	<i>I<sub>D(max)</sub></i> pulse (A)	<i>I<sub>D(max)</sub></i> cont (A)	2 <i>R<sub>ON</sub></i> (Ω)	<i>C<sub>oss</sub></i> <sup>t</sup> (pF)
IXTT02N450\$\$	4500	0.6	0.2	960	19
IXTH02N250\$	2500	0.6	0.2	770	9
STW4N150	1500	12	2	10	120
IXTP3N120	1200	12	3	6.5	100
IXFH16N120P\$	1200	35	10	1.7	390
IRFBG20	1000	5.6	1	16	52
IRFBG30	1000	12	2	8	140
IXFH12N100\$	1000	48	5	2	320
IPP60R520CP	650	17	4	1	32
FCP22N60N	600	66	12	0.28	76
FCH47N60N\$	600	140	30	0.1	200
IRF640N	200	72	12	0.24	190
FQP50N06L	60	210	25	0.08 <sup>d</sup>	450
IRLB3034	40	1400	125	0.003 <sup>e</sup>	2000
FDP8860	30	1800	100	0.004 <sup>d</sup>	1700

**Notes:** (a) all are in TO-220 or TO-247 pkgs. (b) *italics* designate maximum pulsed drain current, for pulse width specified in the part's datasheet (e.g., 80μs); **boldface** designates maximum continuous drain current at  $T_J=70^\circ\text{C}$ . (c) at  $V_{GS}=10\text{V}$ , unless marked otherwise. (d) at  $V_{GS}=5\text{V}$ . (e) at  $V_{GS}=4.5\text{V}$ . (t) typ. (\$ not inexpensive. (\$\$) expensive.

tional Rectifier. These use internal high-voltage transistors to send the control signals up to the high side, with maximum voltage ratings most commonly of 600 V. These typically have switching times in the range of 100 ns to 1 μs. They are intended for cyclic applications such as pulse-width-modulated bridge drivers and use high-side charge pumps to develop the over-the-rail gate drive voltage; but you can adapt them for pulsed applications by substituting a flying 9 V battery, as we've done here. See Chapter 3x's

table “High-voltage Half-Bridge Drivers,” and the section on “High-side high-voltage switching” in Chapter 9x.

Another class of ICs that can be used in applications like this is typified by Avago’s ACPL-300 series of “gate drive optocouplers,” which combine an optocoupler and isolated push-pull output stage. For example, the ACPL-W343’s output stage can source or sink 3 A (minimum), with 40 ns rise and fall times (into a load of 25 nF in series with 10 Ω), and isolation good to 2 kV. You have to supply 15–30 V isolated dc for the output stage,<sup>106</sup> in the manner of Figure 3.107B, with the usual bypass capacitor (for peak output currents); the quiescent current is 2 mA, good for 200 hours of operation if you use a pair of 9 V batteries. See §12.7.3 and Figure 12.87 for additional circuit discussion and suggestions.

### C. Some unusual switching examples

#### Light-at-night

Figure 3.108A shows a simple MOSFET switching example, one that takes advantage of the high gate impedance. You might want to turn on exterior lighting automatically at sunset. The photoresistor has low resistance in sunlight, high resistance in darkness. You make it part of a resistive divider, driving the gate directly (no dc loading!). The light goes on when the gate voltage reaches the value that produces enough drain current to close the relay. Sharp-eyed readers may have noticed that this circuit is not particularly precise or stable; that’s OK, because the photoresistor undergoes an enormous change in resistance (from 10k to 10M, say) when it gets dark. Note that the MOSFET may have to dissipate some power during the time the gate bias is inching up, since we’re operating in the linear region;

<sup>106</sup> A bit high for MOSFET gate driving (these are targeted at IGBTs); but you can get parts with lower minimum output supply voltage, for example the HCPL-3180 or the PS9506 from Renesas (both can operate with 10 V min).

but it's switching only a relay, not the power load, so this is of little concern. The circuit's lack of a precise and stable threshold means that the light may turn on a few minutes early or late – again, no big deal. But an additional worry is the behavior of the relay, which is not happy with a marginal coil-driving voltage (which holds the contacts closed with less than the rated mechanical force, thus potentially shortening the relay's life; see the discussion of relays in Chapter 9x).

These problems are remedied in Figure 3.108B, where a pair of cascaded MOSFETs delivers much higher gain, augmented by some positive feedback through the 10M resistor; the latter adds hysteresis, which causes the circuit to snap on regeneratively as it reaches threshold.

### Hefty piezo driver

Figure 3.109 shows a real power MOSFET job: a 200 watt amplifier to drive a piezoelectric underwater transducer at 200 kHz. We've used a pair of hefty nMOS transistors, driven alternately to create ac drive in the (high-frequency) transformer primary. The series inductor in the secondary resonates with the transducer's capacitance to step up the voltage across the piezo to several kilovolts. The TC4425A is a handy “3A Dual High-Speed Power MOSFET Driver” (like the TC4420 in Figure 3.97), which takes a logic-level input (0 V=LOW,  $\geq 2.4$  V=HIGH), and generates a full-swing (0 to  $+V_{DD}$ ) output pair, one inverted and the other noninverted; see Table 3.8 on page 218. It's needed to overcome capacitive loading, since the MOSFETs must be turned on fully in a fraction of a microsecond. The diodes shunted across the series gate resistors cause a rapid turn-off to prevent undesirable conduction overlap of the power transistors.

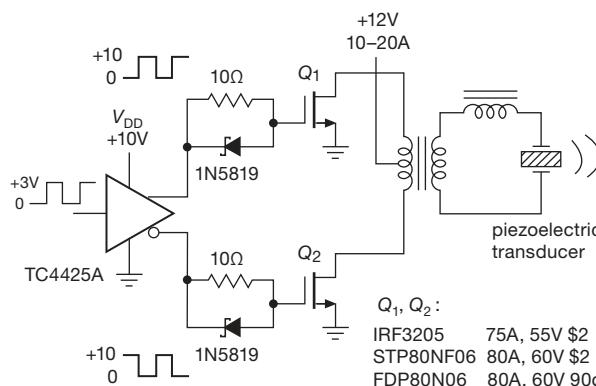


Figure 3.109. MOSFET piezo power driver.

### 3.5.7 IGBTs and other power semiconductors

The contemporary power MOSFET is a versatile transistor for both power switching applications (e.g., dc power control, or dc–dc switching converters), and for linear power applications (such as audio amplifiers). But there are some drawbacks and some useful alternatives.

#### A. Insulated-gate bipolar transistor (IGBT)

The IGBT is an interesting MOSFET-bipolar hybrid, most simply described as an integrated complementary-Darlington-like (Sziklai) connection of an input MOSFET with a power bipolar transistor (Figure 3.110). So it has the input characteristics of a MOSFET (zero dc gate current), combined with the output characteristics of a power bipolar transistor; note, however, that it cannot saturate to less than  $V_{BE}$ . Unlike MOSFETs, IGBTs do not have an intrinsic reverse diode, so inductive ringing, etc., can easily exceed the reverse voltage rating (e.g., 20 V). Many IGBTs include an internal “anti-parallel” diode to protect against this problem.<sup>107</sup>

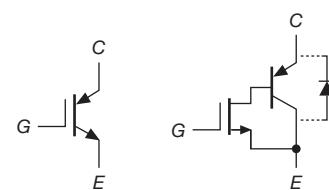


Figure 3.110. IGBT symbol and simplified equivalent circuit showing the optional “anti-parallel” diode.

Nearly all available IGBTs come in the nMOS-*pnp* polarity only, and thus behave as *n*-type devices.<sup>108</sup> They are generally high-voltage and high-power devices, available in discrete transistor power packages like the TO-220, TO-247, and in surface-mount packages like the D<sup>2</sup>PAK and SMD-220, with ratings to 1200 V and 100 A. For higher currents you can get them in larger rectangular power “modules,” with higher voltage ratings and with current ratings to 1000 A or more.

IGBTs excel in the arena of high voltage switching, because high-voltage MOSFETs suffer from greatly increased  $R_{ON}$ : an approximate rule-of-thumb for MOSFETs is that the  $R_{ON}$  increases as the square of the voltage

<sup>107</sup> Some parts are available with and without the added diode, indicated for example by a -D suffix on the part number.

<sup>108</sup> Currently the only *p*-type IGBTs we know of are the Toshiba GT20D200 series.

rating.<sup>109</sup> For example, let us compare two power products from International Rectifier (along with a BJT of comparable ratings):

	<b>MOSFET</b>	<b>IGBT</b>	<b>BJT</b>
<b>Type</b>	IRFPG50	IRG4PH50S	TT2202
$V_{max}$	1000 V	1200 V	1500 V
$I_{max}$	dc pulse	6.1 A 24 A	10 A 25 A
$R_{ON}$ (typ)	25°C 150°C	1.5 Ω 4 Ω	— —
$V_{ON}$	25°C (typ, 15 A) 150°C	23 V 60 V	1.2 V 1.2 V
			1 V (@8 A) 1 V (@8 A)

These are comparably priced (about \$5) and packaged (TO-247), have similar input characteristics (2.8 nF and 3.6 nF input capacitance), and the resulting saturation voltages  $V_{ON}$  when switching 15 A are shown for the same full input drive of  $V_{in} = +15$  V. The IGBT is the clear winner in this high-voltage and high-current regime.<sup>110</sup> And, when compared with a power BJT, it shares the MOSFET advantage of high static input impedance (though still exhibiting the drastically reduced dynamic input impedance during switching, as we saw in §3.5.4B). The BJT does have the advantage of lower saturation voltage (the IGBT's  $V_{ON}$  is at least  $V_{BE}$ ) and lower drive voltage (see the figure in Chapter 3x's section "Power transistors for linear amplifiers"), at the expense of a high static driving current; the latter drawback is exacerbated at high currents, where BJT beta drops rapidly. A saturated BJT also suffers from slow recovery owing to stored charge in the base region.

With the very high voltages and currents encountered with IGBTs, it is mandatory to include fault protection in the circuit design: an IGBT that may be required to switch up to a 50 A load from a 1000 V supply will be destroyed in milliseconds if the load becomes short-circuited, owing to the 50 kW (!) power dissipation. The usual method is to shut off the drive if  $V_{CE}$  has not dropped to just a few volts after 5  $\mu$ s or so of input drive (see Figure 12.87B). We'll revisit these three power transistor technologies in Chapter 3x.

## B. Thyristors

For the utmost in *really* high-power switching (we're talking kiloamperes and kilovolts), the preferred devices are

the *thyristor* family, which include the unidirectional "silicon controlled rectifiers" (SCRs) and the bidirectional "triacs." These three-terminal devices behave somewhat differently from the transistors we've seen (BJTs, FETs, and IGBTs): once triggered into conduction by a small control current (a few milliamps) into their control electrode (the *gate*), they remain ON until external events bring the controlled current (from *anode* to *cathode*) to zero. They are used universally in house-current lamp dimmers, where they are switched on for a fraction of each half-cycle of ac line voltage, thus varying the *conduction angle*.

Thyristors are available in ratings from 1 A to many thousands of amperes, and voltage ratings from 50 V to many kilovolts. They come in small transistor packages, the usual transistor power packages, larger modules, and really scary "hockey puck" packages that are capable of switching megawatts. These are hefty devices; you can hurt yourself just by dropping one on your foot.

## 3.6 MOSFETs in linear applications

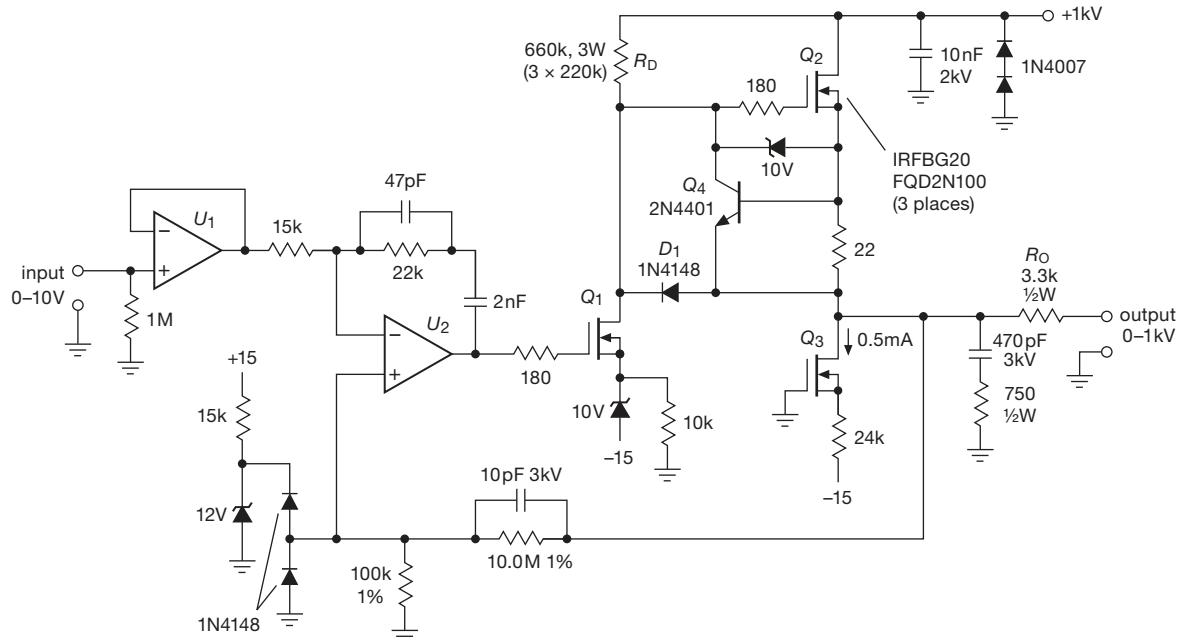
Although we dealt extensively with linear applications in this chapter's treatment of *JFETs*, our discussion of MOSFETs has concentrated almost entirely on switching applications. Lest we leave the wrong impression, we address in this section some applications of discrete power MOSFETs to linear applications, particularly those that benefit from their unique properties. See also additional applications to linear voltage regulators in Chapter 9 (e.g., Figures 9.17, 9.20, 9.104, 9.110, and 9.113), and advanced material in §3x.5.

### 3.6.1 High-voltage piezo amplifier

A nice application of MOSFETs as linear amplifiers takes advantage of available types with high voltage ratings, and their immunity to second breakdown. Ceramic *piezoelectric* transducers are often used in optical systems to produce controlled small motions; for example, in *adaptive optics* you might use a piezoelectrically controlled "rubber mirror" to compensate for local variations in the index of refraction of the atmosphere. Piezo transducers are nice to use, because they're very stiff. Unfortunately, they may require as much as a kilovolt of drive to produce significant motions. Furthermore, they're highly capacitive – typically 0.01  $\mu$ F or more – and have mechanical resonances in the kilohertz range, thus presenting a nasty load. We needed dozens of such driver amplifiers, which for some reason cost a few thousand dollars apiece if you buy them commercially.

<sup>109</sup> You find exponents from 1.6 to 2.5 in the literature; the lower end of this range is likely to be more accurate; see our plots in Chapter 3x.

<sup>110</sup> Where it excels also in maintaining high transconductance, compared with the MOSFET. The advantage goes to the IGBT, starting around 200 V. See also the comparison table in §3.5.5.



**Figure 3.111.** 1 kV low-power piezo driver with totem-pole output stage. A similar design is used for the high-voltage regulated dc supply shown in Figure 9.110.

We solved our problem with the circuit shown in Figure 3.111. The IRFBG20 is an inexpensive ( $\sim \$2$ ) MOSFET, good for 1 kV and 1.4 A; the similar FQD2N100 (1 kV, 1.6 A) costs about \$0.85. The first transistor is a common-source inverting amplifier, driving a source follower with active current-sink load. The *n*p*n* transistor is a current limiter and can be a low-voltage unit, since it floats on the output. One subtle feature of the circuit is the fact that it's actually push-pull, even though it looks single-ended: you need plenty of current (20 mA) to push 10,000 pF around at 2 V/ $\mu$ s; the output transistor can *source* current, but the pulldown resistor can't sink enough (look back to §2.4.1, where we motivated push-pull with the same problem). In this circuit the driver transistor is the pull-down, with conduction through the gate-source diode!<sup>111</sup> The rest of the circuit involves feedback with op-amps, a forbidden subject until the next chapter; in this case the magic of feedback makes the overall circuit linear (100 V of output per volt of input), whereas without it the output voltage would depend on the (nonlinear)  $I_D$ -versus- $V_{GS}$  characteristic of the input transistor. A nice improvement to this circuit consists of replacing the 660 kΩ 3 W pullup resistor (whose current drops at high output

voltages, e.g., to 0.15 mA at 900 V) with a depletion-mode MOSFET current source set to, say, 0.25 mA. See the discussion below (§3.6.2C, and also Figure 3.23, Table 3.6, and §§9.3.14C and 3x.6).

For a detailed analysis of a precision high-voltage amplifier with bipolarity output capability, see the section with that name in Chapter 4x. For an analysis of relevant issues, such as MOSFET transconductance at low currents, see §3x.5.2, and for the response of a MOSFET source follower into a capacitive load see §3x.8.

**Exercise 3.18.** Modify this circuit so the high-voltage output can be turned on and off, under control of an input signal (0 V for off, +3V for on).

### 3.6.2 Some depletion-mode circuits

Depletion-mode MOSFETs are the neglected siblings of the far more popular enhancement-mode MOSFETs. They can do some nice tricks, though, that are worth knowing about. And they are available in high voltage (to 1 kV) and high current (to 6 A) varieties. Table 3.6 on the following page lists nearly all available parts of this somewhat rarefied species. Here are some applications that exploit their property of conduction at zero gate voltage.

<sup>111</sup> This is called a “totem-pole” output stage, and became popular in the early 1970s in bipolar TTL logic, see Figure 10.25A.

**Table 3.6 Depletion-mode *n*-channel MOSFETs**

Part #	Manuf	Available packages						$I_{DSS}$ min	$R_{ON}^t$ @ $V_{GS}=0$	$g_m^t$ @ $I_D$	$V_{GS(th)}$ min	$C_{iss}$ typ	$C_{oss}$ typ	$C_{rss}$ typ	Cost <sup>y</sup>		
		TO-92	SOT-23	SOT-220	TO-247	D-Pak	D <sup>f</sup>										
		(V)	(W)	(mA)	(mA)	(Ω)	(mS)										
<i>small, RF</i>																	
BF998	NXP	- f	- - - - -	- - - - -	12	-	2	18	-	24	0.01	-	-2	2.1	1.1	0.025	
BF999	Vishay	- •	- - - - -	- - - - -	20	-	5	18	-	16	0.01	-	-2.5	2.5	1.0	0.025	
SKY65050 <sup>s</sup>	Skyworks	- g	- - - - -	- - - - -	6	-	40	70	-	80	0.12	-1	-0.7	really small			
<i>small</i>																	
DN2470K4	Supertex	- - - - -	- - - - -	•	700	2.5	500 <sup>t</sup>	42 <sup>m</sup>	100 <sup>n</sup>	0.1	-1.5	-3.5	540	60	25	0.81	
BSS126	Infineon	- •	- - - - -	- - - - -	600	0.5	7	-	320	17	0.01	-2.7	-1.6	21	2.4	1	0.13
BSP135	Infineon	- - - - •	- - - - -	-	600	1.8	20	-	30	160	0.1	-1.8	-1.0	98	8.5	3.4	1.38
LND150	Supertex	◊ k	• - - - -	- - - - -	500	0.7	1	3	850	2	0.001	-1.0	-3	7.5	2	0.5	0.58
DN3145	Supertex	- - •	- - - - -	-	450	1.3	120	-	60 <sup>m</sup>	140 <sup>n</sup>	0.1	-1.5	-3.5	120	15	10	0.68
DN3545	Supertex	• - ◊	- - - - -	-	450	1.6	200	-	20 <sup>m</sup>	150 <sup>n</sup>	0.1	-1.5	-3.5	360	40	15	0.74
DN2540	Supertex	• - ◊	- • - - -	-	400	1.6	150	-	17	325	0.1	-1.5	-3.5	200	12	1	0.81
DN3135	Supertex	- • ◊	- - - - -	-	350	1.3	180	-	35 <sup>m</sup>	140	0.1	-1.5	-3.5	60	6	1	0.62
CPC3720C	Clare	- - •	- - - - -	-	350	1.6	130	-	22 <sup>m</sup>	225	0.1	-1.6	-3.9	70	20	10	0.37
CPC5603C	Clare	- - - •	- - - - -	-	415	2.5	130	-	14 <sup>m</sup>	-	-	-3.6	-2	300	-	-	0.69
CPC5602C	Clare	- - - - •	- - - - -	-	350	2.5	130	-	14 <sup>m</sup>	-	-	-3.6	-2	300	-	-	0.60
DN3535	Supertex	- - •	- - - - -	-	350	1.6	200	-	10 <sup>m</sup>	200 <sup>n</sup>	0.1	-1.5	-3.5	360	40	10	0.68
BSS139	Infineon	- •	- - - - -	-	250	0.4	30	-	12.5	130	0.08	-1.4	-	60	6.7	2.6	0.57
BSP129	Infineon	- - - - •	- - - - -	-	240	1.8	50	-	6.5	360	0.28	-2.1	-1.0	82	12	6	0.47
DN3525N8	Supertex	- - •	- - - - -	-	250	1.6	300	-	6 <sup>m</sup>	225 <sup>n</sup>	0.15	-1.5	-3.5	270	20	5	0.66
CPC3703	Clare	- - •	- - - - -	-	250	1.6	300	-	4 <sup>m</sup>	225 <sup>n</sup>	0.1	-1.6	-3.9	327	51	27	0.57
BSP149	Infineon	- - - •	- - - - -	-	200	1.8	140	-	1.7	800	0.48	-1.8	-1.0	326	41	17	0.80
BSS159	Infineon	- •	- - - - -	-	50	0.4	70	-	4	160	0.16	-2.5	0.0	70	15	6	0.31
<i>large</i>																	
<i>see note x</i>																	
IXTx01N100D <sup>x</sup>	IXYS	- - - - -	P - - Y	1000	25	100 <sup>t</sup>	90	150	0.1	-2.5	-	120	25	5	0.75		
IXTx08N100D2 <sup>x</sup>	IXYS	- - - - -	P - - Y	1000	60	100 <sup>t</sup>	21 <sup>m</sup>	560	0.4	-2	-4	325	24	6.5	0.69		
IXTx1R6N100D2 <sup>x</sup>	IXYS	- - - - -	P - - Y	1000	100	100 <sup>t</sup>	21 <sup>m</sup>	1100	0.8	-2.5	-4.5	645	43	11	1.66		
IXTx3N100D2 <sup>x</sup>	IXYS	- - - - -	P - A -	1000	125	3000	-	5.5 <sup>m</sup>	4200	3	-2.5	-4.5	1020	68	17	2.11	
IXTx6N100D2 <sup>x</sup>	IXYS	- - - - -	P H A -	1000	300	6000	-	5.5 <sup>m</sup>	4200	3	-2.5	-4.5	2650	167	41	4.66	
IXTx02N50D2 <sup>x</sup>	IXYS	- - - - -	P - Y	500	25	250 <sup>t</sup>	20	150	0.2	-2	-5	120	25	5	1.05		
IXTx08N50D2 <sup>x</sup>	IXYS	- - - - -	P - A -	500	60	800	-	4.6 <sup>m</sup>	570	0.4	-2	-4	312	35	11	1.62	
IXTx1R6N50D2 <sup>x</sup>	IXYS	- - - - -	P - Y	500	100	1600 <sup>t</sup>	2.3	1750	0.8	-2	-4	645	65	17	1.66		
IXTx3N50D2 <sup>x</sup>	IXYS	- - - - -	P - A -	500	125	1600 <sup>t</sup>	1.5	2100	1.5	-2	-4	1070	102	24	2.13		
IXTx6N50D2 <sup>x</sup>	IXYS	- - - - -	P H A -	500	300	6000	-	0.5 <sup>m</sup>	4500	3	-2	-4	2800	255	64	4.66	
IXTH20N50D <sup>x,e</sup>	IXYS	- - - - -	H -	500	400	1500 <sup>t,e</sup>	0.5 <sup>m</sup>	7500	10	-1.5	-3.5	2500	400	100	8.61		

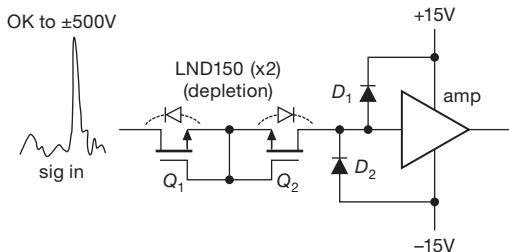
**Notes:** (c)  $P_D$  at  $T_c=25^\circ\text{C}$ , for package marked  $\diamond$ . (e) the IXTH20N50D delivers most of its current capability in the enhancement-mode region. (f) has two gates, 4-lead SOT-143 package. (g) SC-70 package. (k) unusual pinout; also try LND250K1. (m) maximum. (n) minimum. (s) SKY part's full name: SKY65050-372LF. (t) typical. (x) substitute the letter listed in the package column for the "x" in the part number; for example, IXTP01N100D is a TO-220 package. (y) quantity 100.

### A. Input protection

Low-level circuits (such as sensitive amplifiers) don't like to be driven beyond their power supply rails. One simple protection scheme uses a series resistor at the input, with a pair of downstream clamp diodes to the amplifier's supply rails. That's OK for small overdrive; but it's unsatisfac-

tory if the input may go to a few hundred volts (think *powerline!*), because the large resistor value ( $\sim 100\text{k}$  or so, to limit the fault current and dissipated power) compromises the signal bandwidth and noise. Figure 3.112 shows how to use a pair of depletion-mode MOSFETs (instead of a large resistor) as the series element. The particular part shown

is small (SOT23, SOT89, or TO-92), inexpensive (\$0.60), and able to withstand momentary inputs to  $\pm 500$  V. The pair looks like a series resistance of  $\sim 1.7$  k (twice  $R_{DS(ON)}$ ) until the input goes beyond the amplifier's rails, whereupon it limits the current through the clamp diodes to  $\sim 2$  mA. See some subtleties in "A Riff on Robust Input Protection" (§5.15.5).



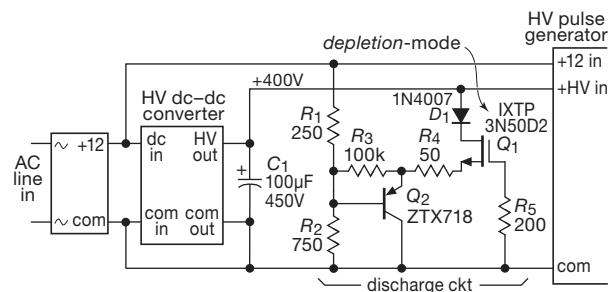
**Figure 3.112.** Protecting a low-level input from outrageous overvoltage "faults." Under normal conditions the series-connected depletion-mode MOSFETs  $Q_1$  and  $Q_2$  (with their intrinsic body diodes shown) conduct with an effective series resistance of  $R_{ON} \sim 1\text{k}\Omega$  each. An input signal beyond the amplifier's  $\pm 15$  V rails is clamped by diodes  $D_1$  or  $D_2$ , with  $Q_1$  and  $Q_2$  safely limiting the clamp current to their  $I_{DSS} \sim 2$  mA. See also Figure 5.81.

### B. HV capacitor discharge

Human contact with a circuit at a few hundred volts can be, well, a *shocking* experience. That's why it's considered good manners to arrange things so that storage capacitors charged to such voltages are discharged promptly after power is removed. Capacitors, after all, have pretty good memory – they can stay charged for hours, or even years (that's how bits are stored in "flash memory," see §14.4.5).

The traditional approach is to put a "bleeder" resistor across the storage capacitor, sized to discharge it in  $\sim 10$  s or so. Good enough. But it is not really satisfactory when you have a large-value capacitor, for example one used to store the energy for a short-duration high-voltage pulse generator. Figure 3.113 shows such an application, with a  $100\text{\textmu F}$  storage capacitor charged to  $+400$  V by a low-power dc–dc converter (say 10 W), the latter powered from a low-voltage dc supply that also powers the other pulse-generator circuitry.

What you'd like is a bleeder resistor that is connected only when the external power is removed. Here depletion-mode MOSFET  $Q_1$  is held in the nonconducting state when the supply is powered ( $V_{GS} = -9$  V), but is sent into conduction ( $V_{GS} \approx 0.6$  V) when the  $+12$  V is absent. It's rated at 500 V, 3 A  $I_{DSS(\min)}$ , and costs about \$2. (You can get depletion-mode MOSFETs up to 1 kV.) We don't need 3 A (which would discharge  $100\text{\textmu F}$  in just 13 ms); but we *do* need a MOSFET large enough to absorb the stored energy,



**Figure 3.113.** Depletion-mode MOSFET  $Q_1$  discharges the  $100\text{\textmu F}$  high voltage capacitor  $C_1$  when power is removed; when powered it is inactive.

here 8 joules – this part can absorb a 25–50 J pulse without exceeding  $T_{J(\max)}$  (see Chapter 9x). Follower  $Q_2$  boosts the discharge current, which otherwise would be just a few millamps (set by bleeder  $R_2$ ).

### C. Current source

Depletion-mode power MOSFETs make excellent 2-terminal current sources, capable of high voltages (to 1000 V for some parts, see Table 3.6) and many watts of power dissipation. They extend the basic idea, seen earlier with JFETs (§3.2.2, Figures 3.20 and 3.23), to higher voltage and power levels. Because these applications are associated with power, we defer the discussion to Chapter 9 (§9.3.14C), where you can see that the circuits are the same as with JFETs (Figure 9.36), and you can delight in curves of measured current versus voltage (Figures 9.40 and 9.41).<sup>112</sup> Such a depletion-mode MOSFET current source is ideal for an application like the high-voltage piezo driver we just saw (§3.6.1), where it can replace the primitive 660k power-resistor pullup and thereby supply approximately constant driver-stage drain current over the signal swing.

### D. Extending regulator $V_{IN}$

Sometimes you need to extend the permissible range of dc input voltage to some low-voltage device. Figure 3.114A shows an example: a linear voltage regulator (§9.3) that provides  $+3.3$  V (for example) from a higher dc input. Such regulators have limited maximum input voltage range – perhaps +20 to +30 V (if made with BJTs), or as little as +6 V (if made with CMOS). Here the *n*-channel depletion-mode MOSFET  $Q_1$  is wired as a follower, providing at the regulator's input a voltage that is greater than  $V_{OUT}$  by the

<sup>112</sup> There's plenty more detail fleshed out in Chapter 3x (§3x.6), where we show tricks for raising the output impedance (i.e., constancy of current) with the ever-wonderful cascode, raising operating voltage with a series stack, and reducing power dissipation.

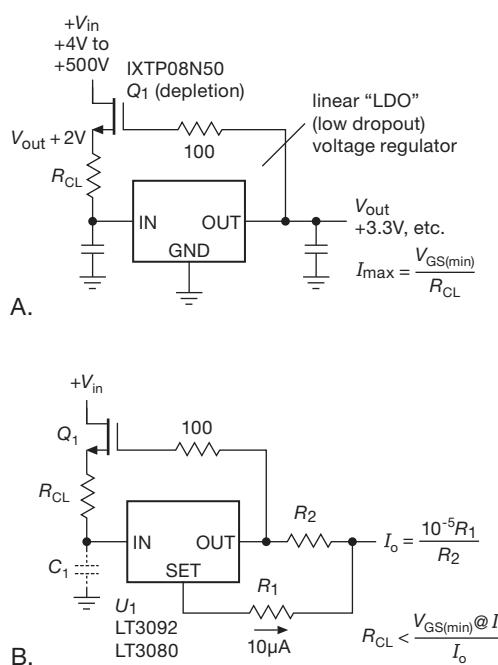
magnitude of its  $V_{GS}$ ; for the IXTP08N50  $V_{GS}$  is between  $-2$  and  $-4$  V, so the regulator's input is held between  $2$  and  $4$  V above its output. The circuit's input voltage can go to  $+500$  V ( $Q_1$ 's rated maximum), with due regard to heat dissipation of course. The resistor  $R_{CL}$  protects  $Q_1$  by setting a rough current limit. You can play the same trick with a constant-current regulator (Figure 3.114B). For details see the fuller discussion in Chapter 9 (§9.13.2). See also §3x.6 for a discussion of the use of depletion-mode MOSFETs as current sources, especially at high voltages.

### 3.6.3 Paralleling MOSFETs

You sometimes hear the statement that power MOSFETs can be paralleled directly (without ballasting resistors in the source leads), because their negative temperature coefficient of  $I_D$  at fixed  $V_{GS}$  guarantees automatic redistribution of drain currents in a paralleled array. Furthermore, the story goes, the same property prevents thermal runaway.

#### A. As switches – Yes!

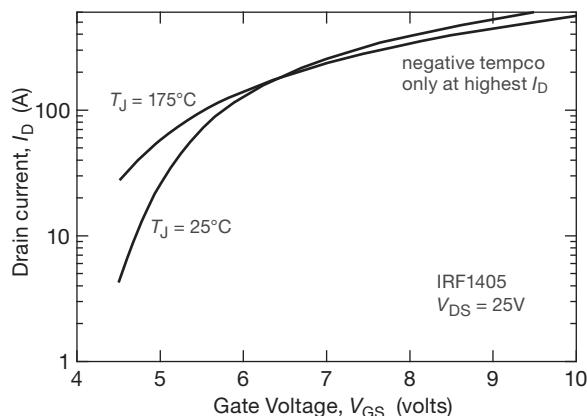
Power MOSFETs exhibit negative tempco of  $I_D$  – but only at high drain currents (or, more accurately, at relatively



**Figure 3.114.** A. A high voltage depletion-mode MOSFET extends the input voltage range of a series voltage regulator. B. An analogous circuit for a current source. For good performance at high frequencies  $C_1$  should be small, or even eliminated altogether. See Figure 9.104, where the  $is$  are dotted,  $ts$  are crossed, and all is explained.

large values of  $V_{GS}$ ), as seen in Figure 3.115. For *switching* applications, in which you are operating at essentially zero  $V_{DS}$  (limited by  $R_{ON}$ ), the large gate drive puts the device into the region of negative  $I_D$  tempco, so you *can* (and should) simply tie multiple MOSFETs in parallel, with no ballasting resistors.<sup>113</sup> Here  $R_{ON}$  increases with increasing temperature (Figure 3.116), and the parallel connection shares drain current (and power) properly.

The positive tempco of  $R_{ON}$ , though helpful for parallel



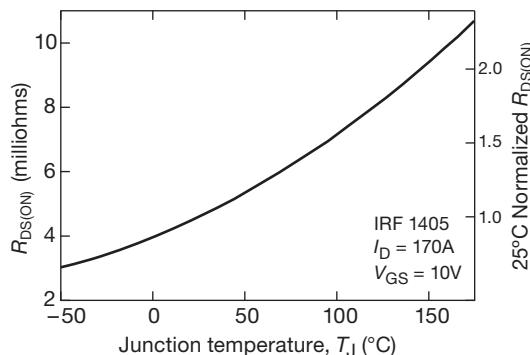
**Figure 3.115.** Transfer characteristics ( $I_D$  versus  $V_{GS}$ ) for the IRF1405 *n*-channel power MOSFET. Note that the temperature coefficient is *positive* except at the highest drain currents ( $>175$  A); for *linear* applications you would rarely exceed 10 A of drain current.

operation of MOSFET switches, creates a new problem, namely the possibility of *thermal runaway*. See §3.6.4.

#### B. In linear power circuits – No!

Here the situation is more complicated: in most *linear* applications (e.g., audio power amplifiers, in which there is substantial voltage  $V_{DS}$  across the transistor) you are operating in the positive tempco region of relatively low drain currents – because otherwise the power dissipation ( $I_D V_{DS}$ ) would be much greater than allowable by thermal considerations (i.e., excessive junction temperature; see §9.4.1A). For example, the transistor of Figure 3.115 is limited to 200 watt dissipation at a case temperature of  $75^\circ\text{C}$ ; so, in a circuit with 25 V across it, the average drain current is limited to 8 A, in which  $I_D$  has a large positive tempco. So, for practical linear applications – in which you operate

<sup>113</sup> However, each FET should have its own series gate resistor to prevent oscillation during the switching transitions; these are typically in the range of a few ohms to a few tens of ohms, and they should ordinarily be used as well for single switching MOSFETs. Ferrite beads on the gate or drain leads can also be helpful to tame oscillations.



**Figure 3.116.** ON-resistance increases with increasing temperature:  $R_{ON}$  versus temperature for the IRF1405 *n*-channel power MOSFET.

with substantial  $V_{DS}$  – unequal current sharing of paralleled MOSFETs is in fact exacerbated. And, because you’re using multiple transistors exactly because a single one won’t handle the power, the circuit is in serious trouble; a single transistor will likely hog an excessive share of the current, putting its dissipation well over the limit set by thermal resistance and heatsinking.

#### Source-ballast resistors

The solution is to add small ballasting resistors in the individual source leads, chosen roughly so that the voltage drop across them is at least comparable to the scatter in gate-source operating voltages (Figure 3.117A). We’ve found that a few tenths of a volt drop is frequently adequate for MOSFETs of a given type, from a single manufacturing batch or from transistors selected for matched  $V_{GS}$ ;<sup>114</sup> however, the datasheet specs would suggest (conservatively) larger drops – a volt or two at full operating current. Unless you are willing to worry about matched transistors (both during initial construction, and later replacement), you should take the conservative approach to produce a robust design, with source-ballasting resistors sized to drop a volt or two at the currents where power dissipation becomes important.

This example illustrates a frequent designer’s quandary, namely a choice between a conservative circuit that meets the strict worst-case design criterion, and is therefore *guaranteed* to work, and a better-performing circuit design that doesn’t meet worst-case specifications, but is overwhelmingly likely to function without problems. There are times when you will find yourself choosing the latter, ignoring the little voice whispering into your ear.

<sup>114</sup> In the example in §3.6.3A you might put four IRF1405s in parallel, with  $0.1\ \Omega$  10 W resistors in each source lead to handle a total current of 25A.

#### Active feedback

This current-matching problem exemplifies a typical circuit tradeoff of robustness versus performance: a conservatively large ballast drop produces increased  $R_{ON}$  and power dissipation. As is often the case, a clever circuit can recover the lost benefits.

Figure 3.117B shows a nice solution, another in our “Designs from the Masters” series; just a snippet this time, but a valuable snippet. The small current-sense resistors in the MOSFET source leads provide active feedback through a primitive differential amplifier. Compared with a conservative source-ballasting circuit (Figure 3.117A), in which the source resistors are chosen to provide a 2 V drop (at a nominal 1 A per transistor operating current), the active circuit uses much smaller  $0.1\ \Omega$  sense resistors, providing 100 mV of drop, which is applied to the *npn* differential pair to adjust the gate voltages as needed to equalize the source currents. This circuit requires larger gate drive voltage, which is seldom a problem; in exchange, it minimizes the voltage drop and impedance in the MOSFET’s high-current path. This scheme is well suited to relatively slow circuits, for example the series pass element in a linear power supply. Note that this arrangement is easily expanded to any number of MOSFETs.<sup>115</sup>

There’s a pleasant exception to this general characteristic of positive  $I_D$  tempco in power MOSFETs: *lateral* devices (as opposed to the *vertical* structure of nearly all power MOSFETs) exhibit negative tempco beginning at very low gate voltage (and very low  $I_D$ ); see Figure 3.118. Lateral power MOSFETs do not attain the high breakdown voltage and low  $R_{ON}$  ratings of vertical power MOSFETs, but they are favored in linear power applications such as audio amplifiers for their linearity and thermal stability. A popular choice is the 2SK1058 (*n*-channel) and 2SJ162 (*p*-channel) complementary pair from Renesas (Hitachi), limited to 160 V and 7 A; their  $R_{ON}$  is an unimpressive  $\sim 1\ \Omega$ . This is not of great concern in the linear amplifier context, where they do not operate near voltage-saturation; but it’s high enough that you often see several of them used in parallel. See the section “Power transistors for linear amplifiers” in Chapter 3x for further discussion.

The positive tempco of  $I_D$  in power MOSFETs creates an additional problem, namely the possibility of *thermal runaway*.

<sup>115</sup> We found this cute circuit trick used in some HP (subsequently Agilent, now Keysight) E3610-series linear power supplies. It’s much simpler than using individual op-amps to bias each transistor, as some MOSFET manufacturers suggest. A different way to benefit from the higher power-dissipation capability of multiple transistors is to wire them in *series*, see for example Figure 9.111. A series connection guarantees equal current distribution.

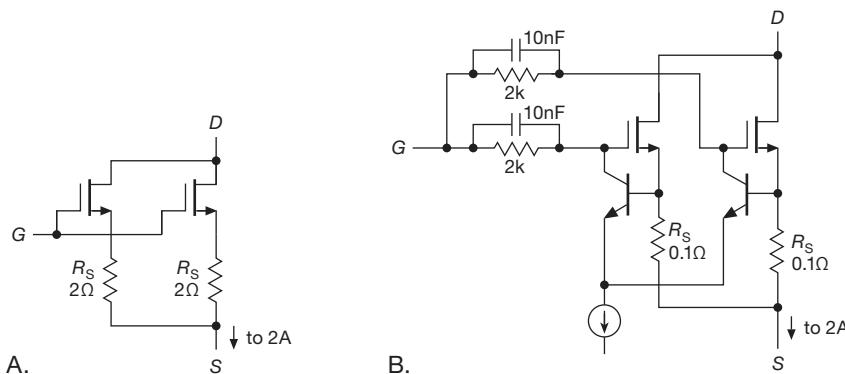


Figure 3.117. Paralleling power MOSFETs: A. with source ballasting resistors; B. with sense resistors and active feedback.

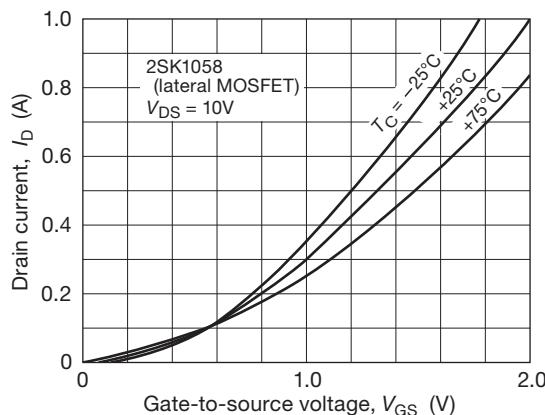


Figure 3.118. Transfer characteristics ( $I_D$  versus  $V_{GS}$ ) for the 2SK1058 *lateral n*-channel power MOSFET, popular for use in high fidelity audio power amplifiers. Here the temperature coefficient is negative over most of the operating region.

### 3.6.4 Thermal runaway

Up to now, we've avoided the R-word, because "thermal runaway" is quite independent of whether transistors are used in parallel; it refers particularly to circuit configurations in which the power dissipation produces a rise in temperature that in turn raises the power that must be dissipated. Two important examples are the push-pull linear amplifier and the saturated power switch.

#### A. Push-pull power amplifier

In the class-AB push-pull power amplifier, commonly used in audio output stages, the push-pull pair is biased with substantial quiescent current (typically  $\sim 100$  mA) to preserve linearity during waveform crossover. The quiescent current varies with temperature because both  $I_D$  (with MOSFETs) and  $I_C$  (with bipolar transistors) have positive

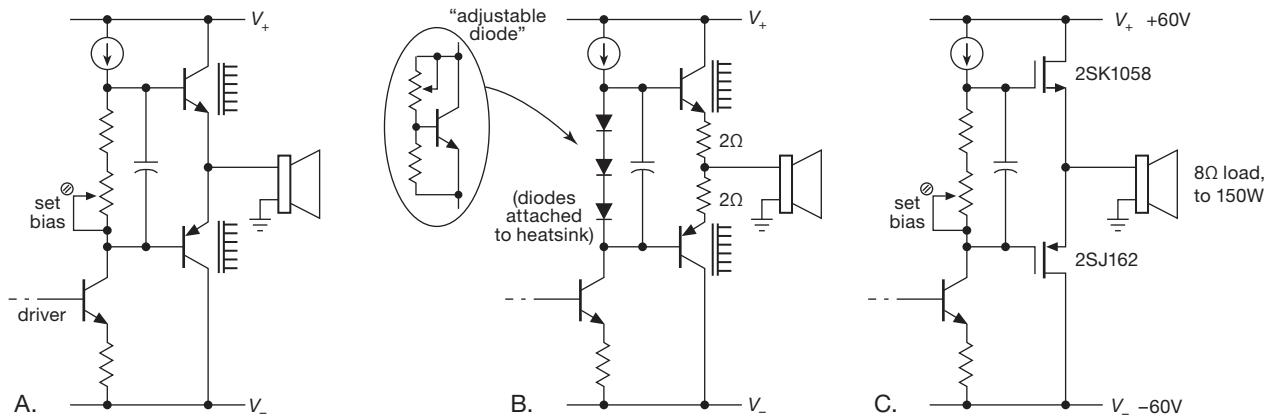
temperature coefficients at constant drive voltage. Depending on the circuit configuration and the degree of heatsinking, the output transistors may or may not reach a stable temperature; if they don't, you've got thermal runaway (independent of whether or not you've paralleled multiple transistors).

We saw this earlier in §2.4.1B, where we introduced the push-pull audio power amplifier built with complementary *bipolar* transistors. Because bipolar transistors have a positive tempco of collector current at fixed  $V_{BE}$ ,<sup>116</sup> the usual approach is to bias the bases apart with a voltage source that tracks the tempco of the output stage  $V_{BE}$ 's – typically by using diodes or transistor base-emitter junctions, thermally coupled to the output stage heatsink – often in conjunction with small emitter resistors in the output stage (Figure 3.119B).

Power MOSFETs in linear push-pull amplifiers present the same problem, because they are operated in the region of positive tempco of  $I_D$  (§3.6.3B). You can use the same trick (bias generator with tracking negative tempco, perhaps in combination with small-value output-stage source resistors; see the section "Power transistors for linear amplifiers" in Chapter 3x). However, the problem is nicely finessed by using lateral power MOSFETs, whose negative tempco of  $I_D$  (Figure 3.118) beginning at  $I_D \approx 100$  mA guarantees no thermal runaway. The usual approach is to bias the output stage gates apart with a (settable) constant dc voltage, as shown in Figure 3.119, bypassed at signal frequencies.<sup>117</sup> The bias is typically set for a quiescent current  $I_Q$  close to the zero-tempco crossover (100 mA for

<sup>116</sup> Or, alternatively, a negative tempco of  $V_{BE}$  for constant  $I_C$ .

<sup>117</sup> The figure shows bare-bones circuitry. In practice the bipolar transistors would be configured as Darlington or Sziklai pairs, and the single-ended driver stage might be replaced with a symmetrical pair of drivers, driven from the differential input stage. For a 150 W amplifier you would probably use paralleled pairs of transistors to stay within



**Figure 3.119.** Thermal stability in push–pull power amplifiers – simplified output stage configurations. A. Fixed  $V_{BE}$  biasing promotes runaway, owing to positive  $I_C$  tempco in bipolar output stage. B. Tracking thermally-coupled bias generator tames runaway. C. Stable quiescent current in lateral MOSFETs biased at fixed  $V_{GS}$ ; no thermal compensation is needed.

the 2SK1058/2SJ162 complementary pair), ensuring that  $I_Q$  remains relatively constant as the amplifier warms up.<sup>118</sup>

### B. Saturated switch

It's widely believed that MOSFETs are immune to thermal runaway when used for power *switching*. The thought process goes like this: "These puppies have really low  $R_{ON}$  when driven to full conduction, so they hardly need any heatsinking; besides, if they do heat up (while carrying some large but bounded current), the thing will stabilize at some elevated temperature, because the power carried off by the heatsink increases roughly proportional to the rise above ambient temperature, and eventually catches up to the power being dissipated; plus, hey, these things are *tough!*"

Nice thoughts. But the reality can be different. That's because  $R_{ON}$  isn't constant, but increases with temperature (Figure 3.116); so the switch dissipates more power as it heats up, and, if the heatsink is too small, the heat it carries off may never catch up – in which case the process runs away!

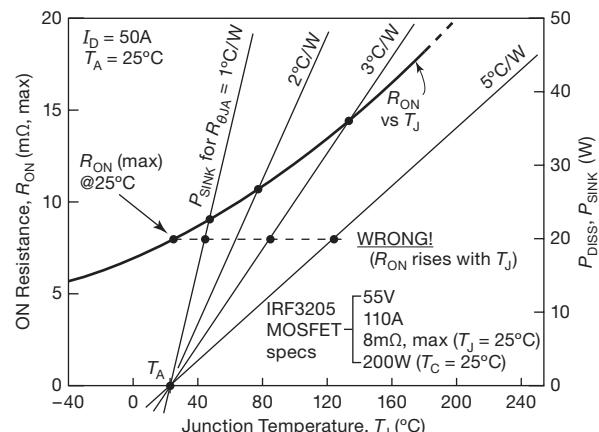
To lend some perspective: you don't need to have actual thermal *runaway* to cause overheating and destruction – an undersized heatsink<sup>119</sup> will do the job just fine by al-

allowable junction temperature; no ballasting resistors would be needed for the MOSFET version.

<sup>118</sup> Because lateral MOSFETs may be hard to get, a regular power MOSFET can be used in place of the BJT "V<sub>BE</sub> diode" (as in Figure 3.119B) to bias a complementary pair of ordinary power MOSFETs. This approach prevents thermal runaway because the tempco of a MOSFET is higher at low currents than at high currents, see Figure 3.115.

<sup>119</sup> Or none at all! The impressive  $I_{D(\max)}$  ratings on the datasheet could

low the junction temperature to soar above  $T_{J(\max)}$ . And, as we'll see shortly, the better approach is to reduce power dissipation by reducing  $R_{ON}$ , rather than by piling on larger heatsinks. With this cautionary comment, let's see how actual thermal runaway could occur in an ill-advised design.



**Figure 3.120.** Thermal runaway in a MOSFET switch. The curved line plots maximum ON-resistance and corresponding power dissipation at 50 A for an IRF3205 n-channel power MOSFET. The straight lines plot power removed by three choices of heatsink thermal resistance  $R_{SINK}$ . Thermal runaway occurs with the smallest heatsink, where there is no graphical intersection.

Figure 3.120 shows an easy graphical way to see what's happening and to figure out how much heatsink you need to

tempt you to omit the heatsink entirely, even in a power-switching circuit operating at substantial drain currents.

prevent runaway (and, by the way, to keep the junction temperature  $T_J$  below the specified  $T_{J(\max)}$ ). We begin by plotting the datasheet's  $R_{ON}$  versus temperature for an inexpensive power MOSFET (their graph tops out at 175°C, because that is the maximum rated junction temperature; we took the liberty of extending the graph another 75°). Then we use that to evaluate the power dissipation, as  $P_{diss} = I^2 R_{ON}$ ; for our chosen 50 A drain current we got the values marked on the right-hand axis. Finally, we plot separately the power carried off by four values of heatsink "thermal resistance"  $R_{\Theta JA}$  (given by  $P_{diss} = (T_J - T_A)/R_{\Theta JA}$ ), assuming an ambient temperature of  $T_A = 25^\circ\text{C}$  (plenty of discussion of this in §9.4.1A).

The heatsinks carry off an amount of power proportional to the temperature rise above ambient, as plotted; the transistor generates power according to its graph. The intersection (if any!) is the equilibrium temperature, which in this case is about 45°C or 75°C, for the two larger heatsinks. But the smallest heatsink has no intersection – it cannot carry off as much heat as the transistor generates, at any temperature: *thermal runaway!* In real life you should assume that the ambient temperature will be higher (equipment is put into racks, or stacked with other equipment; and, you can have hot weather!): you would do this by sliding the heatsink curves to the right.

From this simple example you might conclude that larger heatsinks are the right cure for thermal runaway in saturated switching applications. But look again at the numbers: we're switching 50 A with an  $R_{ON}$  of order 10–15 mΩ – that's a 0.5 to 0.75 V drop, and 25 to 40 W dissipation. At these sorts of currents we really should be using a larger transistor, or several in parallel, to reduce  $R_{ON}$  (and therefore the dissipated power). The "maximum  $I_D$ " specification (here 110 A) looks pretty good on a datasheet, but is not a realistic guide to appropriate dc operation of the part. In this example a better choice would be a low  $R_{ON}$  part like the FDB8832<sup>120</sup> with  $R_{ON}=2.3\text{ m}\Omega$  (max) at 25°C, and with a typical ON voltage of 115 mV at 50 A and power dissipation of 5.8 W.<sup>121</sup> This is a 30 V part (high-voltage MOSFETs have higher  $R_{ON}$ ); if you wanted to switch somewhat higher voltages with low  $R_{ON}$  and  $P_D$ , your choices are to use a high-power MOSFET module<sup>122</sup> or (less expensive) several conventional MOSFETs in parallel. For voltages above 400 V or so the transistor of choice is the IGBT (insulated-gate bipolar transistor; see

§3.5.7A), which has the input properties of a MOSFET and the output properties of a BJT. An example is the Mitsubishi CM1200HC-50H, rated at 2500 V and 1200 A: at full current it saturates at just 3 V (equivalent to a 2.5 mΩ ON-resistance). That's pretty good... but that works out to 3.6 kilowatts dissipation! (These things are used for power switching in applications like electric locomotives.)

Lest we leave the wrong impression, we hasten to point out that thermal calculations and heatsink selection need not require the kind of graphical plotting we've done here (in which we were interested primarily in the possibility of true thermal runaway). More simply, you can just apply a safety factor  $m$  to the datasheet's 25°C  $R_{ON}$  value to get a reasonable estimate of  $R_{ON}$  at the maximum junction temperature (150°C); from that you get

$$T_J \approx T_A + I_D^2 \cdot m R_{ON(25C)} \cdot R_{\Theta JA}. \quad (3.15)$$

The multiplier  $m$  varies somewhat with the voltage rating of the MOSFET; based on data from many datasheets (see the graph and discussion in the section "MOSFET ON-resistance versus temperature" in Chapter 3x), it ranges from roughly  $m \approx 1.5$  (for low-voltage MOSFETs) to roughly  $m \approx 2.5$  (for high-voltage MOSFETs). As a practical rule-of-thumb, you'll be safe if you use  $m=2$  for MOSFETs rated to 100 V and  $m=2.5$  for those of higher voltages (at least to 1kV).

### C. Second breakdown and safe operating area

It's worth emphasizing a related thermal effect ("second breakdown") that was discussed earlier in §3.5.1B: power transistors fail (usually<sup>123</sup>) if operated beyond their maximum voltage, their maximum current, or their maximum junction temperature (the latter dependent on power dissipation, pulse duration, heatsink thermal resistance, and ambient temperature; see §9.4.2). The boundaries define the *safe operating area*, or SOA, for example as shown in Figure 3.95. Bipolar transistors suffer from an additional failure mode known as *second breakdown*, an incompletely understood instability characterized by local heating, reduction of the breakdown voltage, and, often, destruction of the junction. It is second breakdown that imposes the additional constraint on the bipolar SOAs in Figure 3.95. Happily, MOSFETs are less likely to suffer from second breakdown, which contributes to their popularity in power circuits.<sup>124</sup> Note that for both kinds of transistors the maximum current and power limits are higher for short pulses.

<sup>120</sup> From the same manufacturer, Fairchild Semiconductor.

<sup>121</sup> Rising to 3.6 mΩ, 180 mV, and 9 W (maximum) at  $T_J=150^\circ\text{C}$ .

<sup>122</sup> These come in husky "SOT-227" packages, with screw terminals on the top, an isolated metal base, and with names like "ISOTOP" and "miniBLOC."

<sup>123</sup> Or, perhaps more accurately, they are not guaranteed *not* to fail!

<sup>124</sup> Some newer fine-geometry types *are* susceptible, however; see IR App Note AN-1155.

**Table 3.7 Junction Field-Effect Transistors (JFETs)<sup>a</sup>**

Part # <sup>b</sup>	N or P-channel Jellybean?	TO-92: 2N, PN SOT23: MMBF SOT23: PMBF MMBF_LT	Prefixes & Pkgs <sup>d</sup>																
			V <sub>DSS</sub> max (V)	I <sub>DSS</sub> min (mA)	I <sub>DSS</sub> max (mA)	I <sub>DSS</sub> meas (mA)	R <sub>ON</sub> max (Ω)	V <sub>GSS(off)</sub> <sup>c</sup> min (V)	V <sub>GSS(off)</sub> <sup>c</sup> max (V)	V <sub>GS @ I<sub>D</sub></sub> measured (V)	V <sub>GS @ I<sub>D</sub></sub> min (mA)	V <sub>GS @ I<sub>D</sub></sub> max (mA)	g <sub>m</sub> @ I <sub>D</sub> meas (mS)	g <sub>m</sub> @ I <sub>D</sub> meas (mA)	G <sub>max</sub> <sup>e</sup> typ (pF)	C <sub>iss</sub> typ (pF)	C <sub>rss</sub> typ (pF)		
PN4117	N	• A C - -	40	0.03	0.09	0.07	-	-0.6	-1.8	-0.33	0.03	0.07	0.21	z	0.09	0.03	420 <sup>r</sup>	1.2	0.3
'4118	N	- A C - -	40	0.08	0.24	0.20	-	-1	-3	-1.33	0.1	0.08	0.25	z	0.13	0.1	260 <sup>r</sup>	1.2	0.3
'4119	N	- A C - -	40	0.20	0.60	0.30	-	-2	-6	0.0	0.3	0.10	0.33	z	0.18	0.3	140 <sup>r</sup>	1.2	0.3
BFT46	N	- - C - -	25	0.20	1.5	0.63	-	-	-1.2	-0.16	0.3	1	-	z	1.7	0.3	190 <sup>s</sup>	3.5	0.8
BF511	N	- - D - -	20	0.7	3	4.2	-	-	-1.5	-0.75	1	4	-	z	2.7	1	120	-	0.3
2N5457	N	• A C - -	25	1	5	3.5	-	-0.5	-6	-0.81	1	1	5	z	2.3	1	220	4.5	1.5
'5458	N	- A C - -	25	2	9	4.1	-	-1	-7	-0.97	1	1.5	5.5	z	2.2	1	190	4.5	1.5
'5459	N	- A C - -	25	4	16	9.9	-	-2	-8	-1.82	3	2	6	z	2.9	3	100	4.5	1.5
2N5460	P	• A C - -	25	-1	-5	3.4	-	0.75	6	+0.97	1	1	4	z	1.9	1	260	4.5	1.2
'5461	P	• A C - -	25	-2	-9	2.7	-	1	7.5	+0.67	1	1.5	5	z	2	1	210	4.5	1.2
'5462	P	• A C - -	25	-4	-16	5.9	-	1.8	9	+4.15	1	2	6	z	2.5	3	30	4.5	1.2
MMBF4416	N	• A C C C	30	5	15	5.9	-	-	-6	-0.19	5	4.5	7.5	z	3.9	5	70	4	0.8
2N5484	N	- A C - C	25	1	5	3.3	-	-0.3	-3	-0.73	1	3	6	z	2.3	1	230	10	2.2
'5485	N	- A C - -	25	4	10	6.6	-	-0.5	-4	-1.65	1	3.5	7	z	2.1	1	150	10	2.2
'5486	N	- A C - -	25	8	20	14	-	-2	-6	-2.61	1	4	8	z	2.1	1	75	10	2.2
2SK170BL	N	- B - - -	40	6	12	6.1	-	-0.2	-1.5	-0.04	5	22 <sup>t</sup>	z	29	5	470	30	6	
LSK170B	N	- B C - -	40	6	12	7.6	-	-0.2	-2	-0.17	3	10 <sup>t</sup>	1	20	3	160	20	5	
LSK170C	N	- B C - -	40	10	20	13	-	-0.2	-2	-0.26	5	10 <sup>t</sup>	1	24	5	90	20	5	
BF861B	N	- - C - -	25	6	15	8	-	-0.5	-1.5	-0.47	1	16	25	z	16	5	150	7.5	
BF545C	N	- - C - -	30	12	25	19	-	-3.2	-7.8	-1.80	5	3.0	6.5	z	3.7	5	30	1.7	0.8
BF862	N	- - C - -	20	10	25	12	-	-0.3	-1.2	-0.21	5	35	45 <sup>t</sup>	z	26	5	270	10	1.9
PF5103	N	- A C - -	40	10	40	19	30	-1.2	-2.7	-1.00	5	7.5	-	2	10	5	160	16	6
PN4391	N	• A C C C	40	50	150	115	30	-4	-10	-7.15	5	12 <sup>t</sup>	5	8.8	5	30	12	3.5	
'4392	N	• A C C C	40	25	75	38	60	-2	-5	-1.67	5	16 <sup>t</sup>	10	10	5	130	12	3.5	
'4393	N	• A C C C	40	5	30	16	100	-0.5	-3	-1.25	1	13 <sup>t</sup>	10	6.2	1	150	12	3.5	
J105	N	- A C - -	25	500	-	-	3	-4.5	-10	-8.39	5	40 <sup>t</sup>	5	37	10	60	160 <sup>m</sup>	35 <sup>m</sup>	
J106	N	- A C - -	25	200	-	-	6	-2	-6	-2.42	5	53 <sup>t</sup>	5	43	10	230	160 <sup>m</sup>	35 <sup>m</sup>	
J107	N	- A C - -	25	100	-	-	8	-0.5	-4.5	-1.93	5	75 <sup>t</sup>	5	48	10	340	160 <sup>m</sup>	35 <sup>m</sup>	
J108	N	• A C C -	25	80	-	325	8	-3	-10	-5.83	5	37 <sup>t</sup>	5	31	10	60	85	15	
J109	N	• A C C -	25	40	-	201	12	-2	-6	-2.85	5	26 <sup>t</sup>	5	32	10	160	85	15	
J110	N	• A C C -	25	10	-	122	18	-0.5	-4	-1.80	5	20 <sup>t</sup>	5	34	10	220	85	15	
J111	N	- A C C -	35	20	-	115	30	-3	-10	-7.6	5	-	-	8.4	5	30	28	5	
J112	N	• A C C -	35	5	-	47	50	-1	-5	-2.8	5	6.7 <sup>t</sup>	1	9.5	5	100	28	5	
J113	N	- A C C -	35	2	-	21	100	-0.5	-3	-1.0	5	8 <sup>t</sup>	1	11	5	100	28	5	
J174	P	- B C C -	30	-20	-135	26	85	-5	-10	+2.08	5	4.5	-	5	-	-	15	13	6
J175	P	• B C C C	30	-7	-60	13	125	-3	-6	+1.58	1	-	-	-	-	30	13	6	
J176	P	- B C C C	30	-2	-25	6.1	250	-1	-4	+0.86	1	6.3	-	5	-	-	40	13	6
J177	P	• B C C C	30	-1.5	-20	4.2	300	-0.8	-2.5	+0.62	1	-	-	-	-	50	13	6	
J308	N	• A C C -	25	12	60	35	-	-1	-6.5	-	8	-	10	12	5	120	4	2	
J309	N	• A C C C	25	12	30	23	-	-1	-4	-1.2	5	10	20	10	11	5	300	4	2
J310	N	- A C C C C	25	24	60	39	-	-2	-6.5	-2.4	5	8	18	10	8.9	5	100	4	2
<i>dual JFETs</i>																			
LS840-42	N	F	60	0.5	5	3.3	-	-1	-4.5	-0.85	1	0.5	1	0.2	2.1	1	180	4	1.2
'843-5	N	F	60	1.5	15	-	-	-1	-3.5	-	-	1	1.5 <sup>t</sup>	0.5	-	-	8 <sup>m</sup>	3 <sup>m</sup>	
LSK389A	N	F, J	40	2.6	6.5	-	-	-0.15	-2	-	-	8	20 <sup>t</sup>	3	-	-	25	5.5	
'389B	N	F, J	40	6	12	12	-	-0.15	-2	-0.24	5	8	20 <sup>t</sup>	3	23	5	170	25	5.5
'389C	N	F, J	40	10	20	-	-	-0.15	-2	-	-	8	20 <sup>t</sup>	3	-	-	25	5.5	
LS5912	N	F, J, K <sup>p</sup>	25	7	40	18	-	-1	-5	-1.75	5	4	10	5	5.7	5	70	5	1.2
IFN146	N	F <sup>v</sup>	40	-	30	6	-	-0.3	-1.2	-0.19	1	30	40 <sup>t</sup>	z	25	5	660	75 <sup>m</sup>	15 <sup>m</sup>

(a) listed generally by increasing I<sub>DSS</sub>, but also by part number within a family (e.g., J105–J113); see also Table 8.2 for noise parameters.

(b) for families of related parts, **boldface** designates the family matriarch. (c) usually specified at I<sub>D</sub>=1nA or 10nA, though sometimes at 10μA or even 200μA (e.g., for the J105–J113 “switches”); it doesn’t much matter, given the wide range of specified V<sub>GSS(off)</sub>. (d) see the accompanying pinout figure; all JFETs appear to be symmetric (source and drain are interchangeable), but *italic* designates a datasheet pinout in which the S and D terminals are interchanged. (e) G<sub>max</sub>=g<sub>m</sub>/g<sub>DS</sub>, the maximum grounded-source voltage gain into a current source as drain load; listed values measured at I<sub>D</sub>=1mA and V<sub>DS</sub>=5V, unless noted otherwise. G<sub>max</sub> is proportional to V<sub>DS</sub>, and for most JFETs G<sub>max</sub> is relatively constant over varying I<sub>D</sub>. Use tabulated G<sub>max</sub> to find g<sub>DS</sub>=g<sub>m</sub>/G<sub>max</sub>. (m) maximum. (p) several PDIP-8 pkgs available. (r) at I<sub>D</sub>=30μA. (s) at I<sub>D</sub>=300μA. (t) typical. (v) variant of “F” pinout: G and D terminals interchanged. (z) at I<sub>DSS</sub>.

Table 3.8 Low-side MOSFET Gate Drivers<sup>a</sup>

Part #	Mfg <sup>d</sup>	# channels	Speed <sup>b</sup>						logic thresh <sup>p</sup>	source below gnd? current limit	UVLO? enable?	inv?	non-inv?	output rail-to-rail?	Packages			Comments	
			V <sub>min</sub> (V)	V <sub>max</sub> (V)	I <sub>pk</sub> (A)	t <sub>d</sub> + 0.5t <sub>r</sub> (ns, typ)	C <sub>load</sub> (nF)								TO220, Dpak	DIP	SOIC, MSOP	SOT23	
TC4426-28	MC+	2	4.5	18	1.5	55	1	T	-	-	-	n	n	•	-	•	•	-	G,H
TC4423-25	MC+	2	4.5	18	3	70	1.8	T	-	-	-	n	n	•	-	•	•	-	G,H
TC4420,29	MC+	2	4.5	18	6	80	2.5	T	-	-	-	•	•	•	-	•	•	-	G
TC4421-22	MC+	1	4.5	18	9	85	10	T	-	-	-	•	•	•	-	•	•	-	G,J
FAN3111	F	1	4.5	18	1	20	0.5	C	-	-	-	c	c	c	•	-	-	-	-
FAN3100C,T	F	1	4.5	18	2	20	1	C,T	-	-	•	c	c	c	•	-	-	-	•
FAN3180	F	1	5	18	2	30	1	T	-	-	•	-	•	•	-	-	-	-	B
FAN3216-17	F	2	4.5	18	2	25	2.2	T	-	-	•	-	•	•	-	-	-	-	D
FAN3226-29C,T	F	2	4.5	18	2	25	1	C,T	-	-	•	c	c	c	•	-	-	•	C,E
FAN3213-14	F	2	4.5	18	4	20	2.2	T	-	-	•	-	•	•	-	-	-	-	C
FAN3223-25C,T	F	2	4.5	18	4	25	2.2	C,T	-	-	•	c	c	c	•	-	-	-	E
FAN3121-22	F	1	4.5	18	9	21	10	T	-	-	•	•	•	•	-	-	-	-	-
IRS44273L	IR	1	12	20	1.5	50	1	T	-	-	•	-	-	•	-	-	-	-	-
IR25600	IR	2	6	20	1.5	75	1	T	-	-	-	-	-	•	-	-	•	-	-
MAX17600-05	MA	2	4	14	4	15	1	C5,T	-	-	•	n	n	•	-	-	•	-	H
MAX5054-07	MA	2	4	15	4	38	5	C,T	-	-	-	c	c	c	•	-	-	•	-
MAX5048A,B	MA	1	4	12.6	7.6 <sup>h</sup>	18	1	C,T	-	-	•	c	c	c	•	-	-	-	-
UCC37323-25 <sup>k</sup>	TI	2	4.5	15	4	47	1.8	T	-	-	-	•	•	•	-	-	•	-	-
UCC27517	TI	1	4.7	20	4	17	1.8	T	-	-	•	c	c	c	-	-	-	-	-
UCC27516-19	TI	1	4.7	20	4	17,21	1.8	T,C	-	-	•	•	•	•	-	-	-	-	-
UCC27523-26	TI	2	4.7	20	5	17	1.8	T	-	-	•	•	•	•	-	-	•	-	E,H
UCC37321-22 <sup>k</sup>	TI	1	4	15	9	50	10	T	-	-	-	•	•	•	-	-	•	-	-
MIC44F18-20	MI	1	4.5	13.2	6	24	1	T	-	-	•	•	•	•	-	-	•	-	-
ADP3623-25	A	2	4.5	18	4	28	2.2	T	-	-	•	-	•	•	-	-	•	-	H,P
LM5110	TI	2	3.5	14	5 <sup>f</sup>	38	2	T	•	-	•	-	n	n	•	-	•	-	H,L
LM5112	TI	2	3.5	14	79	38	2	T	•	-	•	-	n	n	-	-	•	-	H,L,M
LM5114	TI	1	4	12.6	7.6 <sup>h</sup>	16	1	C	-	-	•	c	c	c	•	-	-	•	-
ISL89367	IN	2	4.5	16	6	45	10	F	•	-	•	n	o	o	•	-	•	-	N
ISL89160-62	IN	2	4.5	16	6	45	10	C5,T	-	-	•	-	•	•	-	-	•	-	O
MC34151	O	2	6.5	18	1.5	50	1	T	-	-	•	-	-	•	-	-	•	-	-
IR2121	IR	1	12	18	2 <sup>e</sup>	200	3.3	T	-	•	-	-	•	-	-	•	-	-	F
UC3708	TI	2	5	35	3	37	1	T	-	-	-	•	-	•	-	-	•	-	-
IXDD602	IX	1	4.5	35	2	50	1	C5	-	-	•	•	•	•	-	•	•	-	H,R
IXDD604	IX	1	4.5	35	4	40	1	C5	-	-	•	•	•	•	-	•	•	-	H,R
IXDD609	IX	1	4.5	35	9	60	10	C5	-	-	•	•	•	•	-	•	•	-	R
IXDD614	IX	1	4.5	35	14	70	15	C5	-	-	•	•	•	•	-	•	•	-	R
IXDD630	IX	1	10	35	30	65	5.6	C5	-	-	•	•	•	•	•	-	-	-	K,R
ZXGD3002-04	D	1	-	20,40	9.5	11	1	-	-	-	-	-	-	-	-	-	•	-	M,S

Notes: (a) sorted by family, within family sorted by lout; except for ZXGD3000-series, all devices swing rail-to-rail, or nearly so. (b) into Cload at V<sub>S</sub>=12V. (c) input gate with inv and non-inv inputs. (d) A=Analog Devices; D=Diodes, Inc; F=Fairchild; IN=Intersil; IR=International Rectifier; IX=Ixys/Clare; L=LTC; MA=Maxim; MC=Microchip; MI=Micrel; O=On Semiconductor; S=STMicroelectronics; TI=Texas Instruments. (e) 1A source, 2A sink. (f) 3A source, 5A sink. (g) 3A source, 7A sink. (h) 1.3A source, 7.6A sink. (k) 37xxx for 0 to 70°C, 27xxx for -40°C to 105°C. (n) see part-specific comments. (o) XOR input sets optional invert. (p) C=CMOS; C5=5V CMOS; F=flexible, set by V<sub>ref-</sub> and V<sub>ref+</sub> input pins; T=TTL.

Comments: (A) suffix specifies logic threshold. (B) includes 3.3V LDO output. (C) 2ns td channel match. (D) 1ns td channel match. (E) dual inv+en, dual non-inv+en, dual inputs. (F) source-resistor current-sense input terminal, suitable for driving an IGBT. (G) industry std, many mfgs. (H) dual inv, dual non-inv, or one each. (J) for 8-pin pkgs, n- and p-ch drains on separate pins. (K) t<sub>r</sub>, t<sub>f</sub> = 50ns into 68nF. (L) output swing to neg rail, can be 5V below logic gnd. (M) n- and p-ch drains on separate pins. (N) resistor-programmed edge-delay timers; 2-input AND signal inputs. (O) ISL89163-65 same, but include enable inputs; ISL89166-68 same, but include resistor-programmed edge-delay timer inputs. (P) overtemp protection and output. (R) full p/n is IXDx6..., where x = N, I, D and F for non-inv, inv, dual non-inv+en, or one of each. (S) series is one each high-current high-gain npn and pnp transistor emitter-followers for pullup and down.

## Review of Chapter 3

An A-to-Z summary of what we have learned in Chapter 3. This summary reviews basic principles and facts in Chapter 3, but it does not cover application circuit diagrams and practical engineering advice presented there.

### ¶A. FETs.

In Chapter 3 we explored the world of Field-Effect Transistors, or FETs. FETs have a conducting channel with terminals named *Drain* and *Source*. Conduction in the channel is controlled by an electric field created by a third *Gate* electrode (§3.1). As with bipolar transistors (BJTs), FETs are transconductance devices (see ¶G below), which means the drain *current* (assuming sufficient drain-to-source voltage) is controlled by the gate *voltage*.

### ¶B. *n*-channel and *p*-channel.

Like BJTs with their *npn* and *pnp* types, FETs come in both *n*- and *p*-channel polarities (§3.1.2). In either case the channel conductance increases if the gate voltage is taken toward the drain voltage. For example, for an *n*-channel FET with a positive drain voltage, the channel can be turned on with a sufficient positive-going voltage, and cutoff with a sufficient negative-going voltage. That's not to say the *n*-channel device requires positive and negative voltages to turn on and off. A threshold voltage  $V_{th}$  can be defined where the FET is just slightly turned on, and the channel responds to gate voltages above and below  $V_{th}$  for control.

### ¶C. Enhancement and Depletion Modes.

See Figure 3.8. Enhancement-mode devices have a high enough threshold voltage  $V_{th}$  that they are nonconducting (i.e., off) when their gate voltage is at  $V_{GS}=0$  V. To bring such a FET into conduction, the gate of is brought positive (if *n*-channel) or negative (if *p*-channel). Depletion-mode devices, by contrast, have their threshold voltage well into the “off” direction, thus they are conducting (i.e., on) with their gate-voltage at  $V_{GS}=0$  V. Thus for example you must apply a considerable negative gate voltage  $V_{GS}$  to turn off an *n*-channel depletion-mode FET. See Figure 3.9 where drain current versus gate voltage is shown for a selection of *n*-channel devices. FETs can be fabricated with the transfer curve shifted left or right (more about this in ¶H below). Figures 3.10 and 3.11 show convenient maps of the FET types.

### ¶D. MOSFETs and JFETs.

In metal-oxide FETs (MOSFETs) the gate electrode is fully insulated from the channel, and can be taken positive or

negative, typically up to  $\pm 20$  V. In junction FETs (JFETs) the semiconductor gate contacts the channel and acts as a diode junction, so it is insulated only in the reverse direction. Therefore JFETs are necessarily depletion mode devices; one cannot make an enhancement-mode JFET. Figures 3.6 and 3.7 show FET symbols.

### ¶E. FET Characteristics, Gate and Drain.

See Figure 3.13. A FET’s channel conductance and current is controlled primarily by its gate voltage, but it’s also affected by the drain voltage  $V_{DS}$ . At very low drain voltages the channel acts like a resistor, whose value is controlled by the gate (§3.1.2 and §3.2.7); this is called the *linear* region. At higher drain voltages the drain current levels off, being controlled by the gate voltage and only weakly dependent upon drain voltage; this is called the *saturated* region. In the saturated region the FET drain acts like a current source (or sink), and the device is characterized by its transconductance  $g_m$  (see ¶G below). MOSFETs are often used as switches. In this mode of operation a large gate voltage (e.g., 10 V) is applied to make the channel resistance low enough to approximate a closed switch. More on FET switches in sections ¶¶O–Q below.

### ¶F. Square-law.

Over a large region of gate voltages greater than  $V_{th}$ , and for drain voltages above a volt or so (i.e., in the saturated region), a FET’s drain current behaves like a square-law device; that is, its drain current is proportional to the square of the excess gate-drive voltage ( $V_{GS}-V_{th}$ )<sup>2</sup>, see Figure 3.14 and eq’n 3.2. This is sometimes called the *quadratic* region. The threshold voltage  $V_{th}$  is generally determined with an extrapolated  $\sqrt{I_D}$  plot, as the figure shows. For  $V_{GS}$  below threshold the FET is in the subthreshold region; see ¶I below.

### ¶G. Transconductance and Amplifiers.

Transconductance  $g_m$  is the change in output drain current caused by a change in gate voltage:  $g_m=i_D/v_{GS}$  (the lower-case *i* and *v* signify small signals). Common-source FET amplifiers (§3.2.3, Figures 3.28 and 3.29) have voltage gain  $G=-g_m R_D$ , where  $R_D$  is the drain load resistance. In contrast to BJTs (where  $g_m \propto I_C$ ), the transconductance of FETs rises only as  $\sqrt{I_D}$  in the important quadratic region; see Figures 3.53 and 3.54. As a consequence FET amplifiers with resistive drain loads have lower gain when designed to operate at higher current, because  $R_D$  is generally chosen inversely proportional to drain current. The FET’s internal output resistance also acts as a load resistance, thus

limiting gain (“ $G_{\max}$ ”) even with an ideal current-source drain load; see §3.3.2 eq’n 3.13, and Table 3.1.

When used as a *follower*, an FET has an output impedance  $r_{\text{out}}=1/g_m$ , see ¶K below.

#### ¶H. Biasing JFET Amplifiers.

JFETs are well suited for making signal amplifiers (by contrast there are few viable small discrete MOSFETs), and they work especially well in low-noise amplifiers. But there’s one very painful issue analog designers face: the uncertain value of the gate operating voltage for any given part. Scanning the min and max columns for  $V_{GS(\text{off})}$  in the JFET Table 3.1 on page 141, we see values for a particular JFET that range from  $-1\text{ V}$  to  $-7\text{ V}$ , or  $-0.4\text{ V}$  to  $-4\text{ V}$ . The latter is a 10:1 ratio! Figure 3.17 shows  $V_{GS}$  histograms for 300 parts, 100 each for three different JFET types in a family. Here we see gate voltage spreads of about 1 V, which you might rely upon if you buy a batch of parts from one manufacturer and measure them. But, *caution*: Figures 3.51 and 3.52 show how the same part type may vary when purchased from different manufacturers. To deal with the uncertainty, special biasing schemes are often required in FET amplifier circuits. Figures 3.25 and 3.41 show examples of the load-line concept for analyzing amplifier biasing.

#### ¶I. Subthreshold Region.

The simple FET formula of eq’n 3.2 predicts zero drain current when the gate voltage reaches threshold ( $V_{GS}=V_{\text{th}}$ ). In reality the drain current is not zero, and transitions smoothly to a subthreshold region (see Figure 3.16) where the FET looks more like a BJT, with its exponential Ebers-Moll characteristic (§2.3.1). In this region (where  $I_D$  rises exponentially with  $V_{GS}$ ) we’re glad to see a higher  $g_m \propto I_D$ ; but sadly the FET proportionality constant is usually  $2\times$  to  $5\times$  smaller than for BJTs, see Figure 3.53.

#### ¶J. Self-biased Amplifiers.

Depletion-mode MOSFETs (and all JFETs) operate with a reverse voltage on their gates, which allows them to be self-biased (§3.2.6A). The source terminal is “higher” than the gate terminal, so a source resistor connected between them sets the drain current to  $I_D=V_{GS}/R$ . This is also a convenient way to make a 2-terminal current source, but the tolerance will be poor due to the wide variability in  $V_{GS}$ , see ¶H. Alternately the  $V_{GS}$  voltage available at the source pin may be used to operate a current-setting IC like the LM334.

#### ¶K. Source Followers.

Source followers (§3.2.6), Figure 3.40 have a nominal gain of 1, analogous to the BJT emitter follower. Because of their lower  $g_m$ , however, they have considerably higher output resistance,  $r_{\text{out}}=1/g_m$ , so the ideal unity gain is reduced by load resistance, see eq’n 3.7.

#### ¶L. FETs as Variable Resistors.

At low drain voltages ( $V_{DS} \ll V_{GS}$ ) FETs act as variable resistors programmed by the gate voltage. Because the slope varies with  $V_{DS}$ , however, the resistance is somewhat non-linear. But there’s a simple trick to linearize this resistance, by exploiting the quadratic behavior of FETs, see Figures 3.46 and 3.47.

#### ¶M. FET Gate Current.

The gate of a JFET forms a diode junction with the channel; it’s normally reverse-biased, with some non-zero dc leakage current (§3.2.8). This current roughly doubles for every  $10^\circ\text{C}$  temperature increase; furthermore it increases dramatically at high drain currents and drain voltages due to impact ionization, see Figure 3.49. MOSFET gates do not suffer from either of these leakage-current-increasing effects. In contrast to the generally negligible dc gate leakage, the input capacitance  $C_{\text{iss}}$  of FETs (which can be quite high, many hundreds of pF for large power MOSFETs) often presents a substantial ac load. Use a gate-driver chip (Table 3.8) to provide the high transient currents needed for rapid switching.

#### ¶N. JFET Switches.

JFETs can be used as analog-signal switches, as in the *n*-channel switch of Figure 3.62. The switch is OFF when the gate is taken at least  $V_{\text{th}}$  below than the most negative input signal. To turn the switch ON the gate voltage must be allowed to equal the source. JFETs are symmetrical, so e.g., for an n-channel part, the “source” would be the most negative pin. Large-die JFETs work well as power switches up to 100 mA; Table 3.1 lists parts with  $R_{\text{ON}}$  as low as  $3\Omega$ .

#### ¶O. CMOS Switches.

CMOS signal switches are made with a parallel pair of complementary *n*- and *p*-channel MOSFETs. This reduces  $R_{\text{ON}}$  as shown in Figure 3.61, and beneficially causes cancellation of most of the injected charge transfer (§3.4.2E), see Figure 3.79. The injected charge scales roughly inversely proportional to  $R_{\text{ON}}$  (Figure 3.81), so there’s a tradeoff between desirably low on-resistance and desirably low self-capacitance. As an example, Table 3.3 lists a switch with an impressive  $R_{\text{ON}}=0.3\Omega$  – but it’s burdened

with a whopping 300 pF of self-capacitance. A T-switch configuration can be used to reduce the signal feedthrough at high-frequencies, see Figure 3.77.

#### ¶P. CMOS Logic Gates.

See Figure 3.90. A series pair of complementary (*n*- and *p*-channel) small-geometry MOSFETs between the positive rail and ground forms the simplest logic inverter (Figure 3.90); more switches can be arranged to make CMOS logic gates (e.g., Figure 3.91, §3.4.4), with the attractive property of nearly zero static power, except when switching. CMOS logic is covered extensively in Chapters 10 and 12, and is the basis for all contemporary digital processors.

#### ¶Q. MOSFET Power Switches.

Most power MOSFETs (§3.5) are enhancement type, available in both *n*- and *p*-channel polarities. They are very popular for use as high-current high-voltage power switches. A few relevant parameters are the breakdown voltage  $V_{DSS}$  (ranging from 20 V to 1.5 kV for *n*-channel, and to 500 V for *p*-channel); the channel on-resistance  $R_{DS(ON)}$  (as low as 2 mΩ); the power-handling ability (as high as 1000 W with the case held unrealistically at 25°C); and the gate capacitance  $C_{iss}$  (as high as 10,000 pF), which must be charged and discharged during MOSFET switching, see ¶S below. Table 3.4a lists representative small-package *n*-channel parts rated to +250 V and *p*-channel parts of all sizes to -100 V; Table 3.4b extends the *n*-channel selection to higher voltage and current; more complete tables are found in Chapter 3x.

#### ¶R. Maximum Current.

MOSFET datasheets list a maximum continuous rated current, specified however at an unrealistic 25°C case temperature. This is calculated from  $I_{D(max)}^2 R_{DS(ON)} = P_{max}$ , substituting a maximum power  $P_{max} R_{OJC} = \Delta T_{JC} = 150^\circ\text{C}$  (see §9.4), where they have assumed  $T_{J(max)} = 175^\circ\text{C}$  (thus a 150°C  $\Delta T_{JC}$ ), and they use the value of  $R_{DS(ON)}$  (max) at 175°C from an  $R_{DS}$  tempco plot (e.g., see Figure 3.116). That is,  $I_{D(max)} = \sqrt{\Delta T_{JC} / R_{OJC} R_{ON}}$ . Some datasheets show the calculation for a more realistic 75°C or 100°C case temperature. Even so, you don't really want to run your MOSFET junction at 175°C, so we recommend using a lower maximum continuous  $I_D$  and corresponding  $P_{diss}$ .

#### ¶S. Gate Charge.

The capacitances in power MOSFETs that slow down switching are most easily analyzed with gate-charge plots, like Figure 3.101. First consider turn-ON: as current flows

into the gate capacitance  $C_{iss} + C_{rss}$  (dominated by  $C_{iss}$ ) the gate voltage rises. There is a switching delay, because the FET's drain remains off until the gate voltage is high enough for the FET to sink the drain current. Then the drain voltage starts to fall, as seen in Figures 3.102 and 3.103. The falling drain creates a reverse gate current  $I = C_{rss} dV_D/dt$  that prevents further increase in the gate voltage. Put another way, the falling slew-rate  $dV_D/dt = I_G/C_{rss}$  is set by the gate current available to charge the feedback (Miller) capacitance  $C_{rss}$ . When  $V_{DS}$  reaches zero the gate resumes charging, now at a slower rate because the  $C_{rss}$  contribution to total gate capacitance is larger at  $V_{DS}=0$ , see Figure 3.100). The MOSFET does not reach its intended low value of  $R_{DS(ON)}$  until the gate attains its full drive voltage. Turn-off proceeds similarly. MOSFET datasheets include values for  $C_{iss}$  and  $C_{rss}$ , but the latter is typically at  $V_{DS}=25$  V, so you need to go to the datasheet plots of capacitances versus drain voltage.

#### ¶T. MOSFET Gate Damage.

MOSFET gates typically have  $\pm 20$  V to  $\pm 30$  V maximum ratings, beyond which the very thin metal-oxide gate-channel insulator can be permanently damaged, see Figure 3.105. Be sure to discharge static charge before installation of discrete MOSFETs and MOS ICs.

#### ¶U. FET versus BJT for Power Switching.

See §3.5.4H; see also ¶Z below.

#### ¶V. MOSFET Switch Polarity.

Both *n*- and *p*-channel polarities of MOSFETs can be used to switch a voltage, see Figure 3.106 where most of the circuits show a conventional approach with a *p*-channel FET switching a positive voltage. But circuit E shows an *n*-channel FET doing the same task, with an additional voltage source powering the gate (the better-performing *n*-channel FET is preferred if it can be easily used, see §3.1.2). Figure 3.107 illustrates the use of photodiodes to power the high-side gates, to make "floating" switches.

#### ¶W. Power MOSFET Amplifiers.

Unlike bipolar power transistors, power MOSFETs have a wide safe-operating area (SOA) and do not suffer from second breakdown (see Figure 3.95), which is due to a localized thermal-runaway heating problem. Figure 3.119 shows typical class-AB biasing techniques necessary for use in linear power amplifiers.

#### ¶X. Depletion-mode Power MOSFETs.

Although most power MOSFETs are enhancement-mode types, *n*-channel depletion-mode types are available; §3.5.6D shows some applications. See also Table 3.6 on page 210.

#### ¶Y. Paralleling Power MOSFETs.

When used as switches, yes, but when used in power amplifiers, no, at least not without high-value source-ballast resistors! Figure 3.117B shows an elegant active-feedback

workaround for use with regulator pass elements.

#### ¶Z. IGBTs.

IGBTs are an alternative to power MOSFETs, see §3.5.7 where we show a comparison between power MOSFETs, IGBTs and BJTs. They're primarily useful at voltages above 300 V and switching rates below 100 kHz, though there are some nice IGBTs for use at RF, for example the IRGB4045, good for 150 W or more at 20 MHz.

# OPERATIONAL AMPLIFIERS

## CHAPTER 4

### 4.1 Introduction to op-amps – the “perfect component”

In the previous three chapters we learned about circuit design with “discrete components,” both active and passive. Our basic building blocks were transistors, both bipolar (BJT) and field-effect (FET), along with the resistors, capacitors, and other components that are needed to set bias, couple and block signals, create load impedances, and so on.

With those tools we have gone quite far. We’ve learned how to design simple power supplies, signal amplifiers and followers, current sources, dc and differential amplifiers, analog switches, power drivers and regulators, and even some rudimentary digital logic.

But we’ve also learned to struggle with imperfections. Voltage amplifiers suffer from nonlinearity (a grounded-emitter amplifier with a 1 mV input signal has  $\sim 1\%$  distortion), which you can trade off against voltage gain (by adding emitter degeneration); differential amplifiers have input unbalance, typically tens of millivolts (with bipolar transistors), ten times more with discrete junction-FETs (JFETs); in bipolar design you have to worry about input current (often substantial), and the ever-present  $V_{BE}$  and its variation with temperature; in FET design you trade absence of input current for unpredictability of  $V_{GS}$ ; and so on.

We’ve seen hints that things can be better, in particular the remarkable linearizing effects of negative feedback (§2.5.3), and its ability to make overall *circuit* performance less dependent on *component* imperfections. It is negative feedback that gives the emitter-degenerated amplifier its linearity advantage over the grounded-emitter amplifier (at the cost of voltage gain). And in the high-loop-gain limit, negative feedback promises circuit performance largely independent of transistor imperfections.

Promised, but not yet delivered: the high-gain amplifier blocks we need to get high loop gain in a feedback arrangement still involve substantial design efforts – the hallmark of complex circuits implemented with discrete (as opposed to integrated) components.

With this chapter we enter the promised land! The op-amp is, essentially, a “perfect part”: a complete integrated amplifier gain block, best thought of as a dc-coupled differential amplifier with single-ended output, and with extraordinarily high gain. It also excels in precise input symmetry and nearly zero input current. Op-amps are designed as “gain engines” for negative feedback, with such high gain that the circuit performance is set almost entirely by the feedback circuitry. Op-amps are small and inexpensive, and they should be the starting point for nearly every analog circuit you design. In most op-amp circuit designs we’re in the regime where they are essentially perfect: with them we will learn to build nearly perfect amplifiers, current sources, integrators, filters, regulators, current-to-voltage converters, and a host of other modules.

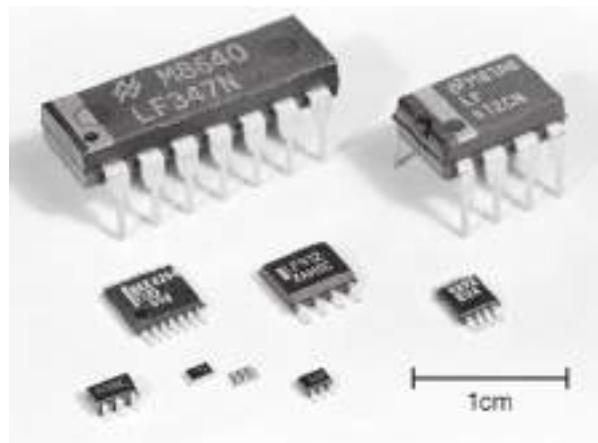
Op-amps are our first example of *integrated circuits* – many individual circuit elements, such as transistors and resistors, fabricated and interconnected on a single “chip” of silicon.<sup>1</sup> Figure 4.1 shows some IC op-amp packaging schemes.

#### 4.1.1 Feedback and op-amps

We first met negative feedback in Chapter 2, where we saw that the process of coupling the output back, in such a way as to cancel some of the input signal, improved characteristics such as linearity, flatness of response, and predictability. As we saw quantitatively, the more negative feedback that is used, the less the resultant amplifier characteristics depend on the characteristics of the open-loop (no-feedback) amplifier, ultimately depending only on the properties of the feedback network itself. Operational amplifiers are typically used in this *high-loop-gain* limit, with *open-loop* voltage gain (no feedback) of a million or so.

A feedback network can be frequency-dependent, to produce an equalization amplifier (for example the

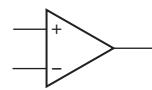
<sup>1</sup> The first operational amplifiers were made with vacuum tubes, followed by implementations with discrete transistors. See §4x.1 for a description (with photo and schematics) of a once-popular vacuum-tube op-amp, the Philbrick K2-W.



**Figure 4.1.** Op-amps (and other linear ICs) come in a bewildering variety of “packages,” most of which are represented in this photograph. Top row, left to right: 14-pin plastic dual in-line package (DIP), 8-pin plastic DIP (“mini-DIP”). Middle row: 14-pin thin-shrink small-outline package (TSSOP), 8-pin small-outline package (SO-8), 8-pin TSSOP (“ $\mu$ MAX”). Bottom row: 5-pin small-outline transistor package (SOT23), 6-ball chip-scale package (CSP – top and bottom views), 5-pin SC-70. The 14-pin packages hold quad op-amps (i.e., four independent op-amps), the 8-pin packages hold duals, and the rest are singles. (TSSOP and smaller packages courtesy of Travis Eichhorn, Maxim Semiconductor.)

treble and bass “tone control” stage of amplification that you find in most audio systems); or it can be amplitude-dependent, producing a nonlinear amplifier (a popular example is a logarithmic amplifier, built with feedback that exploits the logarithmic  $V_{BE}$  versus  $I_C$  of a diode or transistor). It can be arranged to produce a current source (near-infinite output impedance) or a voltage source (near-zero output impedance), and it can be connected to generate very high or very low input impedance. Speaking in general terms, the property that is sampled to produce feedback is the property that is improved. Thus, if you feed back a signal proportional to the output current, you will generate a good current source.

As we remarked in §2.5.1, feedback can be arranged intentionally to be *positive*, for example to make an oscillator, or, as we’ll see later, to make a Schmitt trigger circuit. That’s the *good* kind of positive feedback. The bad kind occurs, uninvited (and unwelcome), when a negative-feedback circuit is burdened with sufficient accumulated phase shifts at some frequency to produce overall positive feedback, and oscillations. This can occur for a variety of reasons. We’ll discuss this important subject, and see how to prevent unwanted oscillations by *frequency compensation*, the topic of §4.9 at the end of the chapter.



**Figure 4.2.** Op-amp symbol.

Having made these general comments, we now look at a few feedback examples with op-amps.

### 4.1.2 Operational amplifiers

The operational amplifier is a very high-gain dc-coupled differential amplifier with a single-ended output. You can think of the classic long-tailed pair (§2.3.8) with its two inputs and single output as a prototype, although real op-amps have much higher gain (typically  $10^5$  to  $10^6$ ) and lower output impedance, and they allow the output to swing through most or all of the supply range (you often use a split supply, for example  $\pm 5$  V). Operational amplifiers are available in literally thousands of types, with the universal symbol shown in Figure 4.2, where the (+) and (−) inputs do as expected: the output goes positive when the noninverting input (+) goes more positive than the inverting input (−), and vice versa. The (+) and (−) symbols don’t mean that you have to keep one positive with respect to the other, or anything like that; they just tell you the relative phase of the output (which is important to keep negative feedback *negative*). Using the words “noninverting” and “inverting,” rather than “plus” and “minus” helps avoid confusion. Power-supply connections are frequently not displayed, and there is no ground terminal. Operational amplifiers have enormous voltage gain, and they are *never* (well, hardly ever) used without feedback. Think of an op-amp as fodder for feedback. The open-loop gain is so high that, for any reasonable closed-loop gain, the characteristics depend on only the feedback network. Of course, at some level of scrutiny this generalization must fail. We will start with a naïve view of op-amp behavior and fill in some of the finer points later, when we need to.

There are literally thousands of different op-amps available, offering various performance tradeoffs that we will explain later (look ahead to Tables 4.2a,b, 5.5, or 8.3 if you want to see a small sample of what’s available). A very good all-around performer is the popular LF411 (“411” for short), originally introduced by National Semiconductor. Like many op-amps, it is a wee beastie packaged in the so-called mini-DIP (dual in-line package) or SOIC (small-outline IC), and it looks as shown in Figure 4.3. It is inexpensive (less than \$1) and easy to use; it comes in an improved grade (LF411A) and also in a version containing

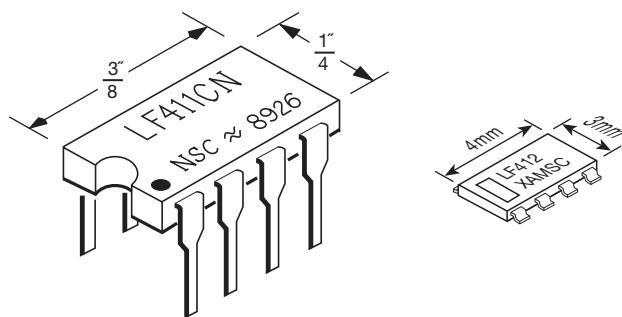


Figure 4.3. Mini-DIP and SOIC packages.

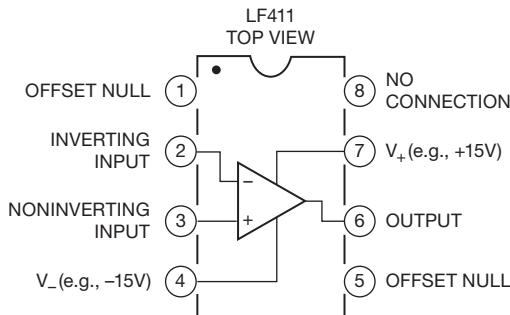


Figure 4.4. Pin connections for LF411 op-amp in 8-pin DIP.

two independent op-amps (LF412, called a “dual” op-amp). We will adopt the LF411/LF412 throughout this chapter as our “standard” op-amp, and we recommend it (or perhaps the versatile LMC6482) as a good starting point for your circuit designs.

Inside the 411 is a piece of silicon containing 24 transistors (21 BJTs, 3 FETs), 11 resistors, and 1 capacitor. (You can look ahead to Figure 4.43 on page 243 to see a simplified circuit diagram of its innards.) The pin connections are shown in Figure 4.4. The dot in the upper-left-hand corner, or notch at the end of the package, identifies the end from which to begin counting the pin numbers. As with most electronic packages, you count pins counterclockwise, viewing from the top. The “offset null” terminals (also known as “balance” or “trim”) have to do with correcting (externally) the small asymmetries that are unavoidable when making the op-amp. More about this later in the chapter.

### 4.1.3 The golden rules

Here are the simple rules for working out op-amp behavior with external negative feedback. They’re good enough for almost everything you’ll ever do.

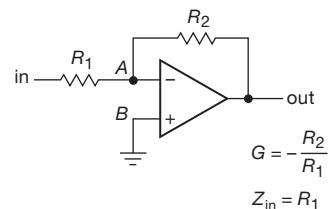


Figure 4.5. Inverting amplifier.

First, the op-amp voltage gain is so high that a fraction of a millivolt between the input terminals will swing the output over its full range, so we ignore that small voltage and state golden rule I.

- I.** The output attempts to do whatever is necessary to make the voltage difference between the inputs zero.

Second, op-amps draw very little input current (about 50 pA for the inexpensive JFET-input LF411, and often less than a picoamp for MOSFET-input types); we round this off, stating golden rule II.

- II.** The inputs draw no current.

One important note of explanation: golden rule I doesn’t mean that the op-amp actually changes the voltage at its *inputs*. It can’t do that. (How could it, and be consistent with golden rule II?) What it does is “look” at its input terminals and swing its output terminal around so that the external feedback-network brings the input differential to zero (if possible).

These two rules get you quite far. We illustrate with some basic and important op-amp circuits, and these will prompt a few cautions listed in §4.2.7.

## 4.2 Basic op-amp circuits

### 4.2.1 Inverting amplifier

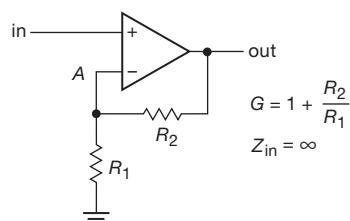
Let’s begin with the circuit shown in Figure 4.5. The analysis is simple, if you remember your golden rules.

1. Point B is at ground, so rule I implies that point A is also.
2. This means that (a) the voltage across  $R_2$  is  $V_{out}$  and (b) the voltage across  $R_1$  is  $V_{in}$ .
3. So, using rule II, we have  $V_{out}/R_2 = -V_{in}/R_1$ .

In other words, the voltage gain ( $G_V \equiv V_{out}/V_{in}$ ) is

$$G_V = -\frac{R_2}{R_1} \quad (4.1)$$

Later you will see that it’s sometimes better not to ground B directly, but through a resistor – but don’t worry about that now.



**Figure 4.6.** Noninverting amplifier.

Our analysis seems almost too easy! In some ways it obscures what is actually happening. To understand how feedback works, just imagine some input level, say +1 volt. For concreteness, imagine that  $R_1$  is 10k and  $R_2$  is 100k. Now, suppose the output decides to be uncooperative, and sits at zero volts. What happens?  $R_1$  and  $R_2$  form a voltage divider, holding the inverting input at +0.91 volts. The op-amp sees an enormous input unbalance, forcing the output to go negative. This action continues until the output is at the required -10.0 volts, at which point both op-amp inputs are at the same voltage, namely ground. Similarly, any tendency for the output to go more negative than -10.0 volts will pull the inverting input below ground, forcing the output voltage to rise.

What is the input impedance? Simple. Point A is always at zero volts (it's called a *virtual ground*). So  $Z_{in} = R_1$ . At this point you don't yet know how to figure the output impedance; for this circuit, it's a fraction of an ohm.

Note that this analysis is true even for dc – it's a dc amplifier. So if you have a signal source that has a dc offset from ground (collector of a previous stage, for instance), you may want to use a coupling capacitor (sometimes called a blocking capacitor, since it blocks dc but couples the signal). For reasons you will see later (having to do with departures of op-amp behavior from the ideal), it is usually a good idea to use a blocking capacitor if you're interested only in ac signals anyway.

This circuit is known as an *inverting amplifier*. Its one undesirable feature is the low input impedance, particularly for amplifiers with large (closed-loop) voltage gain, where  $R_1$  tends to be rather small. That is remedied in the next circuit (Figure 4.6).

### 4.2.2 Noninverting amplifier

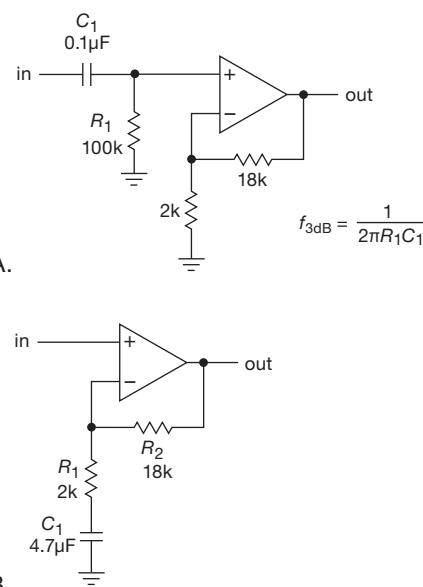
Consider Figure 4.6. Again, the analysis is simplicity itself:

$$V_A = V_{in}.$$

But  $V_A$  comes from a voltage divider:  $V_A = V_{out}R_1/(R_1 + R_2)$ . Set  $V_A = V_{in}$ , and you get a voltage gain of

$$G_V = 1 + R_2/R_1. \quad (4.2)$$

This is a *noninverting amplifier*. In the approximation we are using, the input impedance is infinite (with the JFET-input 411 it would be  $10^{12}\Omega$  or more; a BJT-input op-amp will typically exceed  $10^8\Omega$ ). The output impedance is still a fraction of an ohm. As with the inverting amplifier, a detailed look at the voltages at the inputs will convince you that it works as advertised.



**Figure 4.7.** Amplifiers for ac signals: A. ac-coupled noninverting amplifier, B. blocking capacitor rolls off the gain to unity at dc.

### A. An ac amplifier

The basic noninverting amplifier, like the inverting amplifier earlier, is a dc amplifier. If the signal source is ac-coupled, you must provide a return to ground for the (very small) input current, as in Figure 4.7A. The component values shown give a voltage gain of 10 and a low-frequency 3 dB point of 16 Hz.

If only ac signals are being amplified, it is often a good idea to “roll off” the gain to unity at dc, especially if the amplifier has large voltage gain, to reduce the effects of finite “input offset voltage” (§4.4.1A). The circuit in Figure 4.7B has a low-frequency 3 dB point of 17 Hz, the frequency at which the impedance of the capacitor  $C_1$  equals  $R_1$ , or 2.0k. Note the large capacitor value required. For noninverting amplifiers with high gain, the capacitor in this ac amplifier configuration may be undesirably large. In that

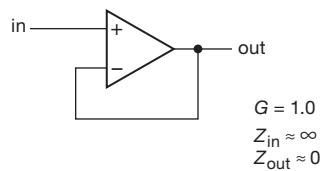


Figure 4.8. Op-amp follower.

case it may be preferable to omit the capacitor and trim the offset voltage to zero, as we will discuss later. An alternative is to raise  $R_1$  and  $R_2$ , perhaps using a T network for the latter (Figure 4.66 on page 259).

In spite of its desirable high input impedance, the non-inverting amplifier configuration is not necessarily to be preferred over the inverting amplifier configuration in all circumstances. As we will see later, the inverting amplifier puts less demand on the op-amp, and therefore gives somewhat better performance. In addition, its virtual ground provides a handy way to combine several signals without interaction. Finally, if the circuit in question is driven from the (stiff) output of another op-amp, it makes no difference whether the input impedance is 10k (say) or infinity, because the previous stage has no trouble driving it in either case.

#### 4.2.3 Follower

Figure 4.8 shows the op-amp version of an emitter follower. It is simply a noninverting amplifier with  $R_1$  infinite and  $R_2$  zero (gain = 1). An amplifier of unity gain is sometimes called a *buffer* because of its isolating properties (high input impedance, low output impedance).

#### 4.2.4 Difference amplifier

The circuit in Figure 4.9A is a *difference amplifier* (sometimes called a *differential amplifier*) with gain  $R_2/R_1$ . This circuit requires precise resistor matching to achieve high common-mode rejection ratios (CMRR). You may be lucky and find a batch of 100k 0.01% resistors at an electronics flea market or surplus outlet; otherwise you can buy precision resistor arrays, with close matching of ratios and temperature coefficients.<sup>2</sup> All your difference amplifiers

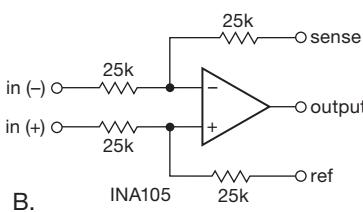
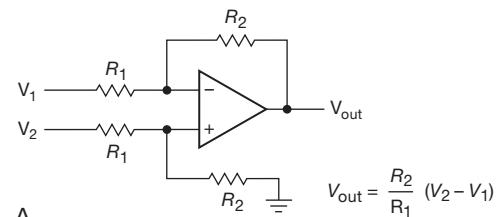


Figure 4.9. Classic difference amplifier: A. Op-amp with matched resistor ratios. B. Integrated version, with uncommitted “sense” and “reference” pins. In the best grade (INA105A) the resistor ratio is matched to better than 0.01%, with a temperature coefficient better than 5 ppm/ $^{\circ}\text{C}$ .

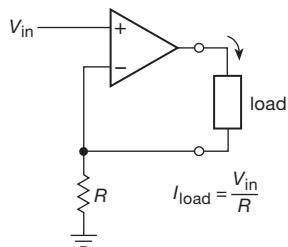
will have unity gain, but that's easily remedied with further (single-ended) stages of gain. If you can't find good resistors (or even if you can!), you should know that you can buy this circuit as a convenient packaged difference amplifier, with well-matched resistors; examples are the INA105 or AMP03 ( $G = 1$ ), INA106 ( $G=10$  or 0.1), and INA117 or AD629 ( $G = 1$  with input dividers; input signals to  $\pm 200$  V) from TI/Burr-Brown and Analog Devices (many more are listed in Table 5.7 on page 353). The unity-gain INA105 configuration is shown in Figure 4.9B, with its uncommitted “sense” and “reference” pins. You get the classic difference amplifier by connecting *sense* to the output and *ref* to ground. But the additional flexibility lets you make all sorts of nifty circuits, such as a precision unity-gain inverter, noninverting gain-of-2 amplifier, and noninverting gain-of-0.5 amplifier. We treat difference amplifiers in greater detail in §5.14.

**Exercise 4.1.** Show how to make these three circuits with an INA105.

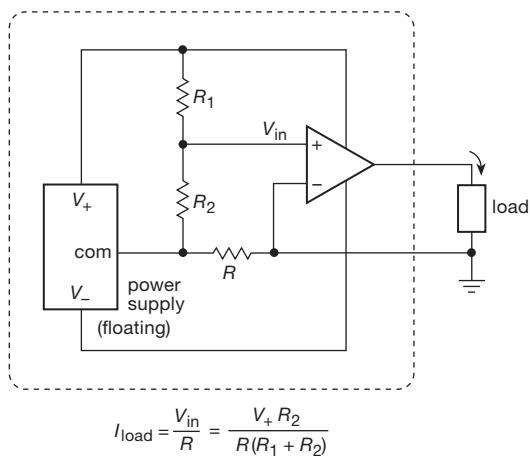
There are, in addition, more sophisticated differential amplifier configurations, known officially as “instrumentation amplifiers”; they are discussed in detail in §§5.15 and 5.16, along with a listing in Table 5.8 on page 363.

<sup>2</sup> For example, the BI Technologies type 664 thin-film quad (four resistors of the same value) in an 8-lead surface-mount IC package (SOIC); these come in accuracies to 0.1%, ratio tracking to 0.05%, and tracking temperature coefficients to  $\pm 5$  ppm/ $^{\circ}\text{C}$ . They are inexpensive (about \$2 for the best grade), and available from Mouser Electronics, among others. Companies like Vishay have offerings with astonishingly good per-

formance: their best resistor arrays specify worst-case ratio tracking to 0.001%, and tracking temperature coefficient (tempco) to  $\pm 0.1$  ppm/ $^{\circ}\text{C}$ .



**Figure 4.10.** Basic op-amp current source (floating load).  $V_{in}$  might come from a voltage divider, or it could be a signal that varies with time.



**Figure 4.11.** Current source with grounded load and floating power supply.

#### 4.2.5 Current sources

The circuit in Figure 4.10 approximates an ideal current source, without the  $V_{BE}$  offset of a transistor current source. Negative feedback results in  $V_{in}$  at the inverting input, producing a current  $I = V_{in}/R$  through the load. The major disadvantage of this circuit is the “floating” load (neither side grounded). You couldn’t generate a usable sawtooth wave with respect to ground with this current source, for example. One solution is to float the whole circuit (power supplies and all) so that you can ground one side of the load (Figure 4.11). The circuit in the box is the previous current source, with its power supplies shown explicitly.  $R_1$  and  $R_2$  form a voltage divider to set the current. If this circuit seems confusing, it may help to remind yourself that “ground” is a relative concept. Any one point in a circuit could be called ground. This circuit is useful for generating currents into a load that is returned to ground, but it has the disadvantage that the control input is now floating, so you cannot program the output current with an in-

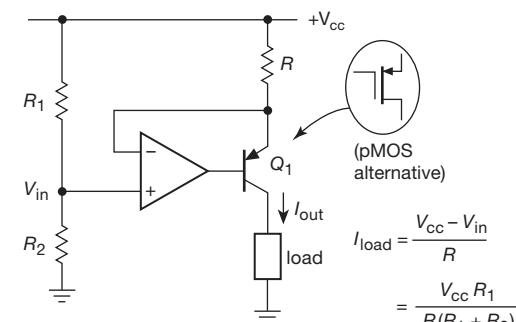
put voltage referenced to ground. In addition, you’ve got to make sure that the floating power supply is truly floating – for example, you’d have trouble making a microamp dc current source this way if you tried to use a standard wall-plug-powered dc power supply, because capacitance between windings in its transformer would introduce reactive currents, at the 60 Hz line frequency, that might well exceed the desired microamp output current; one possible solution would be to use batteries. Some other approaches to this problem are presented in Chapter 9 (§9.3.14) in the discussion of constant-current power supplies.<sup>3</sup>

#### A. Current sources for loads returned to ground

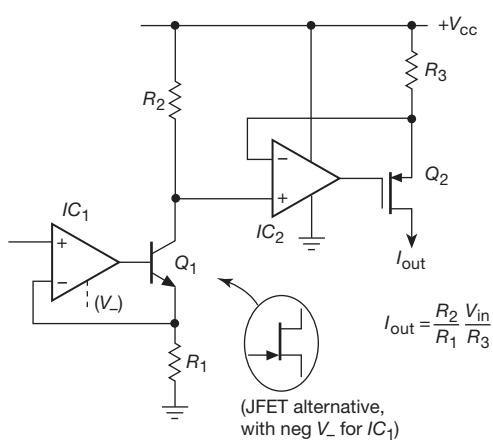
With an op-amp and external transistor it is possible to make a simple high-quality current source for a load returned to ground; a little additional circuitry makes it possible to use a programming input referenced to ground (Figure 4.12). In the first circuit, feedback forces a voltage  $V_{CC} - V_{in}$  across  $R$ , giving an emitter current (and therefore an output current)  $I_E = (V_{CC} - V_{in})/R$ . There are no  $V_{BE}$  offsets, or their variations with temperature, with  $I_C$ , with  $V_{CE}$ , etc., to worry about. The current source is imperfect (ignoring op-amp errors:  $I_B$ ,  $V_{os}$ ) only insofar as the small base current may vary somewhat with  $V_{CE}$  (assuming the op-amp draws no input current), not too high a price to pay for the convenience of a grounded load; a Darlington for  $Q_1$  would reduce this error considerably. This error comes about, of course, because the op-amp stabilizes the *emitter* current, whereas the load sees the *collector* current. A variation of this circuit, using a MOSFET instead of a bipolar transistor, avoids this problem altogether, since FETs draw no dc gate current (but large power MOSFETs have plenty of input capacitance, which can cause problems; see the comment at the end of this subsection).

With this circuit the output current is proportional to the voltage drop below  $V_{CC}$  applied to the op-amp’s non-inverting input; in other words, the programming voltage is referenced to  $V_{CC}$ , which is fine if  $V_{in}$  is a fixed voltage generated by a voltage divider, but an awkward situation if an external input is to be used. This is remedied in the second circuit, in which a similar current source with an *npn* transistor is used to convert an input voltage (referenced to ground) to a  $V_{CC}$ -referenced input to the final current

<sup>3</sup> Another limitation of op-amp current-source circuits is their degraded performance at higher frequencies: an op-amp’s output is inherently low impedance, (typically a push-pull follower, with  $R_{out} \sim 100\Omega$ , look ahead to Figure 4.43), so a current-source circuit must rely on feedback (which declines with increasing frequency) to raise the op-amp’s output impedance. See further discussion in §§4.2.5B, and 4.4.4.



A.

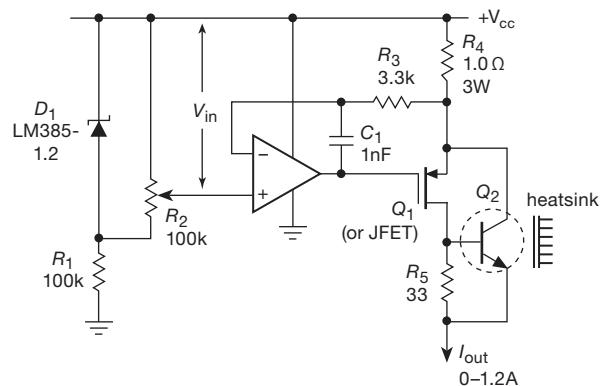


B.

**Figure 4.12.** Current sources for grounded loads that don't require a floating power supply. The op-amps may need to have rail-to-rail input and output capability (RRIO); see text.

source; for the latter we've used a *p*-channel MOSFET for variety (and to eliminate the small base-current error you get with bipolar transistors). Op-amps and transistors are inexpensive. Don't hesitate to use a few extra components to improve performance or convenience in circuit design.

One important note about these circuits: at low output currents the voltage across the emitter (or source) resistors may be quite small, which means that the op-amps must be able to operate with their inputs near or at the positive supply voltage. For example, in the circuit of Figure 4.12B IC<sub>2</sub> needs to operate with its inputs close to the positive supply rail. Don't assume that a given op-amp will do this, without explicit permission from the datasheet! The LF411's datasheet waffles a bit on this, but grudgingly admits that it will work, albeit with degraded performance, with the inputs at the positive rail. (It will not, however, work down to the negative rail; but with IC<sub>1</sub> powered from split supply voltages there's no problem there.) By contrast, op-amps like the LMC7101 or LMC6482 guarantee proper



**Figure 4.13.** FET-bipolar current source suitable for high currents.

operation all the way to (and a bit beyond) the positive rail (see the "Swing to supplies?" column in Table 4.2a on page 271). Alternatively, the op-amp could be powered from a separate  $V_+$  voltage higher than  $V_{\text{CC}}$ .

**Exercise 4.2.** What is the output current in the last circuit for a given input voltage  $V_{\text{in}}$ ? (Did we get it right in the figure?)

Figure 4.13 shows an interesting variation on the op-amp-transistor current source. Although you can get plenty of current with a simple power MOSFET, the high inter-electrode capacitances of high-current FETs may cause problems. When a relatively low-current MOSFET<sup>4</sup> is combined with a high-current *npn* power transistor, this circuit has the advantage of zero base current error (which you get with FETs) along with much smaller input capacitance. In this circuit, which is analogous to the "complementary Darlington" (or Sziklai circuit; see §2.4.2A), bipolar transistor  $Q_2$  kicks in when the output current exceeds about 20 mA.

Lest we leave the wrong impression, we emphasize that the simpler MOSFET-only circuit (in the manner of Figure 4.12B) is a preferable configuration, given the major drawback of power BJTs, namely their susceptibility to "second breakdown" and consequent limit on safe operating area (as we saw in §3.5.1B, see particularly Figure 3.95). Big power MOSFETs have large input capacitance, so in such a circuit you should use a network like Figure 4.13's  $R_3C_1$  to prevent oscillation.

### B. Howland current source

Figure 4.14 shows a nice "textbook" current source. If the resistors are chosen so that  $R_3/R_2 = R_4/R_1$ , then it can be shown that  $I_{\text{load}} = -V_{\text{in}}/R_2$ .

**Exercise 4.3.** Show that the preceding result is correct.

<sup>4</sup> Such as a BS250P or BSS84, see Table 3.4a on page 188.

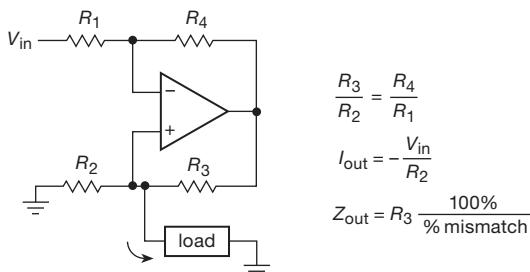


Figure 4.14. Howland current source.

This sounds great, but there's a hitch: the resistor ratios must be matched exactly; otherwise it isn't a perfect current source. Even so, its performance is limited by the op-amp's common-mode rejection ratio (CMRR, §2.3.8). For large output currents the resistors must be small, and the compliance is limited. Also, at high frequencies (where the loop gain is low, as we'll learn shortly) the output impedance can drop from the desired value of infinity to as little as a few hundred ohms (the op-amp's open-loop output impedance). These drawbacks limit the applicability of this clever circuit.

You can convert this circuit into a noninverting current source by grounding  $R_1$  (where  $V_{\text{in}}$  is shown) and applying the control input voltage  $V_{\text{in}}$  instead to  $R_2$ .

Figure 4.15 is a nice improvement on the Howland circuit, because the output current is sourced through a sense resistor  $R_s$  whose value you can choose independently of the matched resistor array (with resistor pairs  $R_1$  and  $R_2$ ). The best way to understand this circuit is to think of  $\text{IC}_1$  as a difference amplifier whose output *sense* and *reference* connections sample the drop across  $R_s$  (i.e., the current); the latter is buffered by follower  $\text{IC}_2$  so there is no current error.

For this configuration you can exploit the internal precision matched resistors in an integrated difference amplifier: use something like an INA106 for  $R_1$ ,  $R_2$ , and  $\text{IC}_1$ , wired "backwards" (for  $G=0.1$ ) to reduce the drop across the sense resistor. See §5.14 and Table 5.7 on page 353.

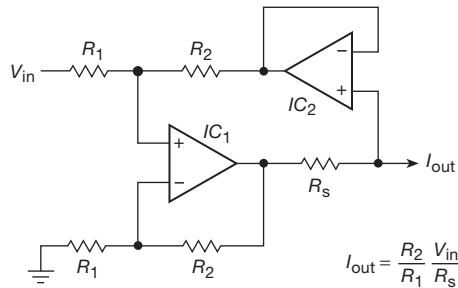


Figure 4.15. Bipolarity current source–sink.

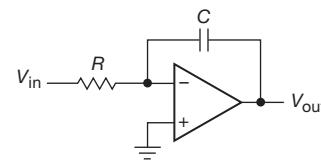


Figure 4.16. Integrator.

or

$$V_{\text{out}}(t) = -\frac{1}{RC} \int V_{\text{in}}(t) dt + \text{const.} \quad (4.3)$$

The input can, of course, be a current, in which case  $R$  is omitted.

As an example, if we choose  $R = 1\text{M}\Omega$  and  $C = 0.1\text{\mu F}$  in this circuit, then a constant dc input of +1 V produces 1  $\mu\text{A}$  of current into the summing junction, hence an output voltage that is ramping downward at  $dV_{\text{out}}/dt = -V_{\text{in}}/RC = -10\text{ V/s}$ . To say it algebraically, for a constant  $V_{\text{in}}$  or constant  $I_{\text{in}}$ ,

$$\Delta V_{\text{out}} = -\frac{V_{\text{in}}}{RC} \Delta t = -\frac{I_{\text{in}}}{C} \Delta t.$$

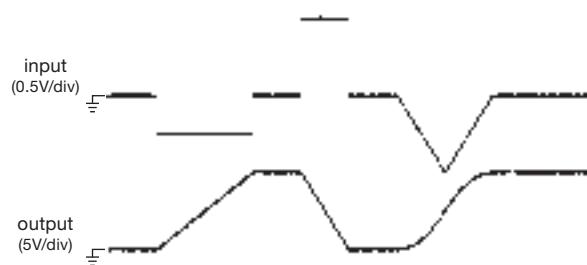
We rigged up the integrator of Figure 4.16, with  $R = 1\text{M}\Omega$  and  $C = 1\text{nF}$ , and drove it with the simple test waveform shown in Figure 4.17. Without having taken a math class, the thing knows calculus!

Sharp-eyed readers may have noticed that this circuit doesn't have any feedback at dc, and so there's no way for it to have a stable quiescent point: for *any* nonzero input voltage  $V_{\text{in}}$ , the output is going *somewhere!* As we'll see shortly, even with  $V_{\text{in}}$  exactly at zero volts, the output tends to wander off, owing to op-amp imperfections (non-zero input current, and "offset voltage"). These latter problems can be minimized by careful choice of op-amp and circuit values; but even so you usually have to provide some way to reset the integrator. Figure 4.18 shows how this is commonly done, either with a reset switch (both discrete JFET and integrated CMOS analog switch examples are shown)

## 4.2.6 Integrators

Op-amps allow you to make nearly perfect integrators, without the restriction that  $V_{\text{out}} \ll V_{\text{in}}$ . Figure 4.16 shows how it's done. Input current  $V_{\text{in}}/R$  flows through  $C$ . Because the inverting input is a virtual ground, the output voltage is given by

$$V_{\text{in}}/R = -C(dV_{\text{out}}/dt)$$

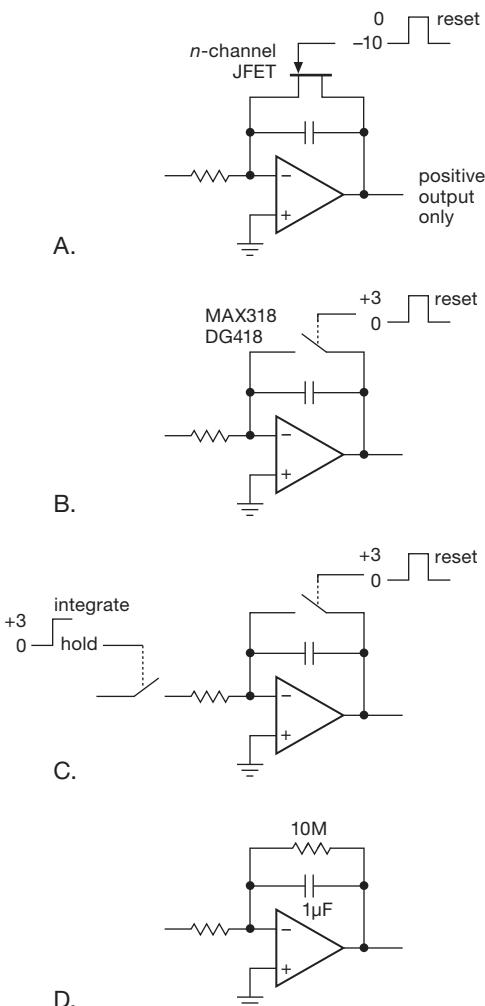


**Figure 4.17.** Integrator waveforms. The output can go anywhere it wants to, unlike our simple  $RC$  “integrator” of §1.4.4. Horizontal: 10 ms/div.

or with a large-value feedback resistor across the integrating capacitor. Closing a reset switch<sup>5</sup> (Figures 4.18A,B) zeroes the integrator by rapidly discharging the capacitor, while allowing perfect integration when open. The use of a feedback resistor (Figure 4.18D) produces stable biasing by restoring feedback at dc (where the circuit behaves like a high-gain inverting amplifier), but the effect is to roll off the integrator action at very low frequencies,  $f < 1/R_f C$ . An additional series analog switch at the input (Figure 4.18C) lets you control the intervals during which the integrator is active; when that switch is open the integrator output is frozen at its last value.

You don’t have to worry about zeroing the integrator, of course, if it’s part of a larger circuit that does the right thing. We’ll see a beautiful example shortly (§4.3.3), namely an elegant triangle-wave generator, in which an untamed integrator is just what you want.

This first look at the op-amp integrator assumes that the op-amp is perfect, in particular that (a) the inputs draw no current, and (b) the amplifier is balanced with both inputs at precisely the same voltage. When our op-amp honeymoon is over we’ll see that real op-amps do have some input current (called “bias current,”  $I_B$ ), and that they exhibit some voltage imbalance (called “offset voltage,”  $V_{OS}$ ). These imperfections are not large – bias currents of picoamps are routine, as are offset voltages of less than a millivolt – but they can cause problems with circuits like integrators, in which the effect of a small error grows with time. We’ll deal with these essential topics later in the chapter (§4.4), after you’re comfortable with the basics.



**Figure 4.18.** Op-amp integrators with reset switches.

#### 4.2.7 Basic cautions for op-amp circuits

- In all op-amp circuits, golden rules I and II (§4.1.3) are obeyed only if the op-amp is in the active region, i.e., inputs and outputs are not saturated at one of the supply voltages.

For instance, overdriving one of the amplifier configurations will cause output clipping at output swings near  $V_{CC}$  or  $V_{EE}$ . During clipping, the inputs will no longer be maintained at the same voltage. The op-amp output cannot swing beyond the supply voltages (typically it can swing only to within 2 V of the supplies, though certain op-amps are designed to swing all the way to one supply or the other, or to both; the latter are known as “rail-to-rail output” op-amps). Likewise, the output compliance of an op-amp current source is set by the same limitation.

<sup>5</sup> Refer back to §3.4 for a detailed discussion of FET switches.

The current source with floating load (Figure 4.10), for instance, can put a maximum of  $V_{CC} - V_{in}$  across the load in the “normal” direction (current in the same direction as applied voltage) and  $V_{in} - V_{EE}$  in the reverse direction.<sup>6</sup>

- The feedback must be arranged so that it is negative. This means (among other things) that you must not mix up the inverting and noninverting inputs. We’ll learn later that you can get yourself into similar problems if you rig up a feedback network that has lots of phase shift at some frequency.
- There must always be feedback at dc in an op-amp circuit. Otherwise the op-amp is guaranteed to go into saturation.

For instance, we were able to put a capacitor from the feedback network to ground in the noninverting amplifier (to reduce gain to 1 at dc, Figure 4.7B), but we could not similarly put a capacitor in series between the output and the inverting input. Likewise, an integrator will ultimately saturate without some additional circuitry such as a reset switch.

- Some op-amps have a relatively small maximum differential-input voltage limit. The maximum voltage difference between the inverting and noninverting inputs may be limited to as little as 5 volts in either polarity. Breaking this rule will cause large input currents to flow, with degradation or destruction of the op-amp.
- Op-amps are high-gain devices, often having plenty of gain even at radiofrequencies, where the inductances in the power-rail wiring can lead to instabilities in the amplifiers. We solve this issue with mandatory (we mean it!) bypass capacitors on the op-amp supply rails.<sup>7</sup> Note: The figures in this chapter and elsewhere (and generally in the real world) do not show bypass capacitors, for simplicity. You have been warned.

We take up some more issues of this type in §4.4, and again in Chapter 5 in connection with precision circuit design.

<sup>6</sup> The load could be rather strange, e.g., it might contain batteries, requiring the reverse sense of voltage to get a forward current; the same thing might happen with an inductive load driven by changing currents.

<sup>7</sup> When we were young we were taught that each op-amp needed its own set of bypass capacitors. But with experience we’ve come to realize that one pair of capacitors can work to stabilize nearby op-amps. Furthermore, local wiring inductance with multiple sets of bypass capacitors can lead to resonances, which allow one op-amp to interfere with another. For example, if  $L=25\text{ nH}$  and  $C=0.01\mu\text{F}$ , then  $f_{LC}=10\text{ MHz}$ , and  $X_{LC}=1.6\Omega$ . The impedance peak at resonance will be  $Q$  times higher. You can solve this problem by adding an additional parallel lossy bypass capacitor, such as a small electrolytic. Its equivalent series resistance, of order  $0.5\Omega$  or more, acts to damp the resonant  $Q$ .

## 4.3 An op-amp smorgasbord

In the following examples we skip the detailed analysis, leaving that fun for the reader.

### 4.3.1 Linear circuits

#### A. Optional inverter

The circuits in Figure 4.19 let you invert, or amplify without inversion, by flipping a switch. The voltage gain is either  $+1$  or  $-1$ , depending on the switch position. The “switches” can be CMOS analog switches<sup>8</sup>, which let you control the sense of inversion with a (digital) signal. The clever variation of Figure 4.20 lets you vary the gain continuously from follower to inverter. And when the pot  $R_1$  is at mid-position, the circuit does nothing at all!

**Exercise 4.4.** Show that the circuits in Figure 4.19 work as advertised.

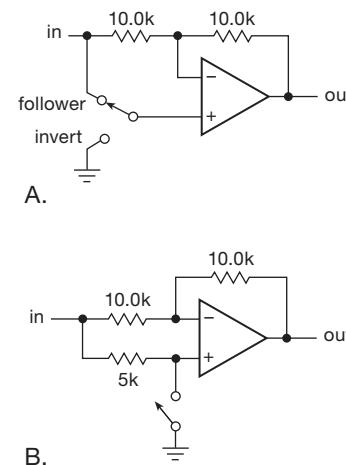


Figure 4.19. Optional inverters;  $G = \pm 1.0$

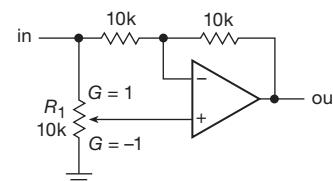


Figure 4.20. Follower-to-inverter: continuously adjustable gain from  $G = +1$  to  $G = -1$ .

<sup>8</sup> For example the ADG419 or MAX319  $\pm 20\text{ V}$  SPDT switches in convenient 8-pin packages, see §3.4 and Table 3.3 on page 176.

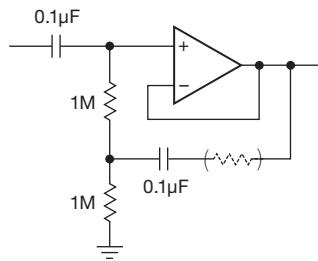


Figure 4.21. Op-amp follower with bootstrap.

### B. Follower with bootstrap

As with transistor amplifiers, the bias path can compromise the high input impedance you would otherwise get with an op-amp, particularly with ac-coupled inputs, for which a resistor to ground is mandatory. If that is a problem, the bootstrap circuit shown in Figure 4.21 is a possible solution. As in the transistor bootstrap circuit (§2.4.3), the  $0.1\ \mu\text{F}$  capacitor makes the upper  $1\text{M}$  resistor look like a high-impedance current source to input signals. The low-frequency rolloff for this circuit will begin at about 10 Hz, dropping at 12 dB per octave for frequencies somewhat below this.<sup>9</sup> This circuit may exhibit some frequency peaking, analogous to the Sallen-and-Key circuit of §4.3.6; this can be tamed by adding a resistor of  $1\text{-}10\text{k}$  in series with the feedback capacitor.

The very low input current (and therefore high input impedance) of FET-input op-amps generally make bootstrapping unnecessary; you can use  $10\text{ M}$  or larger resistors for the input bias path in ac-coupled amplifiers.

### C. Ideal current-to-voltage converter

Remember that the humble resistor is the simplest  $I$ -to- $V$  converter. However, it has the disadvantage of presenting a nonzero impedance to the source of input current; this can be fatal if the device providing the input current has very little compliance or does not produce a constant current as the output voltage changes. A good example is a *photovoltaic cell*, a diode junction that has been optimized as a light detector. Even the garden-variety signal diodes you use in circuits have a small photovoltaic effect (there are amusing stories of bizarre circuit behavior finally traced to this effect). Figure 4.22 shows the good way to convert current to voltage while holding the input strictly at ground. The inverting input is a virtual ground; this is fortunate, because a photovoltaic diode can generate only a few tenths

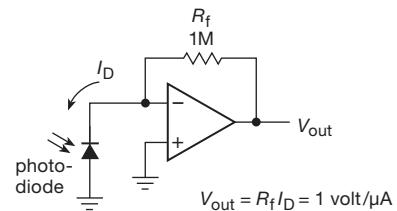


Figure 4.22. Photodiode amplifier.

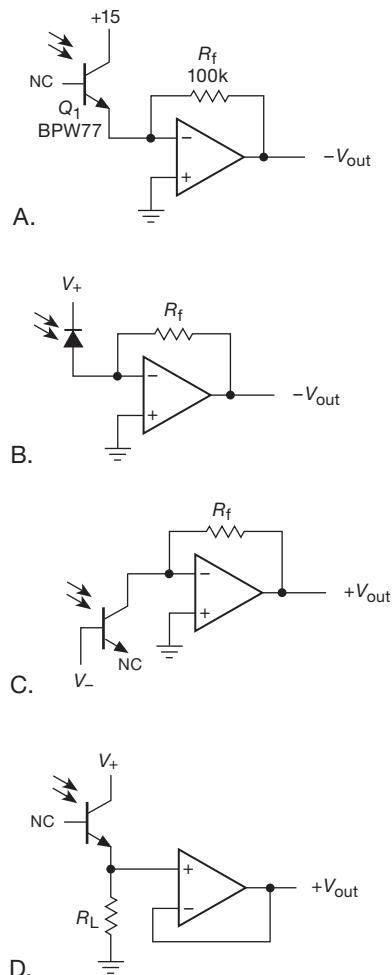
of a volt. This particular circuit has an output of 1 volt per microamp of input current. (With BJT-input op-amps you sometimes see a resistor connected between the noninverting input and ground; its function will be explained shortly in connection with op-amp shortcomings.)

Of course, this *transresistance* configuration can be used equally well for devices that source their current from some positive excitation voltage, such as  $V_{CC}$ . Photodiodes and phototransistors (both devices that source current from a positive supply when exposed to light) are often used this way (Figure 4.23). The photodiode has lower photocurrent, but excels in linearity and speed; very fast photodiodes can operate at *gigahertz* speeds. By contrast, the phototransistor has a considerably higher photocurrent (owing to transistor beta, which boosts the native collector-to-base photocurrent), with poorer linearity and speed. You can even get photo-Darlingtons, which extend this trend.

In real-world applications it is usually necessary to include a small capacitor across the feedback resistor, to ensure stability (i.e., prevent oscillation or ringing). This is because the capacitance of the detector, in combination with the feedback resistor, forms a lowpass filter; the resulting lagging phase shift at high frequencies, combined with the op-amp's own lagging phase shift (see §4.9.3), can add up to  $180^\circ$ , thus producing overall *positive* feedback, and thus oscillation. We treat this interesting problem in some detail in Chapter 4x (“Transresistance amplifiers”); be sure to read that section carefully if you are building amplifiers for photodiodes. (And analogous stability problems occur, for similar reasons, when you drive capacitive loads with op-amps; see §4.6.1B).

**Exercise 4.5.** Use a 411 and a 1 mA (full scale) meter to construct a “perfect” current meter (i.e., one with zero input impedance) with 5 mA full scale. Design the circuit so that the meter will never be driven more than  $\pm 150\%$  full scale. Assume that the 411 output can swing to  $\pm 13$  volts ( $\pm 15$  V supplies) and that the meter has  $500\ \Omega$  internal resistance.

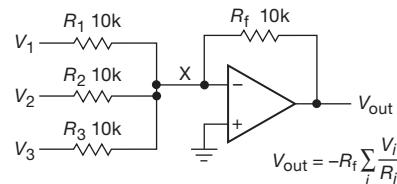
<sup>9</sup> You might be tempted to reduce the input coupling capacitor since its load has been bootstrapped to high impedance. However, this can generate a peak in the frequency response, in the manner of an active filter (see §6.3).



**Figure 4.23.** Photodiode amplifiers with reverse bias: A. Phototransistor; note base terminal is not used. B. Photodiode. C. Phototransistor used as photodiode; for variety we show it current sinking. D. Phototransistor with load resistor driving voltage follower.

#### D. Summing amplifier

The circuit shown in Figure 4.24 is just a variation of the inverting amplifier. Point X is a virtual ground, so the input current is  $V_1/R_1 + V_2/R_2 + V_3/R_3$ . With equal resistor values you get  $V_{\text{out}} = -(V_1 + V_2 + V_3)$ . Note that the inputs can be positive or negative. Also, the input resistors need not be equal; if they're unequal, you get a weighted sum. For instance, you could have four inputs, each of which is +1 volt or zero, representing binary values 1, 2, 4, and 8. By using input resistors of 10k, 5k, 2.5k, and 1.25k, you will get a negative output in volts equal to the binary count input. This scheme can be easily expanded to several dig-



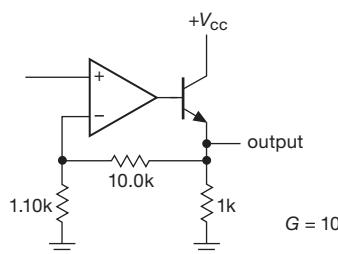
**Figure 4.24.** Summing amplifier.

its. It is the basis of digital-to-analog conversion, although a different input circuit (an  $R$ - $2R$  ladder) is usually used.

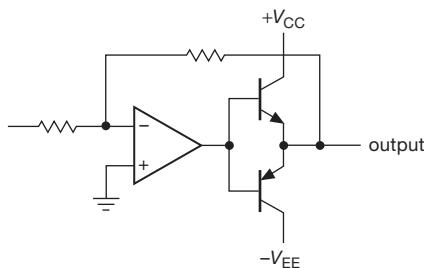
**Exercise 4.6.** Show how to make a two-digit digital-to-analog converter (DAC) by appropriately scaling the input resistors in a summing amplifier. The digital input represents two digits, each consisting of four lines that represent the values 1, 2, 4, and 8 for the respective digits. An input line is either at +1 volt or at ground, i.e., the eight input lines represent 1, 2, 4, 8, 10, 20, 40, and 80. With  $\pm 15$  V supplies, the op-amp's outputs generally cannot swing beyond  $\pm 13$  volts; you will have to settle for an output in volts equal to one-tenth the value of the input number.

#### E. Power booster

For high output current, a power transistor follower can be hung on an op-amp output (Figure 4.25). In this case a non-inverting amplifier has been drawn, though a follower can be added to any op-amp configuration. Notice that feedback is taken from the emitter; thus feedback enforces the desired output voltage in spite of the  $V_{BE}$  drop. This circuit has the usual problem that the follower output can only source current. As with transistor circuits, the remedy is a push-pull booster (Figure 4.26). We'll see later that the limited speed with which the op-amp can move its output (slew rate) seriously limits the speed of this booster in the crossover region, creating distortion. For slow-speed applications you don't need to bias the push-pull pair into quiescent conduction, because feedback will take care of most of the crossover distortion. Complete power booster ICs are available, e.g. the LT1010 and BUF633/4. These



**Figure 4.25.** Single-ended emitter follower boosts op-amp output current (sourcing only).



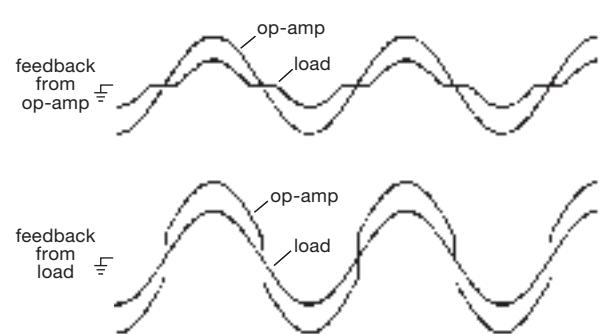
**Figure 4.26.** Push-pull follower boosts op-amp output current, both sourcing and sinking. You commonly see a small resistor ( $\sim 100\Omega$ ) connected between the bases and emitters to reduce crossover nonlinearity by maintaining feedback throughout the signal swing. See Figure 2.71 for improved output-stage biasing.

are unity-gain push-pull amplifiers capable of 200 mA of output current, and operation to 20–100 MHz (see §5.8.4, and also the discussion (and table) of unity-gain buffers in Chapter 4x.); they are carefully biased for low open-loop crossover distortion, and include on-chip protection (current limit, and often thermal shutdown as well). As long as you ensure that the op-amp driving them has significantly less bandwidth, you can include them inside the feedback loop without any worries.<sup>10</sup>

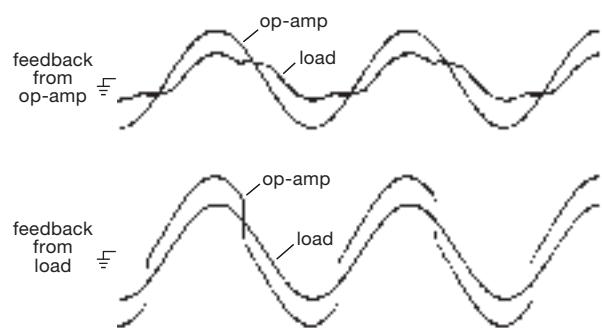
#### Feedback and the push-pull booster

The push-pull booster circuit illustrates nicely the linearizing effect of negative feedback. We hooked up an LF411 op-amp as a noninverting unity-gain follower, driving a BJT push-pull output stage, and we loaded the output with a  $10\Omega$  resistor to ground. Figure 4.27 shows the output signals at the op-amp and at the load, with an input sinewave of 1 V amplitude at 125 Hz. For the upper pair of traces we (foolishly) took the feedback from the op-amp's output, which produced a fine replica of the input signal; but the load sees severe crossover distortion (from the  $2V_{BE}$  dead zone). With the feedback coming from the push-pull output (where the load is connected) we get what we want, as seen in the lower pair of traces. The op-amp cleverly creates an exaggerated waveform to drive the push-pull follower, with precisely the right shape to compensate for the crossover.

Figure 4.28 shows what these waveforms look like when we try driving an actual loudspeaker, a load that is more complicated than a resistor (because it's both a “motor” and a “generator,” it exhibits resonances and other nasty properties; it's also got a reactive crossover network, and



**Figure 4.27.** Feedback cures crossover distortion in the push-pull follower. Vertical: 1 V/div; horizontal: 2 ms/div.



**Figure 4.28.** Same as Figure 4.27, but loaded with a loudspeaker of  $6\Omega$  nominal impedance.

an inductive coil to propel the cone). Once again, the magic of feedback does the job, this time with an op-amp output that is charmingly unsymmetrical.<sup>11</sup>

#### F. Power supply

An op-amp can provide the gain for a feedback voltage regulator (Figure 4.29). The op-amp compares a sample of the output with the zener reference, changing the drive to the Darlington “pass transistor” as needed. This circuit supplies a stable 10 volt output (“regulated”), at up to 1 amp load current. Some notes about this circuit:

<sup>11</sup> We should note, in fairness, that the fine performance seen here is at a rather low frequency (we chose it close to the speaker's bass resonance, to illustrate how clever feedback can be). But the situation degrades at high frequencies, owing to finite slew rate and falling loop gain (topics we'll see in §4.4). It's far better to eliminate most crossover distortion in the push-pull stage itself, by proper “class-AB” biasing (see Figure 2.71 in §2.4.1A), or by using an external unity-gain buffer (see Figure 4.87, and §5.8.4); then using feedback to suppress any residual distortion.

<sup>10</sup> But beware a common error: a working circuit is upgraded by substituting a faster op-amp, whereupon the “improved” circuit oscillates!

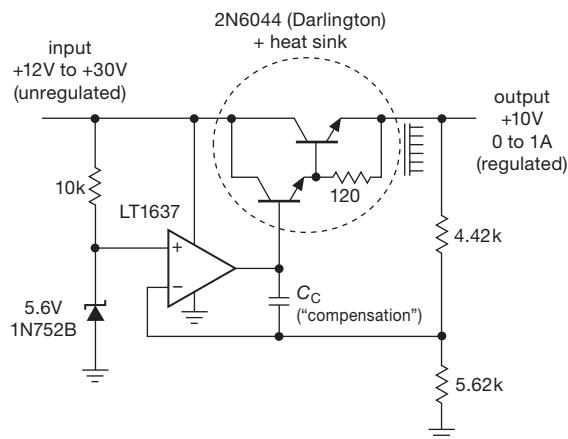


Figure 4.29. Voltage regulator.

- The voltage divider that samples the output could be a potentiometer, for adjustable output voltage.
- For reduced ripple at the zener, the 10k resistor should be replaced with a current source. Another approach is to bias the zener from the output; that way you take advantage of the regulator you have built. *Caution:* when using this trick, you must analyze the circuit carefully to be sure it will start up when power is first applied.
- We used a rail-to-rail op-amp, which can swing its output to the positive rail,<sup>12</sup> so that the input voltage can go as low as +12 V without putting the Darlington pass transistor into saturation. With a 411, by contrast, you would have to allow another 1.5–2 V of margin, because the op-amp's output cannot get closer than that to the positive supply rail.
- The circuit as drawn could be damaged by a temporary short circuit across the output, because the op-amp would attempt to drive the Darlington pair into heavy conduction. Regulated power supplies should always have circuitry to limit “fault” current (see §9.1.1C for more details).
- Without the “compensation capacitor”  $C_C$  the circuit would likely oscillate when the dc output is bypassed (as it would be when powering a circuit) because of the additional lagging phase shift. Capacitor  $C_C$  ensures stability into a capacitive load, a subject we'll visit in §§4.6.1B, 4.6.2, and 9.1.1C.
- Integrated circuit voltage regulators are available in tremendous variety, from the time-honored 723 to the

<sup>12</sup> Our suggested LT1637 is a 44-volt “over-the-top” op-amp that exhibits strikingly higher input-bias currents when its input is near the positive rail (as much as  $I_B=20\mu A$ , about 100 times its normal bias current). The LT1677, with  $I_B=0.2\mu A$ , might be a better choice.

convenient 3-terminal adjustable regulators with internal current limit and thermal shutdown (see §9.3). These devices, complete with temperature-compensated internal voltage reference and pass transistor, are so easy to use that you will almost never use a general-purpose op-amp as a regulator. The exception might be to generate a stable voltage within a circuit that already has a stable power-supply voltage available.

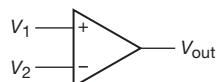
In Chapter 9 we discuss voltage regulators and power supplies in detail, including special ICs intended for use as voltage regulators.

### 4.3.2 Nonlinear circuits

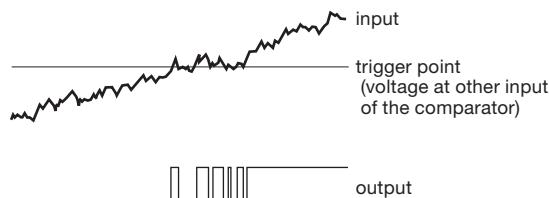
#### A. Comparator – an introduction

It is quite common to want to know which of two signals is larger, or to know when a given input signal exceeds a predetermined voltage. For instance, the usual method of generating triangle waves is to supply positive or negative currents into a capacitor, reversing the polarity of the current when the amplitude reaches a preset peak value. Another example is a digital voltmeter. In order to convert a voltage to a number, the unknown voltage is applied to one input of a comparator, with a linear ramp (capacitor + current source) applied to the other. A digital counter counts cycles of an oscillator during the time that the ramp is less than the unknown voltage and displays the result when equality of amplitudes is reached. The resultant count is proportional to the input voltage. This is called single-slope integration; in most sophisticated instruments a dual-slope integration is used (Chapter 13).

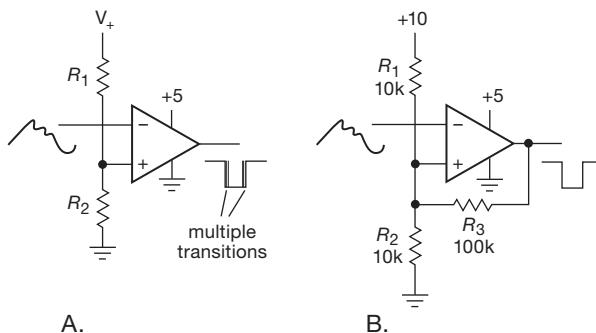
The simplest form of comparator is a high-gain differential amplifier, made either with transistors or with an op-amp (Figure 4.30). In this circuit there's no feedback – the op-amp goes into positive or negative saturation according to the difference of the input voltages. Because of the enormous voltage gain of op-amps (typically  $10^5$ – $10^6$ ), the inputs will have to be equal to within a fraction of a millivolt in order for the output not to be saturated. Although an ordinary op-amp can be used as a comparator (and frequently is), there are special ICs intended for use as comparators. They let you set the output voltage levels independently of the voltages used to power the comparator (e.g., you can have output levels of 0 V and +5 V from a comparator powered from  $\pm 15$  V); and they are generally much faster, because they are not trying to be op-amps, i.e., linear amplifiers intended for use with negative feedback. We'll talk about them in detail in Chapter 12 (§§12.1.7 and 12.3, and Table 12.2.).



**Figure 4.30.** Comparator: an op-amp without feedback.



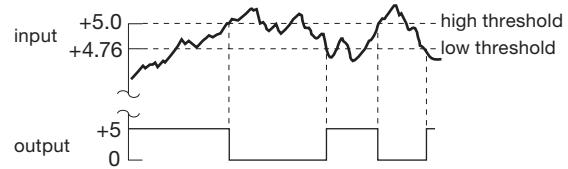
**Figure 4.31.** Comparator without hysteresis produces multiple transitions from noisy input signal.



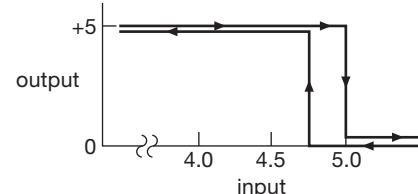
**Figure 4.32.** Positive feedback prevents multiple comparator transitions. A. Comparator without feedback. B. Schmitt trigger configuration uses positive feedback to prevent multiple output transitions. Special comparator ICs are generally preferable, and are drawn with the same symbol.

### B. Schmitt trigger

The simple comparator circuit in Figure 4.30 has two disadvantages. For a very slowly varying input, the output swing can be rather slow. Worse still, if the input is noisy, the output may make several transitions as the input passes through the trigger point (Figure 4.31). Both these problems can be remedied by use of *positive* feedback (Figure 4.32). The effect of  $R_3$  is to make the circuit have two thresholds, depending on the output state. In the example shown, the threshold when the output is at ground (input high) is 4.76 volts, whereas the threshold with the output at +5 volts is 5.0 volts. A noisy input is less likely to produce multiple triggering (Figure 4.33). Furthermore, the positive feedback ensures a rapid output transition, regardless of the speed of the input waveform. (A small “speed-up” capacitor of 10–100 pF is often connected across  $R_3$  to enhance switching speed still further.) This configuration is known



**Figure 4.33.** Hysteresis tames noise-prone comparator.



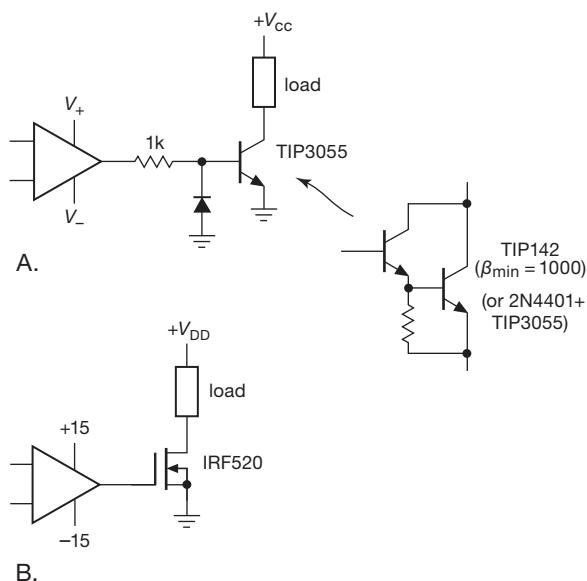
**Figure 4.34.** Output versus input (“transfer function”) for Schmitt trigger.

as a Schmitt trigger, a function that we saw earlier in a discrete transistor implementation (Figure 2.13).

The output depends both on the input voltage and on its recent history, an effect called *hysteresis*. This can be illustrated with a diagram of output versus input, as in Figure 4.34. The design procedure is easy for Schmitt triggers that have a small amount of hysteresis. Use the circuit of Figure 4.32B. First choose a resistive divider ( $R_1$ ,  $R_2$ ) to put the threshold at approximately the right voltage; if you want the threshold near ground, just use a single resistor from noninverting input to ground. Next, choose the (positive) feedback resistor  $R_3$  to produce the required hysteresis, noting that the hysteresis equals the output swing, attenuated by a resistive divider formed by  $R_3$  and  $R_1 \parallel R_2$ . Finally, if you are using a comparator with “open-collector” output, you must add an output pullup resistor small enough to ensure a nearly full supply swing, taking account of the loading by  $R_3$  (read about comparator outputs in §12.3, and see Table 12.2). For the case in which you want thresholds symmetrical about ground, connect an offsetting resistor of appropriate value from the noninverting input to the negative supply. You may wish to scale all resistor values to keep the output current and impedance levels within a reasonable range.

### C. Power-switching driver

The output of a comparator or Schmitt trigger switches abruptly between high and low voltages; it’s not a continuous (or “linear”) signal. You might want to use its output to turn a substantial load on or off. Examples might be a relay, laser, or motor.

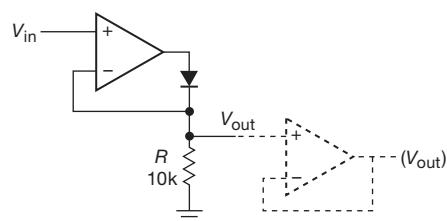


**Figure 4.35.** Power switching with an op-amp; A. With bipolar npn; note base current limit and reverse protection, B. With power MOSFET; note simplified drive circuit.

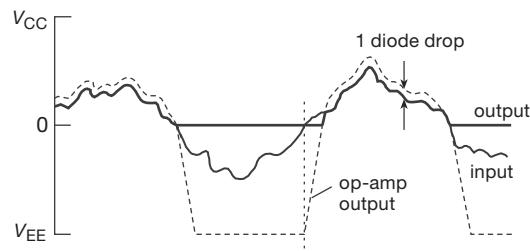
For loads that are either on or off, a switching transistor can be driven from a comparator or op-amp. Figure 4.35A shows how. Note the diode to prevent reverse base-emitter breakdown (op-amps powered from dual supply rails easily swing more than the  $-6\text{ V}$  base-emitter breakdown voltage); it would be omitted if the op-amp's negative supply were no more than  $-5\text{ V}$ . The TIP3055 is a jellybean classic power transistor for noncritical high-current applications, though you'll find plenty of variety of available types with improved maximum voltage, current, power dissipation, and speed (see the listing in Table 2.2 on page 106). A Darlington can be used if currents greater than about 1 amp need to be driven.

In general, however, you're better off using an *n*-channel power MOSFET, in which case you can dispense with the resistor and diode altogether (Figure 4.35B). The IRF520<sup>13</sup> is a near-classic – but the variety of readily available power MOSFETs is overwhelming (see Table 3.4); in general you trade off high breakdown voltage against low ON-resistance.

When switching external loads, don't forget to include a reverse diode if the load is inductive (§1.6.7).



**Figure 4.36.** Simple active half-wave rectifier.



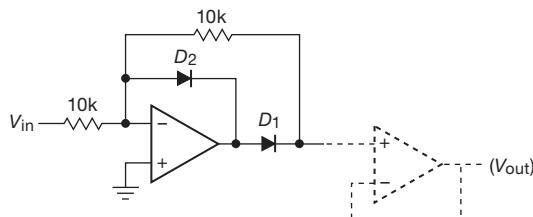
**Figure 4.37.** Effect of finite slew rate on the simple active rectifier.

#### D. Active rectifier

Rectification of signals smaller than a diode drop cannot be done with a simple diode–resistor combination. As usual, op-amps come to the rescue, in this case by putting a diode in the feedback loop (Figure 4.36). For  $V_{in}$  positive, the diode provides negative feedback; the circuit's output follows the input, coupled by the diode, but without a  $V_{BE}$  drop. For  $V_{in}$  negative, the op-amp goes into negative saturation and  $V_{out}$  is at ground.  $R$  could be chosen smaller for lower output impedance, with the tradeoff of higher op-amp output current. A better solution is to use an op-amp follower at the output, as shown, to produce very low output impedance regardless of the resistor value.

There is a problem with this circuit that becomes serious with high-speed signals. Because an op-amp cannot swing its output infinitely fast, the recovery from negative saturation (as the input waveform passes through zero from below) takes some time, during which the output is incorrect. It looks something like the curve shown in Figure 4.37. The output (heavy trace) is an accurate rectified version of the input (light trace), except for a short time interval after the input rises through zero volts. During that interval the op-amp output is racing up from saturation near  $-V_{EE}$ , so the circuit's output is still at ground. A general-purpose op-amp like the 411 has a *slew rate* (maximum rate at which the output can change) of  $15\text{ V}/\mu\text{s}$ ; recovery from negative saturation therefore takes about  $1\ \mu\text{s}$  (when operating from  $\pm 15\text{ V}$  supplies), which may introduce significant output

<sup>13</sup> Along with its higher-current cousins, the IRF530 and IRF540 and the higher-voltage relatives (IRF620–640 and IRF720–740) that fill out the orderly family tree; see “A 30-year MOSFET saga” (§3x.11).



**Figure 4.38.** Improved active half-wave rectifier.

error for fast signals. A circuit modification improves the situation considerably (Figure 4.38).

$D_1$  makes the circuit a unity-gain inverter for negative input signals.  $D_2$  clamps the op-amp's output at one diode drop below ground for positive inputs, and since  $D_1$  is then back-biased,  $V_{\text{out}}$  sits at ground. The improvement comes because the op-amp's output swings only two diode drops as the input signal passes through zero. Because the op-amp output has to slew only about 1.2 volts instead of  $V_{\text{EE}}$  volts, the “glitch” at zero crossings is reduced more than 10-fold. This rectifier is inverting, incidentally. If you require a noninverted output, attach a unity-gain inverter to the output.

The performance of these circuits is improved if you choose an op-amp with a high slew rate. Slew rate also influences the performance of the other op-amp applications we've discussed, for instance the simple voltage amplifier circuits. Shortly we'll take a closer look at the ways in which real op-amps depart from the ideal – input current, offset voltage, bandwidth and slew rate, and so on – because you need to know about those limitations if you want to design good circuits. With that knowledge we'll also look at some active *full-wave* rectifier circuits to complement these half-wave rectifiers.<sup>14</sup> First, though, we'd like to demonstrate some of the fun of designing with op-amps by showing a few real-world circuit examples.

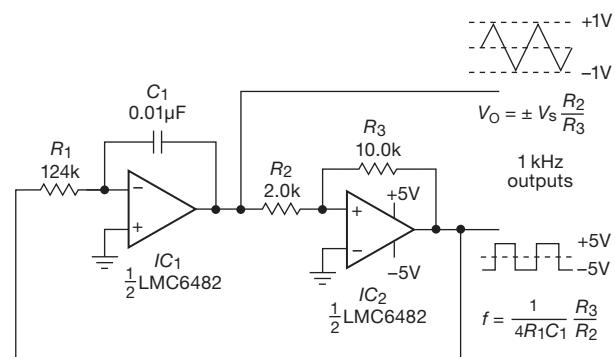
### 4.3.3 Op-amp application: triangle-wave oscillator

These op-amp circuit fragments that we've been exploring – amplifiers, integrators, Schmitt triggers, etc. – are interesting enough; but the real excitement in circuit design comes when you creatively put pieces together to make a complete “something.” A nice example that we can handle now is a triangle-wave oscillator. Unlike any other circuits so far, this one has no input signal; instead it creates an output signal, in this case a symmetrical triangle wave of

1 volt amplitude. As a by-product you also get a square wave, for free. (We'll see many more examples of oscillators in Chapter 7).

The idea is first to use an integrator (with a constant-dc input voltage) to generate a ramp; we need to turn the ramp around when it reaches its  $\pm 1$  V limits, so we let the integrator output (the ramp) drive a Schmitt trigger, with thresholds at  $\pm 1$  V. The output of the Schmitt, then, is what ought to determine the direction of the ramp. Aha! Just use its output (which switches between the supply rail voltages) as the input to the integrator.

Figure 4.39 shows a circuit implementation. It's easiest to start with  $IC_2$ , which is wired as a *noninverting* Schmitt trigger (it looks like an inverting amplifier, but it's not – note that feedback goes to the non-inverting input), for a reason we'll see soon. This configuration is used less frequently than the conventional inverting circuit of Figure 4.32B, because of its lower input impedance (and substantial input current reversal at threshold). Importantly, the LMC6482 has rail-to-rail output swing, so with  $\pm 5$  V power supplies its thresholds are at  $\pm 1$  V, set by a 5:1 ratio of  $R_3$  to  $R_2$ .



**Figure 4.39.** Triangle-wave oscillator.

The Schmitt's  $\pm 5$  V output is the input to the integrator  $IC_1$ . We chose  $C_1$  to be a convenient value of  $0.01 \mu\text{F}$ , then calculated  $R_1$  to ramp through 2 V in a half period (0.5 ms), using  $5\text{V}/R_1 = I_{\text{in}} = C_1[dV/dt]_{\text{ramp}}$ . The calculated resistor value of  $125 \text{ k}\Omega$  (in the figure we show the nearest standard 1% “E96” resistor value; see Appendix C) came out reasonable, given real-world op-amp characteristics, as we'll learn later in the chapter. If it hadn't, we would have changed  $C_1$ ; this is typically how you get to your final circuit component values.

**Exercise 4.7.** Confirm that value of  $R_1$  is correct, and that the Schmitt trigger thresholds are at  $\pm 1$  V.

<sup>14</sup> And in Chapter 4x we'll see additional non-linear circuit applications of op-amps, for example a logarithmic amplifier and a “chaotic” Lorenz-attractor circuit.

Now the reason for connecting IC<sub>2</sub> as a *noninverting* Schmitt trigger becomes clear: if IC<sub>2</sub>'s output is at  $-5\text{ V}$ , say, then the triangle wave is ramping upward toward the Schmitt's  $+1\text{ V}$  threshold, at which point the Schmitt's output will switch to  $+5\text{ V}$ , reversing the cycle. If we had instead used the more conventional inverting Schmitt configuration, the oscillator would not oscillate; in that case it would "latch up" at one limit, as you can verify by walking through one cycle of operation.

The expressions for output frequency and amplitude are shown in the figure. It's interesting to note that the frequency is independent of supply voltage; but if you alter the resistor ratio  $R_2/R_3$  to change the output amplitude, you will also change the frequency. Sometimes it is good to develop algebraic expressions for circuit operation, to see such dependencies. Here's how it goes in this case:

$$\frac{dV}{dt} = \frac{I}{C} = \frac{V_S/R_1}{C_1},$$

$$\text{so } \Delta t = C_1 \frac{R_1}{V_S} \Delta V,$$

$$\text{but } \Delta V = 2 \frac{R_2}{R_3} V_S,$$

$$\text{so } \Delta t = 2 C_1 R_1 \frac{R_2}{R_3},$$

$$\text{and so, finally, } f = \frac{1}{2\Delta t} = \frac{1}{4R_1 C_1 R_2}. \quad (4.4)$$

Note how  $V_S$  cancelled in the fourth step, leading to an output frequency independent of supply voltage.

A warning: it's easy to be dazzled by the apparent power of mathematics and quickly to fall in love with "algebraic circuit design." Our stern advice in this matter (and you can quote us on it) is:

*Resist the temptation to take refuge in equations as a substitute for understanding how a circuit really works.*

#### 4.3.4 Op-amp application: pinch-off voltage tester

Here's another nice application of op-amps: suppose you want to measure a batch of JFETs in order to put them into groups that are matched in pinch-off voltage  $V_{GS}(\text{off})$  (sometimes called  $V_P$ , see §3.1.3). This is useful because the large spread of specified  $V_P$  sometimes makes it difficult to design a good amplifier.<sup>15</sup> We'll assume that you want to find the gate-source back-bias that results in a drain

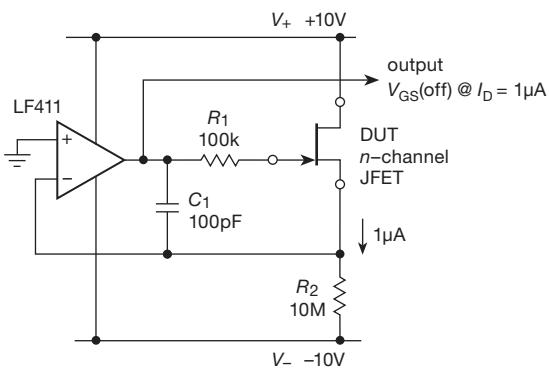


Figure 4.40. Simple pinch-off voltage tester.

current of  $1\text{ }\mu\text{A}$  with the drain at  $+10\text{ V}$  and the source grounded.

If you didn't know about op-amps, you could imagine (a) grounding the source, (b) hooking up a sensitive current meter from the drain to a  $+10\text{ V}$  supply, and then (c) adjusting the gate voltage with a variable negative supply to a value that produces  $1\text{ }\mu\text{A}$  of measured drain current.

Figure 4.40 shows a better way. The device under test (you'll often see the acronym DUT) has its drain tied to  $+10\text{ V}$ ; but the source lead, instead of being grounded, is tied to the inverting input (virtual ground) of an op-amp whose noninverting input is grounded. The op-amp controls the gate voltage, thus holding the source at ground. Because the source is pulled down to  $-10\text{ V}$  through a  $10\text{ M}$  resistor, the source current (and therefore the drain current) is  $1\text{ }\mu\text{A}$ . The op-amp's output is the same as the gate voltage, so the output of this circuit is the pinch-off voltage you wanted to know.

A few details:

- We chose power supply voltages of  $\pm 10\text{ V}$  for the op-amp to make the rest of the circuit simpler, since we wanted to measure  $V_P$  with  $+10\text{ V}$  on the drain. That's OK, because most op-amps work well over a range of supply voltages (in fact, the trend is toward lower operating voltages, driven by the market for battery-powered consumer devices). But if you have only  $\pm 15\text{ V}$  available, you would have to generate  $+10\text{ V}$  within your circuit, either with a voltage divider, a zener, or a 3-terminal voltage regulator (see Chapter 9).
- We put a  $100\text{k}$  resistor ( $R_1$ ) as protection in series with the gate to prevent any significant gate current from flowing during plug-in transients, etc. This can introduce a lagging phase shift around the loop at high frequencies (as can the rather large pull-down resistor  $R_2$ ), so we added a small feedback capacitor  $C_1$  to maintain stability.

<sup>15</sup> You can as well use this same circuit to match the threshold voltage,  $V_{GS\text{th}}$ , of a set of MOSFETs.

We talk about this business of stability toward the end of the chapter, in §4.9.

- For this circuit to work properly, it's important that the op-amp's inverting input not load the source terminal, for example by drawing anything approaching a microamp of current. As we'll learn shortly, this is not always the case. For this example our general-purpose 411 op-amp, with its JFET input transistors, is fine (with input currents in the picoamperes); but an op-amp that uses bipolar transistors for its input stage would generally have input currents in the 10's to 100's of nanoamps, and should be avoided for a low-current application like this.
- The drain current at which pinch-off voltage is specified is not always  $1\ \mu\text{A}$ . You'll see  $V_{GS(\text{off})}$  specified at values of drain current ranging from  $1\ \text{nA}$  to tens of microamps, depending on the size of the JFET, and the whim of the manufacturer. (In an informal survey of datasheets we found  $1\ \text{nA}$  to be the most popular, followed by  $1\ \mu\text{A}$ ,  $10\ \text{nA}$ , and  $0.5\ \text{nA}$ , with five other values used occasionally.) It would be easy to modify the circuit to accommodate higher test currents; but to go to  $10\ \text{nA}$ , say, you would need a  $1\ \text{G}\Omega$  resistor for  $R_2$ ! In that case a better solution is to return the pull-down resistor to a lower voltage, say  $-0.1\ \text{V}$ , which you could generate with a voltage divider from the  $-10\ \text{V}$  negative supply. You'd have to worry again about op-amp input currents with such a small test current.

**Exercise 4.8.** Show how to make the pinch-off tester operate from  $\pm 15\ \text{V}$  supplies, with the measurement still made at  $V_D = +10\ \text{V}$ ; assume that the largest resistor value available is  $10\ \text{M}\Omega$ .

**Exercise 4.9.** Modify the pinch-off tester circuit of Figure 4.40 so that you can measure  $V_{GS}$  at three values of drain current, namely  $1\ \mu\text{A}$ ,  $10\ \mu\text{A}$ , and  $100\ \mu\text{A}$ , by setting a 3-position switch. Assume that the largest resistor value you can conveniently get is  $10\ \text{M}\Omega$ .

**Exercise 4.10.** Now change the circuit so that it measures  $V_{GS(\text{off})}$  at  $I_D=1\ \text{nA}$ . Assume you can get  $100\ \text{M}\Omega$  5% resistors.

### 4.3.5 Programmable pulse-width generator

When triggered by a short input pulse, the circuit in Figure 4.41 generates an output pulse<sup>16</sup> whose width is set by 10-turn pot  $R_1$ . Here's how it works.

$\text{IC}_1$ ,  $\text{IC}_2$ , and  $Q_1$  form a current source that charges timing capacitor  $C$ , as we'll detail below.  $\text{IC}_3$  is a versatile timer IC, whose many exploits we will enjoy in Chapter 7. It holds  $C$  discharged (through a saturated MOSFET

switch whose drain drives the  $\text{DIS}$  pin to ground) and simultaneously holds the output at ground, until it receives a negative-going trigger pulse at its  $\text{TRIG}$  input pin; at that point it releases  $\text{DIS}$  and switches its output to  $V_+$ , in this case  $+5\ \text{V}$ .

The current source now charges  $C$  with a positive-going ramp, according to  $I = C dV/dt$ . This continues until the capacitor voltage, which also drives  $\text{IC}_3$ 's  $\text{TH}$  input, reaches a voltage equal to  $2/3$  of the supply voltage,  $V_{\text{TH}} = \frac{2}{3}V_+$ ; at this point  $\text{IC}_3$  abruptly pulls  $\text{DIS}$  back to ground, simultaneously switching its output to ground. This completes the cycle.

The current source is an elegant circuit. We want to source a current into the capacitor, with compliance from ground to at least  $+3.3\ \text{V}$  ( $2/3$  of  $+5\ \text{V}$ ), with linear control by a pot that returns to ground. For reasons we'll see presently, we want the programmed current proportional to the supply voltage  $V_+$ . In this circuit  $Q_1$  is the current source, with  $\text{IC}_2$  controlling its base to hold its emitter at  $+5\ \text{V}$ .  $\text{IC}_1$  is an inverting amplifier referenced to  $+5\ \text{V}$ ; it pivots its output to a voltage that exceeds  $+5\ \text{V}$  by an amount proportional to the current flowing through  $R_1$  and  $R_2$ . That excess voltage appears across  $R$ , generating the output current. You'll know you understand how it works by doing the following problem.

**Exercise 4.11.** Calculate the current sourced by  $Q_1$  by calculating the output voltage of  $\text{IC}_1$  as a function of  $R_X$  (the sum of  $R_1$  and  $R_2$ ),  $R_3$ , and  $V_+$ . Now use it to calculate the output pulse width, knowing that  $\text{IC}_3$  switches when the voltage at  $\text{TH}$  reaches  $\frac{2}{3}V_+$ .

This circuit is an illustration of the use of *ratiometric* techniques: for a given setting of  $R_1$ , both the capacitor charging current  $I$  and the timer IC threshold voltage  $V_{\text{TH}}$  individually depend on supply voltage  $V_+$ ; but their variation is such that the final pulse width  $T$  does not depend on  $V_+$ . That is why the current source was designed with  $I \propto V_+$ . The use of ratiometric techniques is an elegant way to design circuits with excellent performance, often without requiring precise control of power-supply voltages.

### 4.3.6 Active lowpass filter

The simple  $RC$  filters we saw back in Chapter 1 have a soft rolloff; that is, their response versus frequency does not progress sharply from a passband to a stopband. Perhaps surprisingly, this behavior cannot be remedied by simply cascading multiple stages, as we'll see in detail in Chapter 6 (and particularly in connection with *active filters*, §6.3). Much better filter performance can be achieved if

<sup>16</sup> More about pulsers, for those interested, in §§7.1.4B and 7.2.

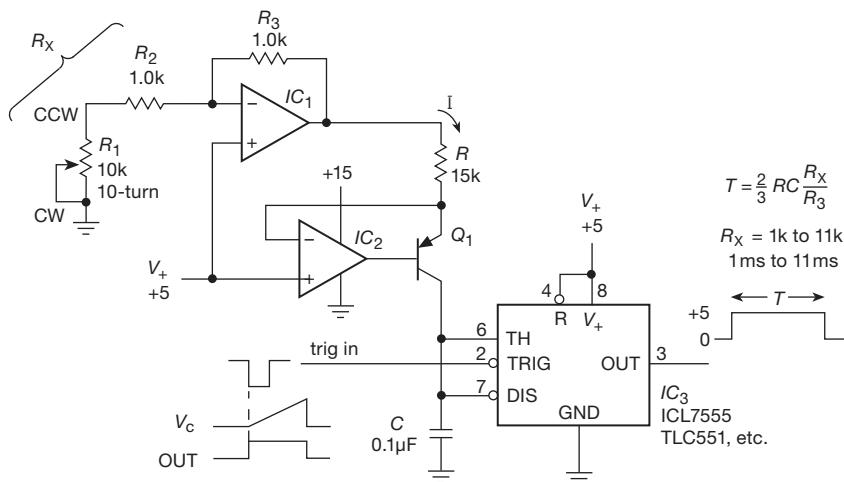
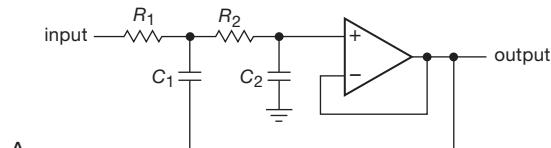


Figure 4.41. Pulse generator with programmable width.

you include both inductors and capacitors, or, equivalently, if you “activate” the filter design by using op-amps.

Figure 4.42 shows an example of a simple and even partly intuitive filter. This configuration is known as a Sallen-and-Key filter, after its inventors. The unity-gain amplifier can be an op-amp connected as a follower, or a unity-gain *buffer* IC, or just an emitter follower. This particular filter is a second-order lowpass filter. Note that it would be simply a pair of cascaded passive  $RC$  lowpass filters, except for the fact that the bottom of the first capacitor is bootstrapped by the output. It is easy to see that at high frequencies (well beyond  $f = 1/2\pi RC$ ) it falls off just like a cascaded  $RC$ , i.e., at  $-12 \text{ dB/octave}$ , because the output is essentially zero (and therefore the first capacitor’s lower end is effectively grounded). As we lower the frequency and approach the passband, however, the bootstrap action tends to reduce the attenuation, thus giving a sharper “knee” to the curve of response versus frequency. We’ve plotted the response versus frequency, with three “tunings” of the  $R$  and  $C$  values.<sup>17</sup>

Of course, such hand-waving cannot substitute for honest analysis, which luckily has been done for a prodigious variety of nice filters. And contemporary general-purpose SPICE-based analog simulation tools, or special filter analysis software, let you design and view filter response curves with relative ease.



A.

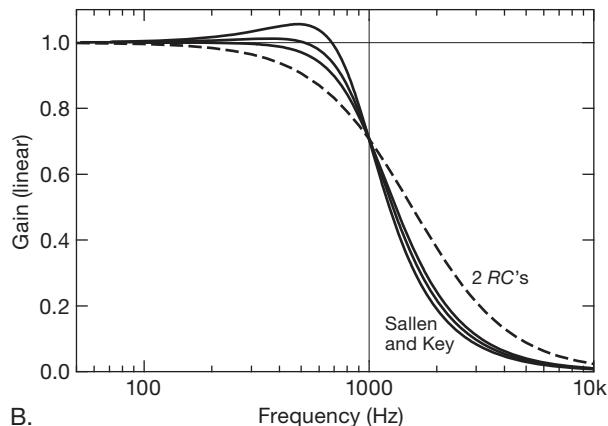


Figure 4.42. Sallen-and-Key active lowpass filter: A. Schematic; B. frequency response, compared with a cascade of two passive  $RC$  sections.

#### 4.4 A detailed look at op-amp behavior

We’ve hinted that op-amps aren’t perfect, and that the performance of circuits such as active rectifiers and Schmitt triggers is limited by op-amp speed, or “slew rate.” For those applications a high-speed op-amp is often required.

<sup>17</sup> Butterworth and two Chebyshevs (0.1 dB and 0.5 dB passband ripple), going from flattest to peakiest response; for the Butterworth, for example, the component values are  $C_1=10 \text{ nF}$ ,  $C_2=2 \text{ nF}$ ,  $R_1=12.7 \text{ k}\Omega$ , and  $R_2=100 \text{ k}\Omega$ . Active filters are discussed in detail in Chapter 6.

But slew rate is just one of a half-dozen important parameters of op-amps, which include input offset voltage, input bias current, input common-mode range, noise, bandwidth, output swing, and supply voltage and current. To state the situation fairly, op-amps are remarkable devices, with near-ideal performance for most applications you are likely to encounter. To put it quantitatively, think of the difficulty of designing, with discrete transistors and other components, a high-gain dc differential amplifier that has an input current less than a picoamp, an offset from perfect balance less than a millivolt, a bandwidth of several megahertz, and that operates with its inputs anywhere between the two supply voltages. You can get such an op-amp for a dollar; it comes in a tiny package measuring  $1.5\text{ mm} \times 3\text{ mm}$ , and it draws less than a milliamp.

But op-amps *do* have performance limitations – that's why there are literally thousands of available types – and in general you're faced with a tradeoff: you can get much lower bias current (for example), at the expense of offset voltage. A good understanding of op-amp limitations and their influence on circuit design and performance will help you choose your op-amps wisely and design with them effectively.

To motivate the subject, imagine that you've been asked to design a dc amplifier, so that small voltages (0–10 mV) can be seen on a handsome analog meter scale. And it should have at least  $10M\Omega$  input resistance, and be accurate to 1% or so. No problem, you say... I'll just use the non-inverting amplifier configuration (to get high input resistance), with lots of gain ( $\times 1000$ , say, so 10 mV is amplified to 10 V). Speed is not an issue, so you don't worry about slew rate. With supreme confidence you draw up the circuit (with an LF411 op-amp), your technician builds it, and ... your boss fires you! The thing was a disaster: it read 20% of full-scale with no input attached, and it drifted like crazy when carried outside. It does work OK – as a *paperweight*.<sup>18</sup>

To get started, look at Figure 4.43, a simplified schematic of the LF411. Its circuit is relatively straightforward, in terms of the kinds of transistor circuits we discussed in the last two chapters. It has a JFET differential input stage, with current-mirror active load, buffered with an *n*p*n* follower (to prevent loading of the high-gain input stage) driving a grounded-emitter *n*p*n* stage (with current-source active load). This drives the push–pull emitter follower output stage ( $Q_7Q_8$ ), with current-limiting circuitry

( $R_5Q_9$  and  $R_6Q_{10}$ ) to protect against output short-circuit.<sup>19</sup> The curious feedback capacitor  $C_C$  ensures stability; we'll learn about it later. This circuit displays the internal circuitry characteristic of the typical op-amp, and from it we can see how and why op-amp performance departs from ideal.

**Exercise 4.12.** Explain how the current-limiting circuitry in Figure 4.43 works. What is the maximum output current?

**Exercise 4.13.** Explain the function of the two diodes in the output stage.

Let's look at these problems, what the consequences are for circuit design, and what to do about it.

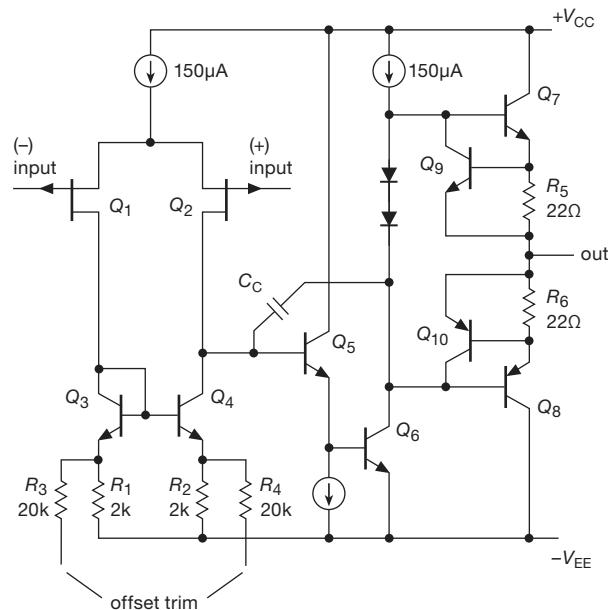


Figure 4.43. Simplified schematic of the LF411 op-amp.

#### 4.4.1 Departure from ideal op-amp performance

The ideal op-amp has these characteristics:

- Input current = 0 (input impedance =  $\infty$ ).
- $V_{out} = 0$  when both inputs are at precisely the same voltage (zero “offset voltage”).
- Output impedance (open loop) = 0.
- Voltage gain =  $\infty$ .
- Common-mode voltage gain = 0.

<sup>18</sup> We'll revisit this example in §4.4.3, and again in more detail in Chapter 5.

<sup>19</sup> The LF411's *detailed* schematic reveals a more elaborate negative current-limit configuration; check it out on the datasheet, to see if you can understand how it works.

- Output can change instantaneously (infinite slew rate).
- Absence of added “noise.”

All of these characteristics should be independent of temperature and supply voltage changes.

In the following paragraphs we describe how real op-amps depart from these ideals. As you struggle through the fact-filled sections, you may want to refer to Table 4.1 to maintain perspective. Tables 4.2a,b, 5.5, and 8.3 may be helpful also, for seeing some actual numbers. And we’ll revisit these in more detail in Chapter 5 (§§5.7 and 5.8) in connection with the design of precision circuits.

### A. Input offset voltage

Op-amps don’t have perfectly balanced input stages, owing to manufacturing variations. The problem is worse with FETs, with their poorer matching of input thresholds. If you connect the two op-amp inputs together to create exactly zero differential input signal, the output will usually saturate at either  $V_+$  or  $V_-$  (you can’t predict which). The difference in input voltages necessary to bring the output to zero is called the input offset voltage,  $V_{OS}$  (it’s as if there were a battery of that voltage in series with one of the inputs). Typical offset voltages are around 1 mV, but “precision” op-amps can have offset voltages as small as 10  $\mu$ V. Some op-amps make provision for trimming the input offset voltage to zero. For a 411 you attach a 10k pot between pins 1 and 5 (“offset trim” in Figure 4.43), with the wiper connected to  $V_{EE}$ , and adjust for zero offset; the effect is to unbalance deliberately the current mirror to compensate for the offset.

### B. Offset voltage drift

Of greater importance for precision applications is the *drift* of the input offset voltage with temperature and time, since any initial offset could be manually trimmed to zero. A 411 has a typical offset voltage of 0.8 mV (2 mV maximum), with a tempco of  $\Delta V_{OS}/\Delta T = 7 \mu\text{V}/^\circ\text{C}$  and unspecified coefficient of offset drift with time. The OP177A, a precision op-amp, is laser-trimmed for a maximum offset of 10 microvolts, with a temperature coefficient of 0.1  $\mu\text{V}/^\circ\text{C}$  (max) and long-term drift of 0.2  $\mu\text{V}/\text{month}$  (typical) – roughly a hundred times better in both offset and tempco.

### C. Input current

The input terminals sink (or source, depending on the op-amp type) a small current called the input bias current,  $I_B$ , which is defined as half the sum of the input currents with the inputs tied together (the two input currents are approximately equal and are simply the base or gate currents of

the input transistors). For the JFET-input 411 the bias current is typically 50 pA (200 pA max) at room temperature (but as much as 4 nA at 70°C), while a typical BJT-input op-amp like the OP27 has a bias current of 15 nA, varying little with temperature. As a rough guide, BJT-input op-amps have bias currents in the tens of nanoamps, whereas JFET-input op-amps have input currents in the tens of picoamps (i.e., 1000 times lower), and MOSFET-input op-amps have input currents of typically a picoamp or less. Generally speaking, you can ignore input current with FET op-amps, but not with bipolar-input op-amps.<sup>20</sup>

The significance of input bias current is that it causes a voltage drop across the resistors of the feedback network, bias network, or source impedance. How small a resistor this restricts you to depends on the dc gain of your circuit and how much output variation you can tolerate. For example, an LF412’s maximum input current of 200 pA means that you can tolerate resistances (seen from the input terminals) up to  $\sim 5 \text{ M}\Omega$  before you have to worry about it at the 1 mV level.

We will see more about how this works later. If your circuit is an integrator, bias current produces a slow ramp even when there is no external input current to the integrator.

Op-amps are available with input bias currents down to a nanoamp or less for bipolar-transistor-input circuit types, or down to a fraction of a picoamp ( $10^{-6} \mu\text{A}$ ) for MOSFET-input circuit types. The very lowest bias currents are typified by the BJT-input LT1012, with a typical input current of 25 pA, the JFET-input OPA129, with an input current of 0.03 pA, and the MOSFET LMC6041, with an input current of 0.002 pA. At the other end, very fast BJT op-amps like the THS4011/21 ( $\sim 300 \text{ MHz}$ ) have input currents of 3  $\mu\text{A}$ . In general, BJT op-amps intended for high-speed operation have higher bias currents.

### D. Input offset current

Input offset current is a fancy name for the difference of the input currents between the two inputs. Unlike input bias current, the offset current,  $I_{OS}$ , is a result of manufacturing variations, since an op-amp’s symmetrical input circuit would otherwise result in identical bias currents at the two inputs. The significance is that, even when it is driven by identical source impedances, the op-amp will see unequal voltage drops and hence a difference voltage between its inputs. You will see shortly how this influences design.

Typically, the offset current is somewhere between one-

<sup>20</sup> There’s a nice trick, called *bias-current cancellation*, exploited in some BJT op-amps to achieve input currents as low as 10s of picoamps. Look back to Figure 2.98; this is discussed further in Chapter 4x.

**Table 4.1 Op-amp Parameters<sup>a</sup>**

Parameter	bipolar (BJT)		JFET-input		CMOS		Units
	jellybean	premium	jellybean	premium	jellybean	premium	
$V_{os}$ (max)	3	0.025	2	0.1	2	0.1	mV
$TCV_{os}$ (max)	5	0.1	20	1	10	3	$\mu\text{V}/^\circ\text{C}$
$I_B$ (typ)	50nA	25pA	50pA	40fA	1pA	2fA	@ 25°C
$e_n$ (typ)	10	1	20	3	30	7	nV/ $\sqrt{\text{Hz}}$ @ 1kHz
$f_T$ (typ)	2	2000	5	400	2	10	MHz
$SR$ (typ)	2	4000	15	300	5	10	V/ $\mu\text{s}$
$V_s$ (min) <sup>b</sup>	5	1.5	10	5	2	1	V
$V_s$ (max) <sup>b</sup>	36	44	36	36	15	15	V

**Notes:** (a) typical and "best" values of important op-amp performance parameters. (b) total supply:  $V_+ - V_-$ .

Typical and "best" values of important op-amp performance parameters. In this chart we list values for run-of-the-mill ("jellybean") parts, and for the best op-amp you can get for each individual parameter. That is, you cannot get a single op-amp that has the combination of excellent performance shown in any of the "premium" columns. In this chart you can clearly see that bipolar op-amps excel in precision, stability, speed, wide supply voltage range, and noise, at the expense of bias current; JFET-input types are intermediate, with CMOS op-amps displaying the lowest bias current.

half and one-tenth the bias current. For the 411,  $I_{offset} = 25 \text{ pA}$ , typical. However, for bias-compensated op-amps (like the OPA177), the specified offset current and bias current are comparable, for reasons we'll see in the advanced Chapter 5.

### E. Input impedance

Input impedance refers to the small-signal<sup>21</sup> differential input resistance (impedance looking into one input, with the other input grounded), which is usually much less than the common-mode resistance (a typical input stage looks like a long-tailed pair with current source). For the FET-input 411 it is about  $10^{12} \Omega$ , whereas for BJT-input op-amps like the LT1013 it is about  $300 \text{ M}\Omega$ . Because of the input bootstrapping effect of negative feedback (it attempts to keep both inputs at the same voltage, thus eliminating most of the differential-input signal),  $Z_{in}$  in practice is raised to very high values and usually is not as important a parameter as input bias current.

### F. Common-mode input range

The inputs to an op-amp must stay within a certain voltage range, typically less than the full supply range, for proper operation. If the inputs go beyond this range, the gain of the op-amp may change drastically, even reversing sign! For a 411 operating from  $\pm 15$  volt supplies, the guaranteed common-mode input range is  $\pm 11$  volts minimum.

<sup>21</sup> Not  $V_{in}/I_{bias}$ !

However, the manufacturer claims that the 411 will operate with common-mode inputs all the way to the positive supply, though performance may be degraded. Bringing either input down to the negative supply voltage causes the amplifier to go berserk, with symptoms like phase reversal<sup>22</sup> and output saturation to the positive supply. From the circuit in Figure 4.43 you can see why the LF411 cannot possibly operate with input voltages to the negative rail, because that would put the source terminals of the input JFET pair below the negative rail, taking them out of the active region. This is discussed further in Chapter 4x, along with some good war stories.

There are many op-amps available with common-mode input ranges down to the negative supply, e.g., the bipolar LT1013 and the CMOS TLC2272 and LMC6082; these are often referred to as "single-supply op-amps" or "ground-sensing op-amps" (see §4.6.3). There are also some op-amps whose common-mode input range includes the positive supply, e.g., the JFET LF356. With the trend toward lower supply voltages for battery-powered equipment, op-amp designers have come up with varieties that accommodate input signals over the full range between supply voltages; these are called rail-to-rail, because supply voltages

<sup>22</sup> The popular and inexpensive (\$0.07 in quantity) LM358 and LM324 single-supply op-amps suffer from input phase reversal for inputs more than 400 mV below the negative rail. Improved replacements like the LT1013 and LT1014 fix this problem (and also an output crossover-distortion problem).

are often called supply *rails*.<sup>23</sup> Examples are the CMOS LMC6482 and TLV2400 series, and the bipolar LM6132, LT1630, and LT6220 series. These have the additional nice feature of being able to swing their outputs all the way to the rails (see the subsection on output swing below). These would seem to be ideal op-amps; however, as we discuss in §§5.7, 5.9, and 5.10, rail-to-rail op-amps typically make compromises that affect other characteristics, notably offset voltage, output impedance, and supply current. There are, in addition, a few (*very few*) op-amps that operate properly for input voltages *above* the positive rail (for example, the “over-the-top” LT1637, listed in Table 4.2a on page 271).

In addition to the *operating* common-mode range, there are maximum allowable input voltages beyond which damage will result. For the 411 they are  $\pm 15$  volts (but not to exceed the negative supply voltage, if it is less).

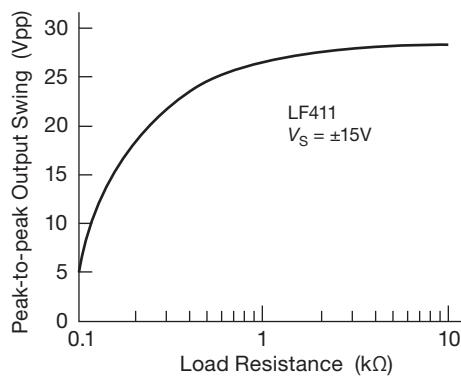
### G. Differential input range

Some bipolar op-amps allow only a limited voltage between the inputs, sometimes as small as  $\pm 0.5$  volt, although most are more forgiving, permitting differential inputs nearly as large as the supply voltages. Exceeding the specified maximum can degrade or destroy the op-amp.

### H. Output swing versus load resistance

The LF411, typical of many op-amps, cannot swing its output closer than a volt or two from either supply rail, even when lightly loaded ( $R_L > 5\text{k}$ , say). That’s because the output stage is a push-pull emitter follower, so even a full rail-to-rail drive to its bases would leave the output a diode drop short of both rails; the drive circuitry has its own difficulties getting close to the rails as well, and the current-limit sense resistors  $R_5$  and  $R_6$  impose an additional voltage drop, which accounts for the shortfall.

For low values of load resistance, the internal current-limit circuit will set the maximum swing. For example, the 411 can swing its output to within about 2 volts of  $V_{CC}$  and  $V_{EE}$  into load resistances greater than about 1k. Load resistances significantly less than that will permit only a small swing. This is frequently shown on datasheets as a graph of peak-to-peak output voltage swing  $V_{om}$  as a function of



**Figure 4.44.** Maximum peak-to-peak output swing versus load (LF411).

load resistance, or sometimes just a few values for typical load resistances. Figure 4.44 shows the datasheet’s graph for the LF411. Many op-amps have asymmetrical output drive capability, with the ability to sink more current than they can source (or vice versa). For that reason you often see maximum output swing plotted, versus load current, as separate curves for output sourcing and sinking current into a load. Figure 4.45 shows such graphs for the LF411.

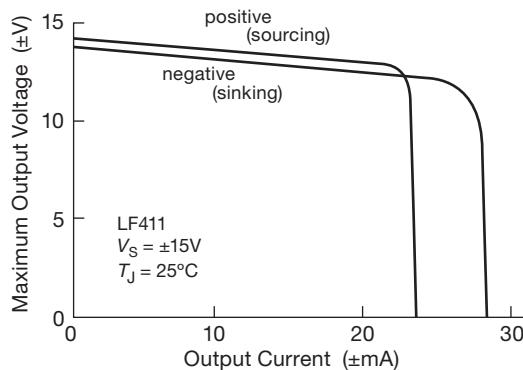
Some op-amps can produce output swings all the way down to the negative supply (e.g., the bipolar LT1013 and CMOS TLC2272), a particularly useful feature for circuits operated from a single positive supply, because output swings all the way to ground are then possible. Finally, op-amps with CMOS transistor outputs in a common-source amplifier configuration<sup>24</sup> (e.g., the LMC6xxx series) can swing all the way to both rails. For such op-amps a much more useful graph plots how close the output can get to each power-supply rail as a function of load current (both sourcing and sinking). An example is shown in Figure 4.46 for the CMOS rail-to-rail LMC6041. Note the effective use of log-log axes, so you can read off accurately the fact that this op-amp can swing to within 1 mV of the rails when supplying  $10\ \mu\text{A}$  of output current, and that its output resistance is approximately  $80\ \Omega$  (sinking) and  $100\ \Omega$  (sourcing). You can find bipolar op-amps that share this property, without the limited supply voltage range of the CMOS op-amps (usually  $\pm 8\text{ V}$  max), for example, the LM6132/42/52 family and the LT1636/7.

### I. Output Impedance

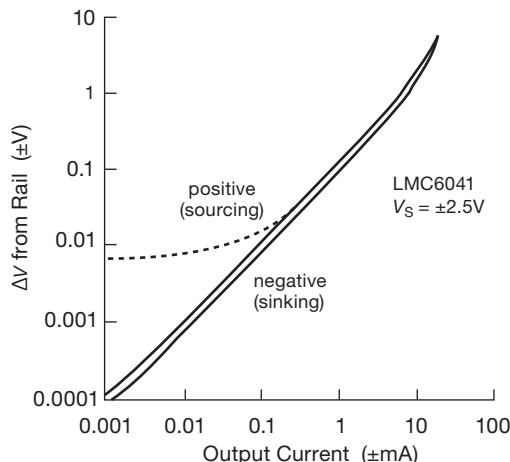
Output impedance  $R_o$  means the op-amp’s intrinsic output impedance *without feedback* (see Figure 2.90). For the 411 it is about  $40\ \Omega$ , but with some low-power op-amps it can

<sup>23</sup> The term “Rail-to-Rail®” is apparently a registered trademark of Nippon Motorola Ltd, though we believe it has been in common use in electronics for decades. This may turn out to be an unwise proprietary claim from their point of view, just as the trademarking of “TRI-STATE®” by National Semiconductor simply drove the industry to adopt the nonproprietary term “3-state” in written references (and, in most cases, to stick with “tristate” in spoken conversation).

<sup>24</sup> Or bipolar-transistor outputs in a common-emitter configuration.



**Figure 4.45.** Maximum output voltage (both sourcing and sinking) versus load current (LF411). The maximum output current capability decreases by  $\sim 25\%$  at  $T_J=125^\circ C$ .

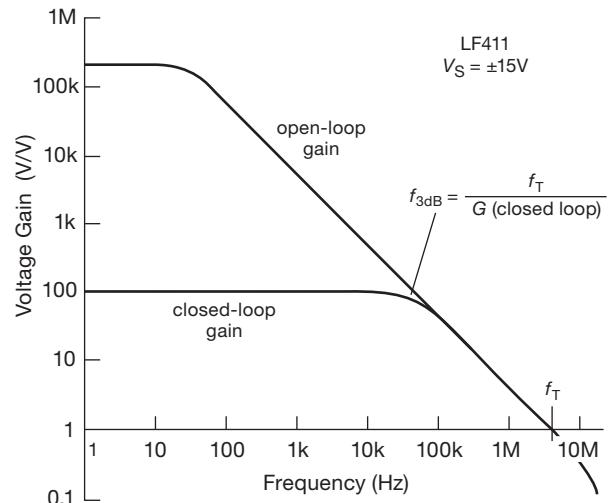


**Figure 4.46.** Maximum swing (as  $\Delta V$  from the respective rails) versus load current for a CMOS rail-to-rail output op-amp. The solid curves are measured values; you can't always trust datasheets – in this case the datasheet's sourcing curve (dashed curve) is evidently in error.

be as high as several thousand ohms, a characteristic shared by some op-amps with rail-to-rail outputs. Feedback lowers the output impedance into insignificance (or raises it, for a current source), by a factor of the loop gain  $AB$  (see §2.5.3C); so what usually matters more is the maximum output current, with typical values of  $\pm 20$  mA or so (but much higher for the rarified group of “high current” op-amps, see Table 4.2b on page 272).

#### J. Voltage gain, bandwidth, and phase shift

Typically the voltage gain  $A_{vo}$  (sometimes called  $A_{VOL}$ ,  $A_V$ ,  $G_V$ , or  $G_{VOL}$ ) at dc is 100,000 to 1,000,000 (often specified

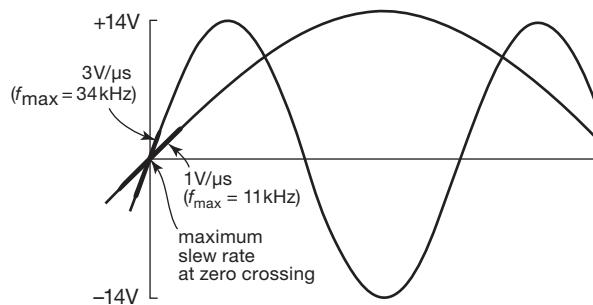


**Figure 4.47.** LF411 gain versus frequency (“Bode plot”).

in decibels, thus 100 dB to 120 dB), dropping to unity gain at a frequency (called  $f_T$ , or sometimes gain-bandwidth product, GBW), most often in the range of 0.1 MHz to 10 MHz. This is usually given as a graph of open-loop voltage gain as a function of frequency, on which the  $f_T$  value is clearly seen; see, for example, Figure 4.47, which shows the curve for our favorite LF411.

For *internally compensated* op-amps this graph is simply a 6 dB/octave rolloff beginning at some fairly low frequency (for the 411 it begins at about 10 Hz), an intentional characteristic necessary for stability, as we’ll see in §4.9. This rolloff (the same as a simple  $RC$  lowpass filter) results in a constant  $90^\circ$  lagging phase shift from input to output (open-loop) at all frequencies above the beginning of the rolloff, increasing to  $120^\circ$  to  $160^\circ$  as the open-loop gain approaches unity. Because a  $180^\circ$  phase shift at a frequency where the voltage gain equals 1 will result in positive feedback (oscillations), the term “phase margin” is used to specify the difference between the phase shift at  $f_T$  and  $180^\circ$ .

There’s a price to pay for greater bandwidth  $f_T$ , namely, higher transistor operating currents, and therefore higher op-amp supply currents. You can get op-amps with supply currents of less than  $1\ \mu A$ , but they have  $f_T$ ’s down around 10 kHz! In addition to high *supply* currents, very fast op-amps can have relatively high input bias currents, often more than a microamp, owing to their bipolar input stages operating at high collector current. Don’t use fast op-amps if you don’t need them – in addition to the drawbacks just mentioned, their high gain at high frequency makes it easier for your circuit to oscillate.



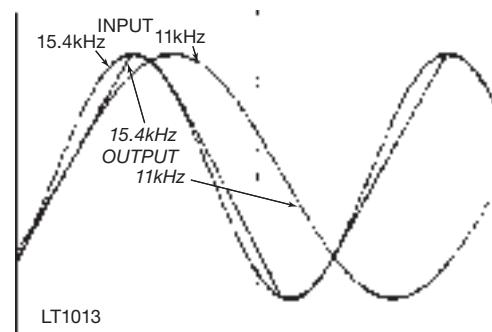
**Figure 4.48.** The maximum slew rate of a sinewave,  $\text{SR}=2\pi Af$ , occurs at the zero crossings.

### K. Slew rate

The op-amp “compensation” capacitance (discussed further in §4.9.2) and small internal drive currents act together to limit the rate at which the output can change, even when a large input unbalance occurs. This limiting speed is usually specified as *slew rate* or slewing rate (SR). For the 411 it is 15 V/μs; low-power op-amps typically have slew rates less than 1 V/μs, whereas a high-speed op-amp might slew at hundreds of volts per microsecond. The slew rate limits the amplitude of an undistorted sine-wave output swing above some critical frequency (the frequency at which the full supply swing requires the maximum slew rate of the op-amp), thus the “output voltage swing as a function of frequency” graph (seen in datasheets; see for example Figure 4.54). A sine wave of frequency  $f$  hertz and amplitude  $A$  volts requires a minimum SR of  $2\pi Af$  volts per second, with the peak slewing occurring at the zero crossings (Figure 4.48). Figure 4.49 shows a ‘scope trace illustrating real-world “slew-rate distortion.”

For externally compensated op-amps, the slew rate depends on the compensation network used. In general, it will be lowest for “unity-gain compensation,” increasing to perhaps 30 times faster for  $\times 100$  gain compensation. This is discussed further in §4.9.2B and in Chapter 4x.<sup>25</sup> As with gain–bandwidth product  $f_T$ , higher SR op-amps run at higher supply currents.

An important note: slew rate is ordinarily specified for a unity-gain configuration (i.e., a follower) with a full-swing step input. So there’s a large differential drive at the op-amp’s input, which really gets the currents flowing in there. The slew rate will be considerably less for a small input, say 10 mV.



**Figure 4.49.** Slew-rate-induced distortion. This scope trace of an LT1013 op-amp follower, for which the datasheet specifies a typical SR of 0.4 V/μs, shows the input and output waveforms for a sinewave whose peak SR is 0.6 V/μs ( $A = 6.0 \text{ V}$ ,  $f=15.4 \text{ kHz}$ ); also shown is a slower sinewave, which overlays its (identical) output ( $A = 6.0 \text{ V}$ ,  $f=11 \text{ kHz}$ ; SR = 0.4 V/μs). Scales: 2 V/div, 10 μs/div.

### L. Temperature dependence

Most of these parameters have some temperature dependence. However, this usually doesn’t make any difference, since small variations in gain, for example, are almost entirely compensated by feedback. Furthermore, the variations of these parameters with temperature are typically small compared with the variations from unit to unit.

The exceptions are input offset voltage and input offset current; these input errors will matter, particularly if you’ve trimmed the offsets approximately to zero, and will cause drifts in the output. When high precision is important, a low-drift “instrumentation” op-amp should be used, with external loads kept above 10k to minimize the horrendous effects on input-stage performance caused by temperature gradients. We will have much more to say about this subject in Chapter 5.

### M. Supply voltage and current

Traditionally, most op-amps were designed for  $\pm 15 \text{ V}$  power supplies, with a smattering of “single-supply” op-amps that operated on single supplies (i.e.,  $+V$  and ground), typically from  $+5 \text{ V}$  to  $+15 \text{ V}$ . The traditional split-supply op-amps were somewhat flexible; for example, the third-generation LF411 accepts supplies from  $\pm 5 \text{ V}$  to  $\pm 18 \text{ V}$ . Most of these early op-amps ran at supply currents of a few millamps.

There has been an important trend to lower-current and especially lower-voltage operation to accommodate battery-powered equipment. So, for example, it is now common to see op-amps that operate with total supply voltages (the span from  $V_+$  to  $V_-$ ) of 5 V, or even 3 V, and run on supply currents of 10 μA to 100 μA. These are usually

<sup>25</sup> Where we show, among other things, the fact that slew rate in conventional BJT op-amps is limited by bandwidth:  $S=0.32f_T$ . Happily, this can be circumvented, with a bit of *unconventional* design.

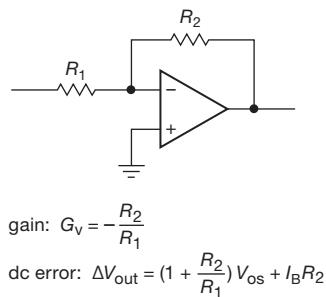


Figure 4.50. Inverting amplifier.

built with 100% CMOS circuitry, but there are some bipolar designs as well. These are usually rail-to-rail output stages – obviously such op-amps cannot afford the luxury of the “no-closer-than-2-volts-from-either-rail” mantra!

When considering these op-amps, watch out for unusually low maximum supply voltage restrictions. Many such op-amps are limited to as little as 10 V total supply (i.e.,  $\pm 5$  V), and an increasing number are limited to 5 volts or less. Also, note that an op-amp with microamp *quiescent* current will necessarily draw plenty of current if you ask it to supply that amount of current to an attached load; output current doesn’t come out of thin air.

#### N. Miscellany: CMRR, PSRR, $e_n$ , $i_n$

For completeness, we should mention here that op-amps are also limited in common-mode rejection ratio (CMRR) and power-supply rejection ratio (PSRR), i.e., their incomplete rejection of common-mode input variations and power-supply fluctuations. This becomes more important at high frequencies, where the loop gain is decreasing and where the compensation capacitor  $C_C$  couples negative-rail fluctuations into the signal chain.

In addition, op-amps are not noiseless – they introduce both voltage noise ( $e_n$ ) and current noise ( $i_n$ ) at their input. These become significant limitations primarily in connection with precision circuits and low-noise amplifiers, and they will be treated in Chapters 5 and 8.

#### 4.4.2 Effects of op-amp limitations on circuit behavior

Let’s go back and look at the inverting amplifier with these limitations in mind. We’ll see how they affect performance, and we’ll learn how to design effectively in spite of them. With the understanding we’ll get from this example, you should be able to handle other op-amp circuits. Figure 4.50 shows the circuit again.

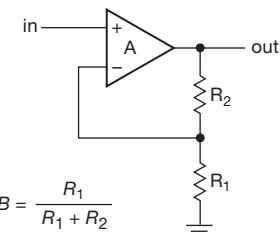


Figure 4.51. Op-amp noninverting amplifier with finite open-loop gain.

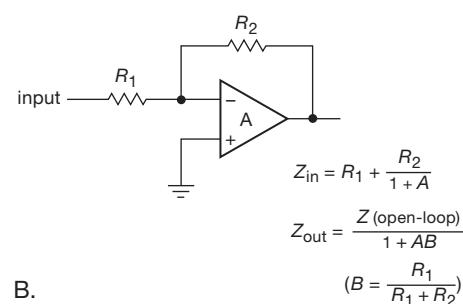
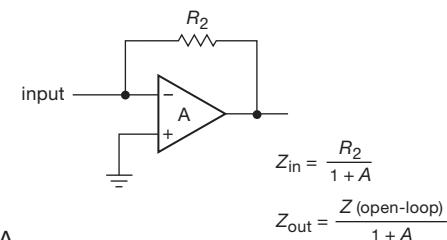


Figure 4.52. Input and output impedances: A. transresistance amplifier, B. inverting voltage amplifier.

#### A. Open-loop gain

Finite open-loop gain affects bandwidth, input and output impedances, and linearity. We saw this earlier, in the context of discrete transistor amplifiers, when we introduced negative feedback in Chapter 2 (§2.5.3). That material forms an essential background to what follows here; be sure to review it if you are foggy on this stuff.

#### Bandwidth

Because of finite open-loop gain, the voltage gain of the amplifier with feedback (closed-loop gain) will begin dropping at a frequency where the open-loop gain approaches  $R_2/R_1$  (Figure 4.47). For garden-variety op-amps like the 411, this means that you’re dealing with a relatively low-frequency amplifier; the open-loop gain is down to 100 at

40 kHz, and  $f_T$  is 4 MHz. Note that the closed-loop gain is always less than the open-loop gain, so the overall amplifier will exhibit a noticeable falloff of gain at a fraction of  $f_T$ . Recall from Chapter 2 that the closed-loop gain of the noninverting amplifier in Figure 4.51 is given by

$$G = \frac{A}{1+AB},$$

where  $B$  is the fraction of the output fed back, in this case  $B = R_1/(R_1 + R_2)$ . The output will therefore be down 3 dB at the frequency where the magnitude of the loop gain  $AB$  is unity (i.e., where the magnitude of the open-loop gain  $A$  equals the desired closed-loop gain  $1/B$ ), approximately 40 kHz for the LF411.<sup>26</sup>

Back in §4.2.5 we remarked that op-amp current sources rely on the op-amp's voltage gain (thus loop gain) to raise its inherently low output resistance  $R_o$  (of order  $\sim 100\Omega$ , see Figure 5.20), and that the decrease of open-loop gain with increasing frequency degrades the current-source's output impedance. This can be made quantitative:  $Z_{out}$  at increasing frequencies is of the form  $R_o \cdot f_T/f$ .

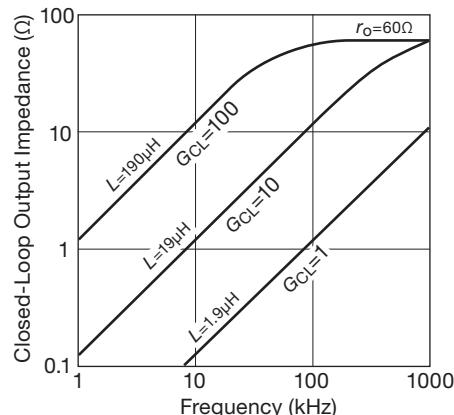
### Output impedance

Finite loop gain also affects the input and output impedances of a closed-loop op-amp circuit. Feedback can extract a sample of the output voltage (e.g., the noninverting voltage amplifiers we've been considering) or the output current (e.g., an op-amp current source). For voltage feedback the op-amp's open-loop output impedance is lowered by a factor of  $1+AB$ , bringing typical open-loop output impedances of tens to hundreds of ohms down to milliohms (for large loop gain), but rising back up to open-loop values as the loop gain falls to unity at higher frequencies.

This linear rise in closed-loop output impedance is nicely illustrated in Figure 4.53, adapted from the LT1055 datasheet. You can see how greater loop gain (feedback configured for lower closed-loop gain) produces correspondingly lower output impedance; and you can see the linear rise up to the op-amp's native  $R_{out}$  (sometimes designated  $r_o$ ), here about  $60\Omega$ . Note also that an impedance that rises linearly with frequency is like an inductor. And, in fact, that's just what the output looks like for signals in this frequency range. This can have important conse-

quences, for example, creating a series  $LC$  resonant circuit when the op-amp's load is capacitive.

The effect of the lowered loop gain (at high frequencies) is to degrade the beneficial effects of negative feedback. So a *voltage* amplifier suffers from increased output impedance, as we've seen. And the reverse is true for an amplifier with feedback that senses output *current*: here feedback normally acts to *raise* the native output impedance by a factor of loop gain (that's *good*: you want high output impedance in a current source), which then drops back to its open-loop values as the loop gain falls. Some op-amps (most notably those with rail-to-rail outputs) use an output stage with intrinsically high output impedance; for these op-amps a high loop gain is essential to achieve low output impedance.



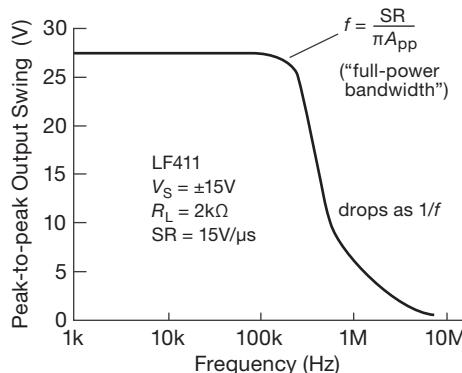
**Figure 4.53.** An op-amp's closed-loop output impedance rises approximately linearly with frequency over a large portion of its bandwidth, thus behaving like an inductance  $L_{out} \approx r_o G_{CL}/2\pi f_T$ . After the loop gain drops to unity,  $Z_{out}$  looks like the op-amp's open-loop output resistance  $r_o$ . These curves were adapted from the LT1055 datasheet.

### Input impedance

The input impedance of a noninverting amplifier is raised by a factor of  $1+AB$  from its open-loop value, a matter usually of little consequence because of the high native input impedances of op-amps.

The *inverting* amplifier circuit is different from the noninverting circuit and has to be analyzed separately. It's best to think of it as a combination of an input resistor driving a shunt feedback stage (Figure 4.52). The shunt stage alone has its input at the "summing junction" (the inverting input of the amplifier), where the currents from feedback and input signals are combined (this amplifier connection is really a "transresistance" configuration; it converts a current

<sup>26</sup> The open-loop gain  $A$  has a lagging 90° phase shift over most of the op-amp's bandwidth, as can be seen from a Bode plot like Figure 4.47, i.e., you can approximate the open-loop gain, then, by  $A(f)=j \cdot f_T/f$ . That's why the closed-loop gain is down 3 dB, and not 6 dB, when the loop gain  $AB$  has unit magnitude.



**Figure 4.54.** Peak-to-peak output swing versus frequency (LF411).

input to a voltage output). Feedback reduces the impedance looking into the summing junction,  $R_2$ , by a factor of  $1+A$  (see if you can prove this). In cases of very high loop gain, the input impedance is reduced to a fraction of an ohm, a good characteristic for a current-input amplifier.

The classic op-amp inverting amplifier connection is a combination of a shunt feedback transresistance amplifier and a series input resistor, as in the figure. As a result, the input impedance equals the sum of  $R_1$  and the impedance looking into the summing junction. For high loop gain,  $R_{in}$  approximately equals  $R_1$ .

It is a straightforward exercise to derive an expression for the closed-loop voltage gain of the inverting amplifier with finite loop gain. The answer is

$$G = -A(1 - B)/(1 + AB), \quad (4.5)$$

where  $B$  is defined as before,  $B = R_1/(R_1 + R_2)$ . In the limit of large open-loop gain  $A$ ,  $G = -1/B + 1$  (i.e.,  $G = -R_2/R_1$ ).

**Exercise 4.14.** Derive the foregoing expressions for input impedance and gain of the inverting amplifier.

### Linearity

In the limit of infinite loop gain, a feedback circuit's behavior depends on only the feedback network; the native nonlinearities of the op-amp (e.g., voltage dependence of gain, crossover distortion, and so on) are compensated by feedback. These defects reappear as loop gain is reduced, for example at higher frequencies. It is for this reason that you have to choose your op-amps with care, for example if you want to design low-distortion audio amplifier circuits. Op-amps intended for this sort of application have carefully designed output stages, and often they specify distor-

tion as a function of frequency and gain. An example is the excellent AD797, which specifies a maximum distortion of 0.0003% at 20 kHz and 3 V (rms) output.

### B. Slew rate

Because of limited slew rate, the maximum undistorted sinewave output swing drops above a certain frequency. Figure 4.54 shows the curve for a 411, with its  $15V/\mu s$  slew rate. For slew rate  $S$ , the output amplitude is limited to  $A(pp) \leq S/\pi f$  for a sinewave of frequency  $f$ , thus explaining the  $1/f$  dropoff of the curve. The flat portion of the curve reflects the power-supply limits of output voltage swing. An easy formula to remember is<sup>27</sup>

$$S_{min} = \omega A = 2\pi f A \quad (4.6)$$

where  $S_{min}$  is the minimum required SR for a sinewave of amplitude  $A$  (that's half the peak-to-peak amplitude:  $A_{pp} = 2A$ ) and angular frequency  $\omega$ ; recall that  $\omega = 2\pi f$ . As an aside, the slew-rate limitation of op-amps can be usefully exploited to filter sharp noise spikes from a desired signal, with a technique known as *nonlinear lowpass filtering*: if the slew rate is deliberately limited, the fast spikes can be dramatically reduced with little distortion of the underlying signal.

### C. Output current

Because of limited output-current capability, an op-amp's output swing is reduced for small load resistances, as we saw in Figure 4.44. For precision applications it is a good idea to avoid large output currents in order to prevent on-chip thermal gradients produced by excessive power dissipation in the output stage.

### D. Offset voltage

Because of input offset voltage, a zero input produces an output<sup>28</sup> of  $V_{out} = G_{dc}V_{OS} = (1 + R_2/R_1)V_{OS}$ . For an inverting amplifier with a voltage gain of 100 built with a 411, the output could be as large as  $\pm 0.2$  volt when the input

<sup>27</sup> Readers comfortable with calculus will recognize this simply as the magnitude of the time derivative of a sinusoid, which brings out one factor of  $\omega$ .

<sup>28</sup> Note that the relevant gain is the *noninverting* gain; that is because the  $V_{OS}$  error acts not at the *circuit's* input, but at the *op-amp's* input terminals. So the effect is as if the error  $V_{OS}$  were applied to the noninverting terminal of the amplifier.

is grounded ( $V_{OS} = 2 \text{ mV}$  max). Solutions: (a) If you don't need gain at dc, use a capacitor to drop the gain to unity at dc, as in Figure 4.7B. In this case you could do that by capacitively coupling the input signal. (b) Trim the voltage offset to zero with the manufacturer's recommended trimming network. (c) Use an op-amp with smaller  $V_{OS}$ . (d) Trim the voltage offset to zero with an external trimming network, as for example in §4.8.3 (see Figure 4.91).

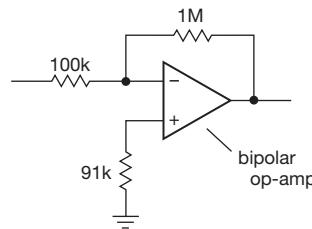
### E. Input bias current

Even with a perfectly trimmed op-amp (i.e.,  $V_{OS} = 0$ ), our inverting amplifier circuit will produce a nonzero output voltage when its input terminal is connected to ground. That is because the finite input bias current,  $I_B$ , produces a voltage drop across the resistors, which is then amplified by the circuit's voltage gain. In this circuit the inverting input sees a driving impedance of  $R_1 \parallel R_2$ , so the bias current produces a voltage  $V_{in} = I_B(R_1 \parallel R_2)$ , which is then amplified by the gain at dc,  $1 + R_2/R_1$  (see footnote 28); the result is an output error voltage of  $V_{out} = I_B R_2$ .

With FET-input op-amps the effect is usually negligible, but the substantial input current of bipolar op-amps (and also *current-feedback* op-amps; see Chapter 4x) can cause real problems. For example, consider an inverting amplifier with  $R_1=10\text{k}$  and  $R_2=1\text{M}$ ; these are reasonable values for an audiofrequency inverting stage, where we might like to keep  $Z_{in}$  at least  $10\text{k}$ . If we chose the low-noise bipolar NE5534 ( $I_B=2\mu\text{A}$ , max), the output (for grounded input) could be as large as  $100 \times 2\mu\text{A} \times 9.9\text{k}$ , or 1.98 volt ( $G_{dc}I_B R_{unbalance}$ ), which is unacceptable. By comparison, for our jellybean LF411 (JFET-input) op-amp, the corresponding worst-case output (for grounded input) is 0.2 mV; for most applications this is negligible, and in any case is dwarfed by the  $V_{OS}$ -produced output error (200 mV, worst-case untrimmed, for the LF411).

There are several solutions to the problem of bias-current errors. If you must use an op-amp with large bias current, it is a good idea to ensure that both inputs see the same dc driving resistance, as in Figure 4.55. In this case, 91k is chosen as the parallel resistance of 100k and 1M. In addition, it is best to keep the resistance of the feedback network small enough so that bias current doesn't produce large offsets; typical values for the resistance seen from the op-amp inputs are 1k to 100k or so. A third cure involves reducing the gain to unity at dc, as in Figure 4.7B.

In most cases, though, the simplest solution is to use op-amps with negligible input current. Op-amps with JFET- or MOSFET-input stages generally have input currents in the picoamp range (watch out for its rapid rise versus temperature, though, roughly doubling every  $10^\circ\text{C}$ ),



**Figure 4.55.** With bipolar op-amps, use a compensation resistor to reduce errors caused by input bias current.

and many modern bipolar designs use superbeta transistors or bias-cancellation schemes to achieve bias currents nearly as low and *decreasing* slightly with temperature. With these op-amps, you can have the advantages of bipolar op-amps (precision, low noise) without the annoying problems caused by input current. For example, the precision low-noise bipolar OP177 has  $I_B < 2\text{nA}$ , and the bias-compensated bipolar LT1012 has  $I_B = \pm 25\text{pA}$  (typ). Among inexpensive FET op-amps, the JFET LF411 has  $I_B = 50\text{pA}$  (typ), and the MOSFET TLC270 series, priced under a dollar, has  $I_B = 1\text{pA}$  (typ).

### F. Input offset current

As we just described, it is usually best to design circuits so that circuit impedances, combined with op-amp bias current, produce negligible errors. However, occasionally it may be necessary to use an op-amp with high bias current, or to deal with signals of extraordinarily high Thévenin impedances. Examples of high-bias-current amplifiers are current-feedback op-amps (e.g., the AD844), low-noise ( $e_n$ ) op-amps (e.g., the AD797), and wideband op-amps (e.g., the LM7171), each with input currents of several *microamps*.

In these cases the best you can do is to balance the dc driving resistances seen by the op-amp at its input terminals. There will still be some error at the output ( $G_{dc}I_{offset}R_{source}$ ) that is due to unavoidable asymmetry in the op-amp input currents. In general,  $I_{offset}$  is smaller than  $I_{bias}$  by a factor of 2 to 20 (with bipolar op-amps generally showing better matching than FET op-amps).

### G. Limitations imply tradeoffs

In the preceding paragraphs we discussed the effects of op-amp limitations, taking the example of the simple inverting voltage amplifier circuit. Thus, for example, op-amp input current caused a *voltage* error at the output. In a different op-amp application you may get a different effect; for example, in an op-amp integrator circuit, a finite input current produces an output *ramp* (rather than a

constant) with zero applied input. As you become familiar with op-amp circuits you will be able to predict the effects of op-amp limitations in a given circuit and therefore choose which op-amp to use in a given application. In general, there is no “best” op-amp (even when price is no object): for example, op-amps with the very lowest input currents (MOSFET types) generally have larger voltage offsets and greater noise, and vice versa. Good circuit designers choose their components with the right tradeoffs to optimize performance, without going overboard on unnecessary “gold-plated” parts.

To help put this discussion of “op-amp realities” into perspective, you may want to look again at Table 4.1 on page 245, where we’ve summarized the kinds of performance you can expect from op-amps that might be described as average, or “jellybean” (for example, the LF412 is a jellybean JFET op-amp), and from those that are among the best available (“premium”) *for each given parameter*. Sadly, you can’t get an op-amp that combines all the characteristics in a “premium” column; *engineering is the art of compromise*.

The limitations of op-amp performance we have talked about will have an influence on component values in nearly all circuits. For instance, the feedback resistors must be large enough so that they don’t load the output significantly, but they must not be so large that input bias current produces sizable offsets. High impedances in the feedback network cause both loading effects, and destabilizing phase shifts, from stray capacitances; they also increase susceptibility to capacitive pickup of interfering signals. These tradeoffs typically dictate resistor values of 2k to 100k with general-purpose op-amps.

Similar sorts of tradeoffs are involved in almost all electronic design, including the simplest circuits constructed with transistors. For example, the choice of quiescent current in a transistor amplifier is limited at the high end by device dissipation, increased input current, excessive supply current, and reduced current gain, whereas the lower limit of operating current is limited by leakage current, reduced current gain, and reduced speed (from stray capacitance in combination with the high resistance values). For these reasons you typically wind up with collector currents in the range of a few tens of microamps to a few tens of milliamps (higher for power circuits, sometimes a bit lower in “micropower” applications), as mentioned in Chapter 2.

In later chapters we look more carefully at some of these problems in order to convey a good understanding of the tradeoffs involved.

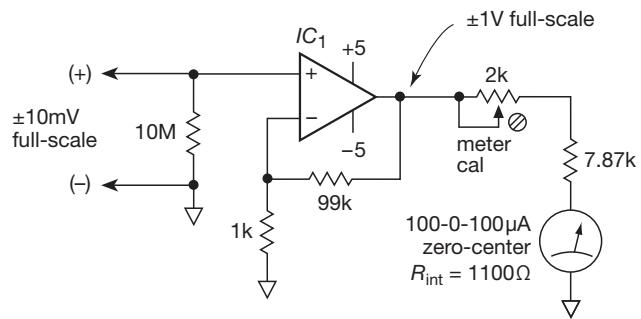
**Exercise 4.15.** Draw a dc-coupled inverting amplifier with a gain

of 100 and  $Z_{in} = 10k$ . Include compensation for input bias current and show offset voltage trimming network (10k pot between pins 1 and 5, wiper tied to  $V_-$ ). Now add circuitry so that  $Z_{in} \geq 10^8 \Omega$ .

#### 4.4.3 Example: sensitive millivoltmeter

To put flesh on these bones, let’s look at a very simple design example – a dc amplifier with lots of gain, high input impedance, and (for variety in today’s too-digital world) an *analog* zero-center panel meter readout. We’ll aim for  $\pm 10 \text{ mV}$  full-scale sensitivity, and 10 megohms input impedance.

Figure 4.56 shows the initial design, where we assume we’ve got  $\pm 5 \text{ V}$  supplies available (more on this later), and we use a noninverting gain of 100 to produce a  $\pm 1 \text{ V}$  op-amp output at fullscale. That drives a 100–0–100  $\mu\text{A}$  zero-center meter movement, which we decorate with a relabeled scale that reads “ $-10 \text{ mV} \dots 0 \dots 10 \text{ mV}$ .”



**Figure 4.56.** Sensitive millivoltmeter with analog readout; see text for choice of IC<sub>1</sub>.

It looks simple, and it is. But it won’t work well if we’re not careful. Imagine we choose our default LF411 for IC<sub>1</sub>, figuring that the low bias current of a JFET-input op-amp is just what we need. We short the input leads together and find to our horror that the meter reads way off center, as much as  $\pm 2 \text{ mV}$ . That’s because the 411 has  $V_{OS}=2 \text{ mV}$  (max). Ideally we’d like the thing to read zero with the input leads shorted or open, where “zero” might realistically mean no more than 1% of the full-scale reading.

OK, we add an offset trimmer and adjust it until the output reads zero with shorted input. We leave it on the bench, go to lunch, then come back and find it now reads  $-0.2 \text{ mV}$  with shorted input. That’s because it sat in the sun, warmed up by  $10^\circ\text{C}$ , and therefore drifted by  $200 \mu\text{V}$  (the LF411 has a temperature coefficient of offset voltage  $TCV_{OS}=20 \mu\text{V}/^\circ\text{C}$ , max). Well, we won’t use it in the sun!

So, we wait for it to cool down, and note with satisfaction that it reads zero again.

Now we go to test it on a voltage divider, but we find that when we remove the short (test leads open circuited) the meter reads +2 mV! This time the problem is bias current, specified as 200 pA (max) at room temperature. That's not much current, but it develops 2000  $\mu$ V across the 10M input resistor, which the meter dutifully reports as an input.

We could solve the  $V_{OS}$  problem with a precision bipolar op-amp, but we'd be in worse trouble with  $I_B$ . We need an op-amp with  $V_{OS} < 100 \mu$ V and  $I_B < 10$  pA. The solution is a precision FET-input op-amp, for example, the OPA336 (125  $\mu$ V untrimmed, and 10 pA), which is just about good enough without trimming, and certainly fine if we're willing to trim the initial offset.

A better solution here is to use a “chopper” op-amp like the LTC1050C or AD8638 (see Table 5.6). These are sometimes called “zero-drift,” or “auto-zeroing” amplifiers. We'll learn about them shortly, and in more detail in the next chapter; for now all you need to know is that they offer specifications like  $V_{OS} < 5 \mu$ V,  $TCV_{OS} < 0.05 \mu$ V/ $^{\circ}$ C, and  $I_B < 50$  pA (those are, in fact, the worst-case specs for the LTC1050 in its inexpensive C-suffix grade).

*A final thought:* having a calibration adjustment is OK if you're building just a few of these things. But in production it would be nice to avoid manual calibration steps. An elegant circuit solution that circumvents calibration is to use a current-sensing resistor on the low side of the meter. We include this feature when we revisit this example early in Chapter 5 (§5.2), in a more rigorous approach to precision design.

#### 4.4.4 Bandwidth and the op-amp current source

Back in §4.2.5 we remarked that op-amp current sources rely on the op-amp's voltage gain (thus loop gain) to raise its inherently low output resistance  $R_o$  (of order  $\sim 100 \Omega$ , see Figure 5.20), and that the decrease of open-loop gain with increasing frequency degrades the current-source's output impedance. Put another way, the op-amp current source is a peculiar circuit, because an op-amp virtue (inherently low output impedance, i.e., a voltage source) becomes a vice, which must be punished with the cudgel of plentiful loop gain. This can be made quantitative: because of finite bandwidth  $f_T$ , the output impedance of an op-amp current source at increasing frequencies is of the form  $R_o \cdot f_T / f$ , dropping ultimately to the op-amp's native output resistance  $R_o$  at the unity-gain frequency  $f_T$ .

Likewise, finite slew rate affects the current-source's output impedance, making it look like a shunt capacitance.

Here's how to think about it: an ideal current source with a *real* capacitive load slews at a rate  $S = dV/dt = I/C$ ; so a current source that is afflicted with a maximum slew rate  $S$  looks like an ideal current source burdened with an effective shunt capacitance  $C_{eff} = I_{out}/S$ . For example, a 10 mA current source made with an op-amp of slew rate 1 V/ $\mu$ s has an effective capacitive load of 10 nF; this is rather large, compared even with a large MOSFET.

### 4.5 A detailed look at selected op-amp circuits

The performance of the next few circuits is affected significantly by the limitations of op-amps; we will go into a bit more detail in their description.

#### 4.5.1 Active peak detector

There are numerous applications in which it is necessary to determine the peak value of some input waveform. The simplest method is a diode and capacitor (Figure 4.57). The highest point of the input waveform charges up  $C$ , which holds that value while the diode is back-biased.

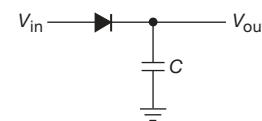


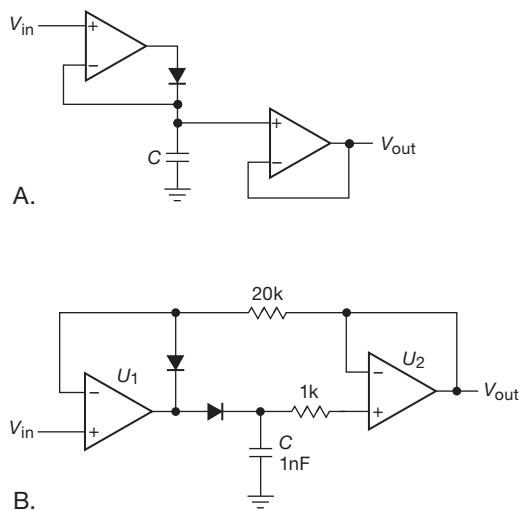
Figure 4.57. Passive peak detector.

This method has some serious problems. The input impedance is variable and is very low during peaks of the input waveform. Also, the diode drop makes the circuit insensitive to peaks less than about 0.6 volt, and inaccurate (by one diode drop) for larger peak voltages. Furthermore, since the diode drop depends on temperature and current, the circuit's inaccuracies depend on the ambient temperature and on the rate of change of output; recall that  $I = C(dV/dt)$ . An input emitter follower would improve the first problem only.

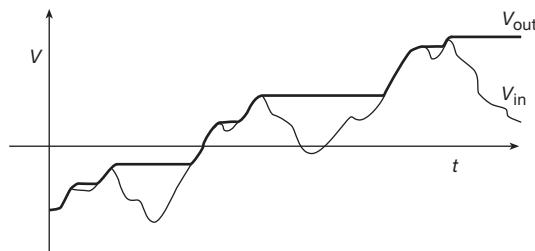
Figure 4.58A shows a better circuit, which exhibits the benefits of feedback. By taking feedback from the voltage at the capacitor, the diode drop doesn't cause any problems. The sort of output waveform you might get is shown in Figure 4.59.

Op-amp limitations affect this circuit in three ways.

- A finite op-amp slew rate causes a problem, even with relatively slow input waveforms. To understand this, note that the op-amp's output goes into negative saturation when the input is less positive than the output



**Figure 4.58.** A. Op-amp peak detector (more accurately, a “peak tracker”). B. Improved peak tracker responds to short peaks, because the input op-amp does not have to slew from negative saturation.



**Figure 4.59.** Peak detector output waveform.

(try sketching the op-amp voltage on the graph; don’t forget about diode forward drop). So the op-amp’s output has to race back up to the output voltage (plus a diode drop) when the input waveform next exceeds the output. At slew rate  $S$ , this takes roughly  $(V_{out} - V_-)/S$ , where  $V_-$  is the negative supply voltage. Improved circuit 4.58B solves this problem.

- (b) Input bias current causes a slow discharge (or charge, depending on the sign of the bias current) of the capacitor. This is sometimes called “droop,” and it is best avoided by using op-amps with very low bias current. For the same reason, the diode must be a low-leakage type (e.g., the FJH1100, with less than 1 pA reverse current at 20 V, an “FET diode” such as the PAD5, or a diode-connected JFET such as the 2N4417; see the diode discussion in Chapter 1x), and the follow-

ing stage must also present high impedance (ideally it should also be an FET or FET-input op-amp).

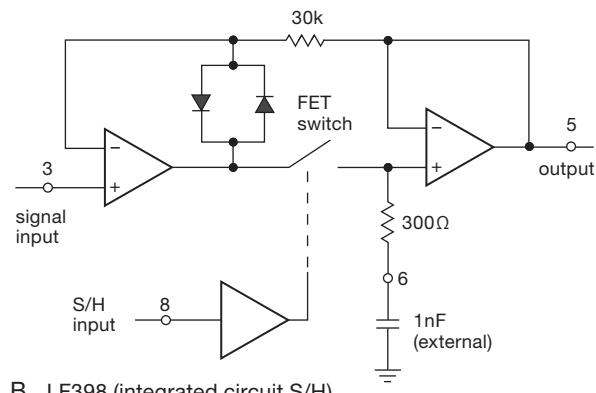
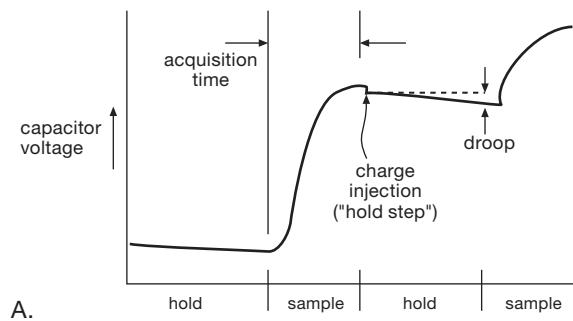
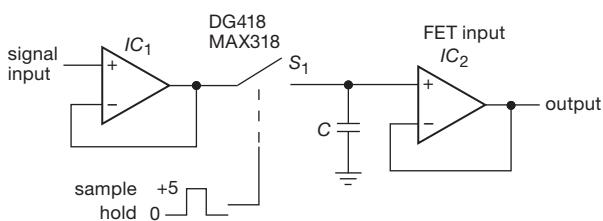
- (c) The maximum op-amp output current limits the rate of change of voltage across the capacitor, i.e., the rate at which the output can follow a rising input. Thus the choice of capacitor value is a compromise between low droop and high output slew rate.

For instance, a  $1 \mu\text{F}$  capacitor used in this circuit with the common LM358 (which would be a poor choice because of its high bias current) would droop at  $dV/dt = I_B/C = 0.04 \text{ V/s}$  (using the “typical” value  $I_B = 40 \text{ nA}$ ; the worst-case value of  $I_B = 500 \text{ nA}$  produces a droop of  $0.5 \text{ V/s}$ ) and would follow input changes only up to  $dV/dt = I_{\text{output}}/C = 0.02 \text{ V}/\mu\text{s}$ . This maximum follow rate is much less than the op-amp’s slew rate of  $0.5 \text{ V}/\mu\text{s}$ , being limited by the maximum output current of  $20 \text{ mA}$  driving  $1 \mu\text{F}$ . By decreasing  $C$  you could achieve greater output slewing rate, at the expense of greater droop. A more realistic choice of components would be the popular TLC2272 MOSFET-input op-amp as driver and output follower ( $1 \text{ pA}$  typical bias current) and a value of  $C=0.01 \mu\text{F}$ . With this combination you would get a droop of only  $0.0001 \text{ V/s}$  and an overall circuit slew rate of  $2 \text{ V}/\mu\text{s}$ . For even better performance, use a MOSFET op-amp like the LMC660 or LMC6041, with a typical input current of  $2 \text{ femtoamps}$ . Capacitor leakage (or diode leakage or both) may then limit performance even if unusually good capacitors are used, e.g., polystyrene or polypropylene (see Chapter 1x).<sup>29</sup>

### A. Resetting a peak detector

In practice it is usually desirable to reset the output of a peak detector in some way. One possibility is to put a resistor across the peak-hold capacitor so that the circuit’s output decays with a time constant  $RC$ . In this way it holds only the most recent peak values. A better method is to put a transistor switch across  $C$ ; a short pulse to the base then zeros the output. An FET switch is often used instead. For example, in Figure 4.58 you could connect an *n*-channel MOSFET such as a 2N7000 across  $C$ ; bringing the gate

<sup>29</sup> There’s more to capacitor “leakage” than one might at first suspect: an effect known as *dielectric absorption* (“DA”) can cause serious mischief in circuits that rely on ideal capacitor performance. It manifests itself rather clearly in the following simple experiment: charge a tantalum capacitor up to 10 volts or so, let it sit there for a while, then rapidly discharge it by momentarily putting a  $100\Omega$  resistor across it. Remove the resistor and watch the capacitor’s voltage on a high-impedance voltmeter. You will be amazed to see the capacitor *charge back up*, reaching perhaps a volt or so after a few seconds! This unhelpful effect is treated in more detail in Chapter 1x and §5.6.2.



B. LF398 (integrated circuit S/H)

**Figure 4.60.** Sample-and-hold: A. Standard configuration, with exaggerated waveform; B. LF398 single-chip S/H.

momentarily positive then zeros the capacitor voltage. An integrated CMOS analog switch (such as the MAX318, with a small series resistor to limit the current) can be used instead of a discrete nMOS (*n*-type metal-oxide semiconductor) transistor.

#### 4.5.2 Sample-and-hold

Closely related to the peak detector is the “sample-and-hold” (S/H) circuit (sometimes called “follow-and-hold”). These are especially popular in digital systems in which you want to convert one or more analog voltages to num-

bers so that a computer can digest them: the favorite method is to grab and hold the voltage(s), then do the digital conversion at your leisure. The basic ingredients of an S/H circuit are an op-amp and an FET switch; Figure 4.60A shows the idea. IC<sub>1</sub> is a follower to provide a low-impedance replica of the input. CMOS analog switch S<sub>1</sub> passes the signal through during “sample” and disconnects it during “hold.” Whatever signal was present when S<sub>1</sub> was turned OFF is held on capacitor C. IC<sub>2</sub> is a high-input-impedance follower (FET inputs), so that capacitor current during “hold” is minimized. The value of C is a compromise: leakage currents in S<sub>1</sub> and the follower cause C’s voltage to “droop” during the hold interval according to  $dV/dt = I_{leakage}/C$ . Thus C should be large to minimize droop. But S<sub>1</sub>’s ON-resistance forms a lowpass filter in combination with C, so C should be small if high-speed signals are to be followed accurately. IC<sub>1</sub> must be able to supply C’s charging current ( $I = CdV/dt$ ) and must have sufficient slew rate to follow the input signal. In practice, the slew rate of the whole circuit will usually be limited by IC<sub>1</sub>’s output current and S<sub>1</sub>’s ON-resistance.

**Exercise 4.16.** Suppose IC<sub>1</sub> can supply 10 mA of output current and  $C = 0.01 \mu\text{F}$ . What is the maximum input slew rate the circuit can accurately follow? If S<sub>1</sub> has  $50 \Omega$  ON resistance, what will be the output error for an input signal slewing at  $0.1 \text{ V}/\mu\text{s}$ ? If the combined leakage of S<sub>1</sub> and IC<sub>2</sub> is 1 nA, what is the droop rate during the “hold” state?

For both the S/H circuit and the peak detector, an op-amp drives a capacitive load. When designing such circuits, make sure you choose an op-amp that is stable at unity gain when loaded by the capacitor C. Some op-amps, (e.g., the LT1457, a member of Linear Technology’s “CLOAD™ stable” op-amps) are specifically designed to drive large ( $0.01 \mu\text{F}$ ) capacitive loads directly. Some other tricks you can use are discussed in §4.6.1B and in Chapter 4x.

You don’t have to design S/H circuits from scratch, because there are nice monolithic ICs that contain all the parts you need. National’s LF398 is a popular part, containing the FET switch and two op-amps in an inexpensive (\$1.25) 8-pin package. Figure 4.60B shows how to use it. Note how feedback closes the loop around *both* op-amps. There are plenty of fancy S/H chips available, if you need better performance than the LF398 offers. For example, the AD783 from Analog Devices includes an internal capacitor and guarantees a maximum acquisition time of  $0.4 \mu\text{s}$  for 0.01% accuracy following a 5 volt step.

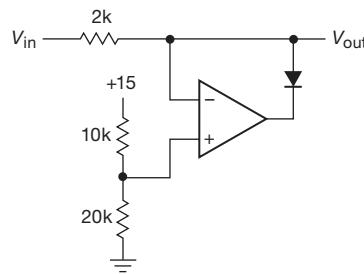


Figure 4.61. Active clamp.

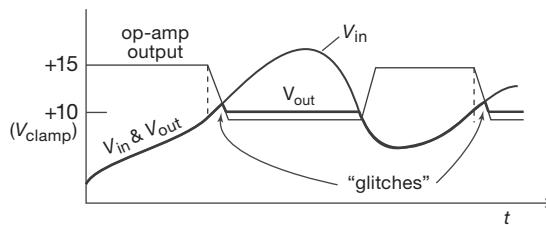


Figure 4.62. Finite slew-rate causes clamp output "glitches".

### 4.5.3 Active clamp

Figure 4.61 shows a circuit that is an active version of the clamp function we discussed in Chapter 1. For the values shown,  $V_{in} < +10$  volts puts the op-amp output at positive saturation, and  $V_{out} = V_{in}$ . When  $V_{in}$  exceeds  $+10$  volts the diode closes the feedback loop, clamping the output at  $+10$  volts. In this circuit, op-amp slew-rate limitations allow small glitches as the input reaches the clamp voltage from below (Figure 4.62).

**Exercise 4.17.** The active clamp in Figure 4.61 suffers from a slew-rate speed limitation similar to that of the peak tracker of Figure 4.58A. Figure out an improvement to the clamp circuit, analogous to the trick used in Figure 4.58B.

### 4.5.4 Absolute-value circuit

The circuit shown in Figure 4.63 gives a positive output equal to the magnitude of the input signal; it is a full-wave rectifier. As usual, the use of op-amps and feedback eliminates the diode drops of a passive full-wave rectifier.

You can imagine situations in which you want an output proportional to the *logarithm* of the absolute value. A simple circuit change might be to substitute a diode (or transistor with base tied to collector) for the feedback resistor of the second op-amp, exploiting the Ebers–Moll relation between diode voltage and summing-junction current. As we will see in Chapter 4x, this is the basis of the *logarithmic amplifier*; the circuit needs a few

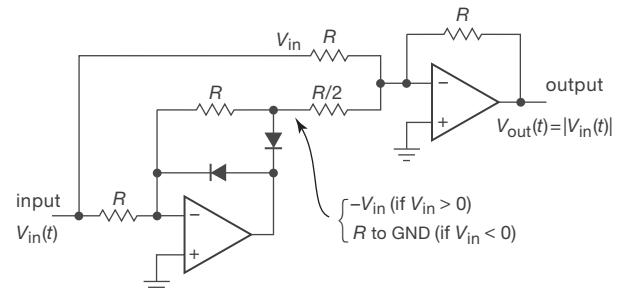
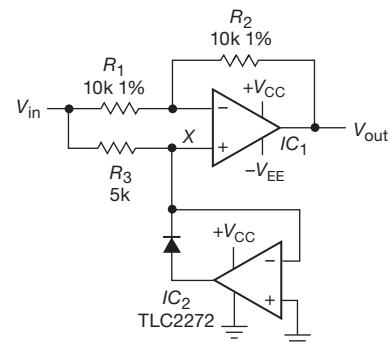


Figure 4.63. Active full-wave rectifier.

Figure 4.64. Another full-wave rectifier; note that ground is the negative supply voltage for IC<sub>2</sub>.

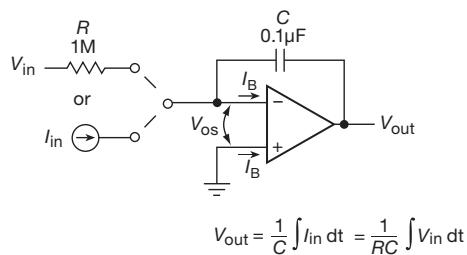
additional components, however, to compensate for the temperature coefficient of  $V_{BE}$ .

**Exercise 4.18.** Figure out how the circuit in Figure 4.63 works. Hint: apply first a positive input voltage and see what happens; then do negative.

Figure 4.64 shows another absolute-value circuit. It is readily understandable as a simple combination of an optional inverter (IC<sub>1</sub>) and an active clamp (IC<sub>2</sub>). For negative input levels the clamp holds point X at ground, making IC<sub>1</sub> a unity-gain inverter; for positive input levels, the clamp is out of the circuit, with its output at negative saturation, making IC<sub>1</sub> a follower. Thus the output is equal to the absolute value of the input voltage. By running IC<sub>2</sub> from a single positive supply, you avoid problems of slew-rate limitations in the clamp, since its output moves over only one diode drop. Note that no great accuracy is required of  $R_3$ .

### 4.5.5 A closer look at the integrator

We introduced the op-amp integrator in §4.2.6, before dealing with input bias current and offset voltage. One problem with that circuit (Figure 4.16) is that the output tends



Current Error					
$I_e = I_B$		if voltage input		Total Current Error	
		$[+ V_{\text{os}}/R]$		$I_{\text{input}}$	$V_{\text{input}}$
OP27E	40nA	[+ 25pA] (25μV/1M)		40nA	40nA
LMC6042A	4pA	[+ 3nA] (3mV/1M)		0.004nA	3nA
OP97E	100pA	[+ 25pA] (25μV/1M)		0.1nA	0.13nA

Figure 4.65. Integrator errors: bias current and offset voltage.

to wander off, even with the input grounded, because of op-amp offsets and bias current (there's no feedback at dc, which violates the third item in §4.2.7). This problem can be minimized by using a FET op-amp for low input current and offset, trimming the op-amp input offset voltage, and using large  $R$  and  $C$  values. Furthermore, in applications in which the integrator is zeroed periodically by closing a switch placed across the capacitor (Figures 4.18A–C), only the drift over short time scales matters.

It's worth looking at this in a bit more detail. Look at the integrator in Figure 4.65, shown with a choice of voltage input  $V_{\text{in}}$  (which, in the absence of op-amp errors, produces a current into the summing junction of  $I = V_{\text{in}}/R$ ), or a current input  $I_{\text{in}}$  (in which case you omit the input resistor  $R$ ). The ideal integrator produces an output

$$V_{\text{out}}(t) = -\frac{1}{C} \int I_{\text{in}}(t) dt = -\frac{1}{RC} \int V_{\text{in}}(t) dt.$$

It's easy enough to figure out the effect of the op-amp's input errors  $I_B$  and  $V_{\text{os}}$ . Let's first take the case of an integrator circuit with *current* input.<sup>30</sup> The op-amp's bias current  $I_B$  adds (or subtracts) from the true input current  $I_{\text{in}}$ ; in the absence of any external input current the integrator's output would ramp at a rate  $dV_{\text{out}}/dt = I_B/C$ . The effect of op-amp input offset voltage, on the other hand, is simply

<sup>30</sup> Examples of signals that are naturally in the form of a current include those from a photodiode, a PMT, or an ion detector, or from dielectric, semiconductor or nanomaterial measurements.

to offset the output voltage by  $V_{\text{os}}$ , without ramping;<sup>31</sup> so, when you reset the integrator by shorting the feedback capacitor  $C$ , the output goes to a voltage equal to  $V_{\text{os}}$  rather than zero.

Let's look at some actual values. In Figure 4.65 we've chosen, rather arbitrarily, values of  $0.1 \mu\text{F}$  for  $C$  and (for voltage input)  $1\text{M}\Omega$  for  $R$ . So a positive input current of  $1 \mu\text{A}$  produces an output ramp of  $-10 \text{ V/s}$ . If we were to choose the precision bipolar OP27E, its relatively high input current of  $\pm 40 \text{ nA}$  (max) would cause an output ramp of as much as  $dV_{\text{out}}/dt = I_B/C = \pm 0.4 \text{ V/s}$ .

This isn't good, particularly if you want to integrate for a few seconds or more. So let's fix things by choosing an op-amp that excels in low bias current, for example the CMOS LMC6041A (the suffixes denote the particular grade; we've chosen the best in all cases). It has a specified maximum bias current of  $4 \text{ pA}$  over its temperature range (but an astounding<sup>32</sup> "typical" value of  $2 \text{ fA}$ , or  $2 \times 10^{-15} \text{ A}$ ). Now the worst-case output ramp, in the absence of any input signal current, is reduced to  $dV_{\text{out}}/dt = I_B/C = \pm 40 \mu\text{V/s}$ . The "typical" ramping rate is 2000 times smaller, if the specs can be believed; that's a mere  $0.02 \mu\text{V/s}$ .

At this point, the lesson seems to be that the best op-amp for any integrator is the one with the smallest bias current  $I_B$ . But, alas, life is more complicated. In particular, if the integrator is wired for *voltage* input, with a series input resistor  $R$ , then the op-amp's offset voltage  $V_{\text{os}}$  now produces a ramp when the input to the circuit is held at ground. Imagine the input is grounded ( $V_{\text{in}} = 0$ ), and think about it this way: the op-amp strives to align its inputs with a voltage  $V_{\text{os}}$  between them; that small voltage then produces a current  $I = V_{\text{os}}/R$  through the input resistor. That current has to come through the feedback capacitor, that is, the output must ramp to produce the current needed to satisfy the op-amp's twisted belief that its inputs should differ by  $V_{\text{os}}$ . Another way to say it is that the current acts just like an input current of  $I = -V_{\text{os}}/R$ .

Now the choice of op-amp is not so clear! Look at Figure 4.65 again. The CMOS op-amp with its very low bias current has a pretty large offset voltage,  $V_{\text{os}} = 3 \text{ mV}$  (max). So in this circuit it can produce an equivalent input current of  $3 \text{ nA}$  ( $3 \text{ mV}$  across  $1 \text{ M}\Omega$ ); that's nearly a thousand times

<sup>31</sup> If the input signal is not a true current source, but rather comes from a voltage  $V_{\text{in}}$  in series with a resistor  $R_{\text{in}}$ , then the op-amp's  $V_{\text{os}}$  causes an additional small error current of  $V_{\text{os}}/R_{\text{in}}$ .

<sup>32</sup> Particularly when considering that its input is protected by diode clamps to the supply rails. How did these magicians accomplish this? (The same way you shake hands with a gorilla – *very carefully!*).

larger than the worst-case contribution of its bias current, and it's getting into the same ballpark as the input current of the OP27E bipolar op-amp we considered first.

If minimum drift is needed with these particular circuit values, the solution is to choose an op-amp with the best compromise of low bias current and low offset voltage; to be precise, it should have the minimum value of the total worst-case error current  $I_E = I_B + V_{OS}/R$ . A good choice would be the bipolar OP97E, a precision (low-offset) op-amp with internal bias-cancellation circuitry. It sports maximum values of  $I_B = 0.1 \text{ nA}$  and  $V_{OS} = 25 \mu\text{V}$ ; the corresponding worst-case current error is  $I_E = 0.125 \text{ nA}$ , which is 25 times better than that of the LMC6041A and 320 times better than that of the OP27.

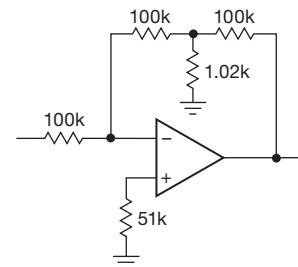
Note that the relative contribution of  $V_{OS}$  and  $I_B$  to integrator error is scaled by the value of  $R$ . So you can simply choose a larger resistor value if you have an op-amp with excellent  $I_B$  but only modest  $V_{OS}$ .

If the residual drift of an integrator circuit is still too large for a given application, or if long-term accuracy is unimportant, one solution is to put a large resistor  $R_2$  across  $C$  to provide dc feedback for stable biasing, as shown in Figure 4.18D. The effect is to roll off the integrator action at very low frequencies,  $f < 1/R_2C$ . The feedback resistor may become rather large in this sort of application. Figure 4.66 shows a trick for producing the effect of a large feedback resistor using smaller values in any op-amp circuit. In this case – an inverting amplifier circuit – the feedback network behaves like a single  $10 \text{ M}\Omega$  resistor, thus producing a voltage gain of  $-100$ . This technique has the advantage of using resistors of convenient values without the problems of stray capacitance, etc., that occur with very large resistor values. Note that this “T-network” trick may increase the effective input offset voltage if used in a trans-resistance configuration (§4.3.1C). For example, the circuit of Figure 4.66, driven from a high-impedance source (e.g., the current from a photodiode, with the input resistor omitted), has an output offset of 100 times  $V_{OS}$ , whereas the same circuit with a  $10 \text{ M}\Omega$  feedback resistor has an output equal to  $V_{OS}$  (assuming the offset that is due to input current is negligible).

#### 4.5.6 A circuit cure for FET leakage

Sometimes a circuit technique is so elegant and fascinating that we feel compelled to tell others about it. That's the case for the circuit in this section, brought over and updated from our 2nd edition. Read it and delight in its cleverness; but then check our remarks in the concluding paragraph.

In the integrator with a FET reset switch (Figure 4.18),



**Figure 4.66.** “T-network” simulates large-value resistor (here  $10 \text{ M}\Omega$ ).

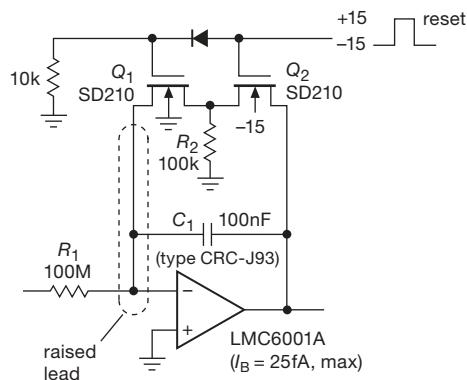
drain-source leakage sources a small current into the summing junction even when the FET is OFF. With an ultra-low-input-current op-amp and low-leakage capacitor, this can be the dominant error in the integrator. For example, the excellent LMC6001A JFET-input “electrometer” op-amp has a maximum input current of  $0.025 \text{ pA}$ , and a high-quality  $0.1 \mu\text{F}$  metallized Teflon or polystyrene capacitor specifies leakage resistance as  $10^7 \text{ megohms}$ , minimum. Thus the integrator, exclusive of reset circuit, keeps stray currents at the summing junction below  $1 \text{ pA}$  (for a worst-case  $10 \text{ V}$  full-scale output), corresponding to an output  $dV/dt$  of less than  $0.01 \text{ mV/s}$ . Compare this with the leakage contribution of a MOSFET such as the SD210 (enhancement mode), which specifies a maximum leakage current of  $10 \text{ nA}$  at  $V_{DS}=10 \text{ V}$  and  $V_{GS}=-5 \text{ V}$ ! In other words, the reset FET can contribute up to 10,000 times as much leakage as everything else combined.

Figure 4.67 shows a clever circuit solution. Although both  $n$ -channel MOSFETs are switched together,  $Q_1$  is switched with gate voltages of 0 and  $+15$  volts so that gate leakage (as well as drain-source leakage) is entirely eliminated during the OFF state (zero gate voltage). In the ON state the capacitor is discharged as before, but with twice  $R_{ON}$ . In the OFF state,  $Q_2$ 's small leakage passes to ground through  $R_2$  with negligible drop. There is no leakage current at the summing junction because  $Q_1$ 's source, drain, and substrate are all at the same voltage. (Sharp-eyed readers may have noticed that the virtual ground at the op-amp's inverting input is imperfect to the extent of its offset voltage  $V_{OS}$ .<sup>33</sup> This can be trimmed to eliminate completely any leakage current from  $Q_1$ .)

The ultimate limit to capacitor “droop” in this circuit, once FET switch leakage has been eliminated, is set by the op-amp's input current and by the capacitor's self-discharge. The capacitor shown has a specified<sup>34</sup> leakage

<sup>33</sup> “There isn't a *literal* connection, Dude.”

<sup>34</sup> In careful multiyear measurements of leakage in some polyester and



**Figure 4.67.** MOSFET leakage defeated by clever circuit.

resistance of  $10^7 \text{ M}\Omega$ , i.e.,  $10^{13} \Omega$ . The resulting leakage currents, of order  $10^{-16} \text{ A}$  (following reset), are entirely negligible compared with op-amp bias currents. For the op-amp shown, the specified bias current is 25 fA maximum (10 fA typ) at  $25^\circ\text{C}$ ; that bias current produces a maximum droop of  $0.25 \mu\text{V}/\text{s}$ . There are no op-amps with lower specified maximum bias current currently available; but you can find op-amps whose “typical” bias current is lower – for example, the LMC6041, an inexpensive op-amp whose datasheet proclaims  $I_B = 2 \text{ fA}$  (typ) at  $25^\circ\text{C}$  (no maximum is given at room temperature;  $I_B = 4 \text{ pA}$  max over the full temperature range). What is one to make of an op-amp whose typical input current is 2000 times smaller than the guaranteed maximum? Just that the manufacturer *knows* that it’s very good, but it’s too painful to test in production. You’d do well to use these inexpensive units in such a circuit if you’re willing to screen the op-amps yourself; otherwise pay the bounty for a unit that has a guaranteed limit (but note that the LMC6001A, such a unit, has a typical  $I_B$  that is five times higher than that of the less expensive LMC6041).

When designing circuits where low input current is needed, watch out for temperature effects: all FET op-amps (both JFET and CMOS types) exhibit dramatic increases in input current with rising temperature, typically doubling each  $10^\circ\text{C}$ ; the LMC6001A’s guaranteed maximum bias current jumps from 25 fA at  $25^\circ\text{C}$  to 2000 fA at  $85^\circ\text{C}$ . At high temperatures, the input (leakage) currents

polypropylene capacitors at low voltages, Tom Bruins of Agilent found actual leakage resistances to be some 10,000 times larger than specified. Using 160 V capacitors at 10 V, he measured time constants (in seconds, sometimes called “megohm-microfarad products”) of  $2 \times 10^7$  seconds (polyester) and  $>10^9$  seconds (polypropylene). In fact, what is commonly called “leakage” in such capacitors appears mostly to be dielectric absorption currents; see §5.6.2 and Chapter 1x.

of FET op-amps may often be higher than the input (bias) currents of low- $I_B$  bipolar types; that is because leakage currents rise exponentially with temperature, whereas transistor bias currents remain roughly constant (or decrease slightly). Look back at Figure 3.48 for a good illustration; see also Figures 5.6 and 5.38.

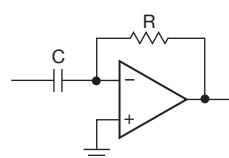
It may be difficult to find discrete low-capacitance MOSFETs with substrate pins; currently the SD210 family (with SST-prefix SMT versions) is available from Linear Systems (Fremont, CA). The two-switch T-configuration is sound, though it may be challenging to find suitable switch components without substrate-diode conduction, etc. These MOSFETs work well, but if they become “unobtanium” we suggest you modify the circuit to use JFETs, in the manner of Figure 5.5.

#### 4.5.7 Differentiators

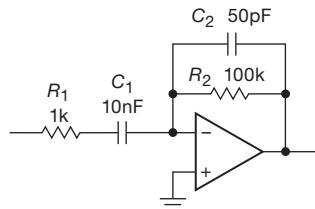
Differentiators are similar to integrators, but with  $R$  and  $C$  reversed (Figure 4.68). Because the inverting input is at ground, the rate of change of input voltage produces a current  $I = C(dV_{\text{in}}/dt)$  and hence an output voltage

$$V_{\text{out}} = -RC \frac{dV_{\text{in}}}{dt}. \quad (4.7)$$

Differentiators are bias stable, but they generally have problems with noise and instabilities at high frequencies because of the op-amp’s high gain and internal phase shifts. For this reason it is necessary to roll off the differentiator action at some maximum frequency. The usual method is shown in Figure 4.69. The choice of the rolloff components  $R_1$  and  $C_2$  depends on the noise level of the signal and the bandwidth of the op-amp, with larger values providing greater stability and less noise, at the expense of differentiator bandwidth. A minimum recommended value for  $R_1$  is given by  $R_1 = 0.5\sqrt{R_2/C_1f_T}$ ;  $C_2$  can be added for further noise reduction, with a starting value of  $C_2 \approx C_1R_1/R_2$ . At high frequencies ( $f \gg 1/2\pi R_1 C_1$ ) this circuit becomes an integrator because of  $R_1$  and  $C_2$ . We’ll explain what’s going on here in more detail in §4.9.3.



**Figure 4.68.** Op-amp differentiator (noisy, probably unstable!).



**Figure 4.69.** Adding  $R_1$  and  $C_2$  stabilizes the basic op-amp differentiator (consisting of  $C_1$ ,  $R_2$ , and the op-amp); they also reduce high-frequency noise.

#### 4.6 Op-amp operation with a single power supply

Op-amps don't require  $\pm 15$  volt regulated supplies. They can be operated from split supplies of lower voltages<sup>35</sup> or from unsymmetrical supply voltages (e.g., +12 and -3), as long as the total supply voltage ( $V_+ - V_-$ ) is within specifications (see Table 4.1 on page 245 for generic values, and Tables 4.2a,b on pages 271–272 for specific parts). Unregulated supply voltages are often adequate because of the high "power-supply rejection ratio" you get from negative feedback (for the 411 it's 90 dB typ). But there are many occasions when it would be nice to operate an op-amp from a single supply, say +9 volts. This can be done with ordinary op-amps by generating a "reference" voltage above ground, if you are careful about minimum supply voltages, output-swing limitations, and maximum common-mode input range.

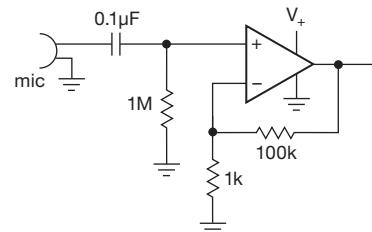
In many cases, however, you can simplify these circuits by taking advantage of a class of op-amps designed for single supply operation. With characteristic directness, engineers call these "single-supply op-amps." Their common feature is that both their input common-mode range and their output swing extends to the negative supply rail (i.e., ground, when run from a single positive supply). A subclass of these can swing their outputs to *both* supplies ("rail-to-rail outputs"), and some of those even permit input swings to both rails ("rail-to-rail I/O"). Keep in mind, though, that operation with symmetrical split supplies should be considered the normal op-amp technique for most applications.

#### 4.6.1 Biasing single-supply ac amplifiers

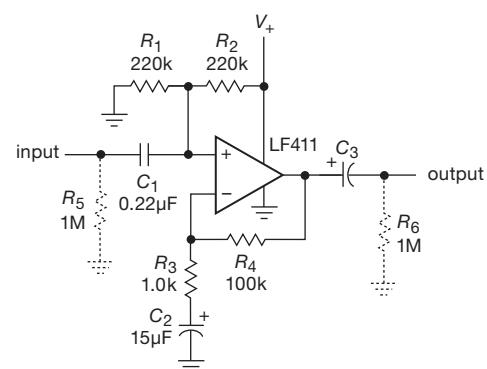
For a general-purpose op-amp like the 411, the inputs and output can typically swing to within about 1.5 volts of either supply. With  $V_-$  connected to ground, you can't have either of the inputs or the output at ground; that is, it won't work properly if you drive the inputs to ground, and it simply cannot swing its output to ground.

Thus one reason why the circuit in Figure 4.70 won't work is that the ac-coupled low-level signal from the microphone is centered on ground, where the op-amp will not work. But even if the op-amp's input common-mode range included the negative rail (ground, here), we'd still be in trouble, because in this circuit the amplified output would also be centered on ground (so that it would have to swing both above and below ground). It is important to understand that this problem with the output cannot be solved in this manner – an op-amp simply cannot swing beyond its supply rails. Even an op-amp with rail-to-rail inputs *and* outputs would not work.

**Exercise 4.19.** Draw a sketch of the output waveform from the circuit of Figure 4.70, when driven with a 10 mV input sinewave, assuming that the op-amp is of the special class with rail-to-rail inputs and outputs.

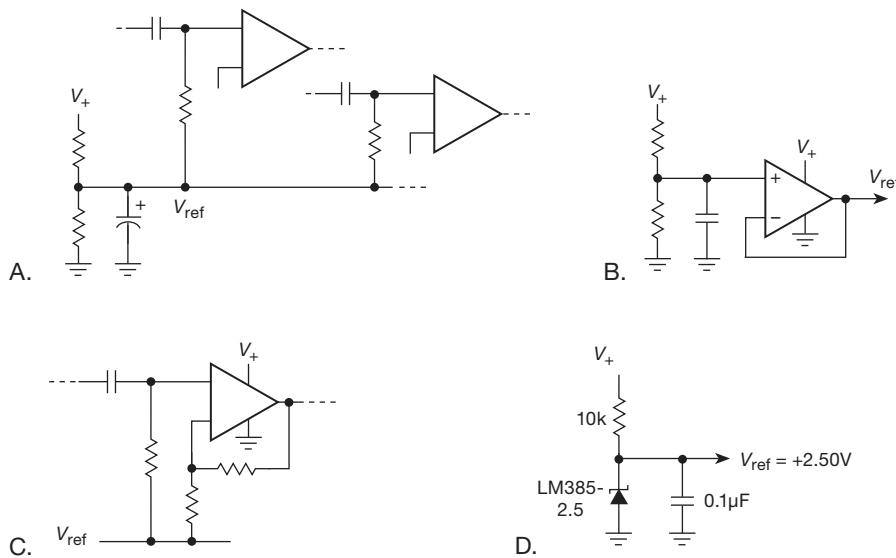


**Figure 4.70.** Defective single-supply microphone amplifier.



**Figure 4.71.** A reference voltage at  $\frac{1}{2}V_+$  (created by divider  $R_1R_2$ ) allows single-supply operation with an ordinary op-amp.

<sup>35</sup> In a world of lower and lower operating voltages, op-amps that can work at  $\pm 15$  V are now called "high-voltage op-amps."



**Figure 4.72.** Biasing schemes for single-supply operation A. Common reference (also known, confusingly, as a “virtual ground”) for multiple stages; note bypass capacitor. B. Follower generates low-impedance reference. C. The reference can serve as the return path for feedback, with significant signal currents. D. Zener-type fixed voltage reference.

### A. Reference voltage

One solution is to generate a *reference voltage* somewhere between ground and the positive supply (e.g. at half of  $V_+$ ) to bias the op-amp for successful operation (Figure 4.71). This circuit is an audio amplifier with 40 dB gain. Choosing  $V_+=12\text{ V}$  and  $V_{ref}=0.5V_+$  gives an output swing of about 9 volts pp before the onset of clipping. Capacitive coupling is used at the input and output to block the dc level, which equals  $V_{ref}$ . The optional resistors should be used if this circuit connects to the outside world; they ensure that there is no dc voltage at the input and output, which prevents loud clicks and pops when external stuff is connected.

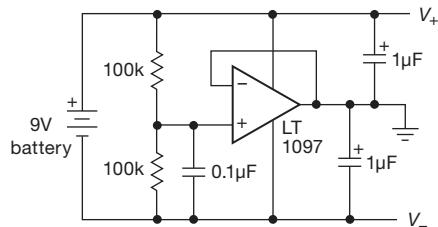
The reference voltage can be generated at the op-amp input with a simple resistive divider, as shown. If the circuit requires several op-amp stages, it is simpler to generate a common reference, with a single bias resistor to each stage, as in Figure 4.72A. Be sure to bypass the reference, to prevent signal coupling. You can also buffer the reference with a follower (Figure 4.72B), which is particularly useful if any significant dc or signal currents flow through that path, as in Figure 4.72C. Note that the follower can be any ordinary op-amp, because it is operating with a mid-supply signal. In this circuit the reference voltage doesn't have to be half of the supply voltage; it may be best to split the supply unsymmetrically, to allow maximum signal swing. In some instances it may be preferable to put it at a fixed voltage from one rail, using an IC zener-like fixed

reference, as in Figure 4.72D; that rail is then a regulated supply with respect to the common reference.

Contemporary circuit design is moving to lower supply voltages, often in the form of a single positive supply. For operation with a single +5 V supply, for example, a conventional op-amp like the 411 simply will not do: Not only can its outputs not swing typically closer than 1.5 V to the supply rails; in fact, it is not even specified for operation from a total supply voltage of less than 10 V. So for such circuits you should use op-amps designed for low-voltage operation. These are often called “single-supply” op-amps and come in several forms, some of which include specified operation of both inputs and outputs down to the negative rail; others feature output swings to both *rails*, of which a subset permits both inputs and outputs to go to both rails. We'll deal with these shortly, in §4.6.3.

### B. Supply splitter

The circuit in Figure 4.72C suggests a different approach to operation with a battery. Instead of piping a line called  $V_{ref}$  around as a signal common, with the negative battery terminal called ground, why not ground the “reference” output, effectively splitting the single supply into a positive-negative pair? This is a common technique in battery operated equipment, and is shown in Figure 4.73. The battery voltage is split by the resistive divider, which feeds a follower to generate a low-impedance common voltage. To



**Figure 4.73.** Op-amp split-supply generator. A follower generates a low-impedance mid-battery output voltage, which becomes circuit ground.

the outside world that common voltage is “ground,” with both ends of the battery floating.

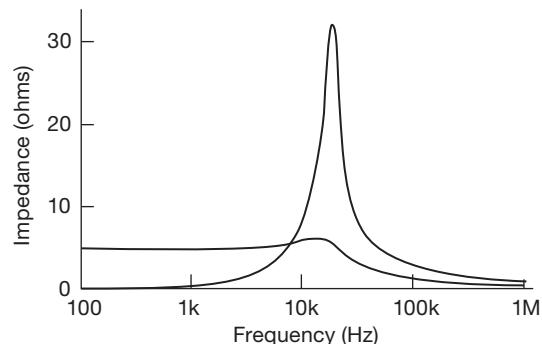
The output should be bypassed, as always, to maintain low-impedance supply rails, relative to ground, at signal frequencies. That is necessary because ground is usually the common return for filters, biasing networks, loads, etc. Look at almost any normal split-supply circuit and you’ll find dc and signal currents flowing into and out of ground.

This raises an interesting problem, which we discuss in detail in Chapter 4x and in §9.1.1C, namely that the op-amp’s output resistance, in combination with the bypass capacitor, creates a lagging phase shift at high frequencies that can cause the feedback loop to go into oscillation. Some op-amps are designed to circumvent this problem, for example the LT1097 shown in the figure (whose datasheet states that it is stable with any capacitive load). Even so, this circuit exhibits a peak in its output impedance versus frequency (Figure 4.74), and a related effect, namely, a ringing transient with that same characteristic frequency (Figure 4.75); you can think of these effects as the not-quite-banished ghost of an oscillation. As the figures show, a small series damping resistor at the op-amp’s output (Figure 4.76A) effectively stops this resonant behavior, at the expense of an increase in dc output impedance.

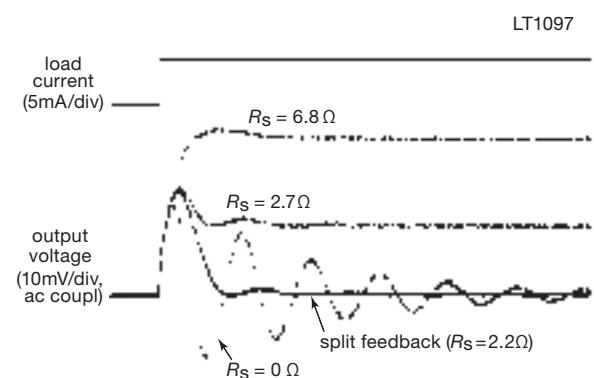
If increased output impedance is undesirable (which often it is not), another approach is to take “slow” feedback downstream of the damping resistor (which preserves accurate dc performance, i.e., low dc output impedance), with a parallel “fast” feedback path from the upstream side (Figure 4.76B) to prevent ringing. You can see the result in Figure 4.75, where we used  $R_1=2.7\Omega$ ,  $R_2=10k$ , and  $C=2.7\text{nF}$ : the initial transient looks just like that with a  $2.7\Omega$  damping resistor, but then returns to the correct dc level because dc feedback is taken from the point of load. A third possibility is to “overcompensate” the op-amp, for which the LT1097 provides hospitality in the form of a convenient “overcomp” pin; adding a capacitor to ground from this

pin increases the phase margin by shifting the dominant pole downward in frequency (§4.9).

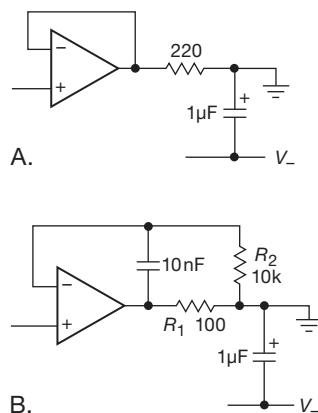
There’s a nice integrated solution from Texas Instruments, the TLE2425 and TLE2426 “rail-splitter” ICs. These come in a convenient 3-terminal TO-92 (small transistor) package, draw less than 0.2 mA quiescent current, are stable into any capacitive load greater than  $0.33\mu\text{F}$ , and can source or sink an unbalanced current of 20 mA (Figure 4.77). The TLE2426 splits the rails 50% with an internal resistive divider, whereas the TLE2425 uses an internal voltage reference to put the output common 2.50 V above the negative rail.



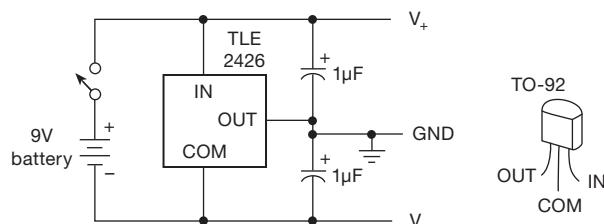
**Figure 4.74.** One effect of the capacitive load is a bump in the output impedance, which is greatly reduced with a  $5\Omega$  damping resistor; see text.



**Figure 4.75.** Measured output voltage transient of the rail-splitter circuit of Figure 4.76A caused by a 4.5 mA load current step, with several values of series damping resistor. The latter eliminates ringing, at the expense of dc output impedance. An alternative is the “split-feedback” scheme of Figure 4.76B. Scales: 5 mA/div and 10 mV/div; 40  $\mu\text{s}/\text{div}$ .



**Figure 4.76.** Stabilizing the split-supply generator: A. Decoupling resistor, B. Decoupling resistor with fast and slow feedback paths.



**Figure 4.77.** Integrated 3-terminal rail splitter.

#### 4.6.2 Capacitive loads

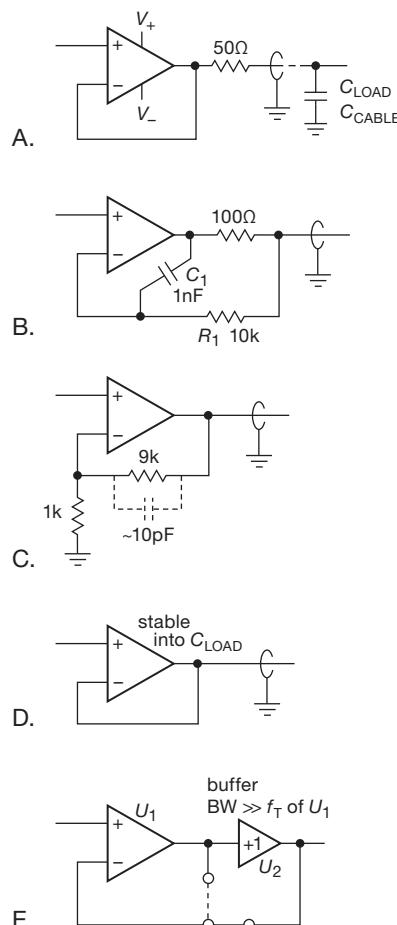
This particular example of a supply-splitter illuminates a more general problem, namely, the effect of capacitive loading at the output of *any* op-amp circuit. Although we'll deal with this further in the advanced Chapter 4x, it's important to appreciate now the causes, and cures, because it can cause mischief in even the simplest of op-amp circuits.

Let's say you build a little box, with some op-amps inside and the output(s) brought out through the ever-popular BNC panel connectors. It's easy to forget that something like a length of shielded cable – say a 2 m BNC cable going from an output connector to some other instrument – has plenty of capacitance: the standard RG-58 shielded cable patch cords have 100 pF per meter (see Appendix H). So an innocuous connecting cable alone loads the op-amp's output with 200 pF. That's sometimes enough to make an op-amp follower oscillate (we rig up just such a demonstration in our circuit design class, where an LF411 follower screams loudly when asked to drive 8 ft of cable). And even if it doesn't break into oscillation, it will likely exhibit response peaks at high frequencies, evident as overshoot and ringing.

The causes are the same as with the supply splitter: the capacitive load creates a lagging phase shift, and it's within the feedback loop.<sup>36</sup>

And the possible cures are the same (Figure 4.78); taking the figure's circuits in order (A–E):

- You can add a small series resistor (perhaps 25–100 Ω) at the op-amp's output, outside the feedback loop. (It's quite common to see a 50 Ω output resistor, which forms a matched source to "50 Ω cable"; see Appendix H.) This is OK, and easy; but it does mean that feedback does not act on the actual output signal, which may be significant with nasty loads, or at high frequencies, etc.



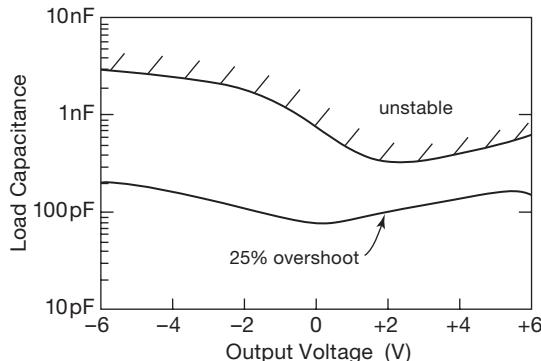
**Figure 4.78.** Driving capacitive loads.

<sup>36</sup> You can also think of this as an effect of the op-amp's inductive-like output impedance (see Figure 4.53) combining with the capacitance to form a resonance, with all its phase shifts, again within the feedback loop.

- You can split the feedback loop, as shown, so that feedback comes directly from the op-amp’s output at high frequencies, where instability lurks. And at lower frequencies the feedback accurately controls the signal seen by the load. This is not really a compromise, because those high frequencies are exactly where the thing would oscillate anyway if you were to allow feedback from the load.
- You can reduce the loop gain, for example by increasing the closed-loop gain, to regain stability.
- You can seek an op-amp that guarantees stability into the range of load capacitances you expect. Many op-amps provide good data in the form of plots of “Stability versus Capacitive Load.” Figure 4.79 shows an example, taken from the datasheet for the LMC6482.
- You can add a unity-gain buffer, with its low native output impedance, either within or outside the feedback loop. If you add it inside the loop, you need to worry about phase shifts introduced by the buffer; it should have significantly higher  $f_T$  than the op-amp, and it’s often a good idea to include a  $50\text{--}100\Omega$  series resistor at the buffer’s input (not shown). You may need to rolloff the op-amp’s response with a small capacitor, as in Figure 4.87 on page 274.

### 4.6.3 “Single-supply” op-amps

As we just remarked, some op-amps are designed specifically to allow inputs and outputs to go to the negative rail. These are called “single-supply” (or “ground-sensing”) op-amps, the idea being that their negative rail is actually tied to ground. The input range actually extends slightly below ground, typically to  $-0.3\text{ V}$ . In some cases the outputs can swing also to the positive rail (“rail-to-rail output”), and a subset of these permits input swings to (and slightly be-



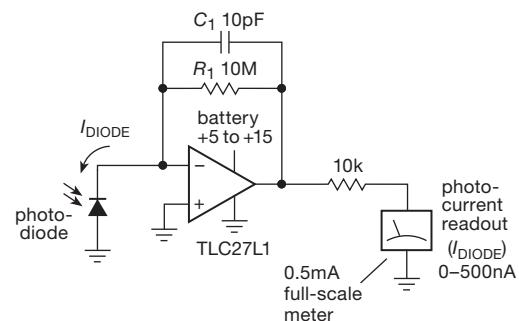
**Figure 4.79.** Stability versus capacitive load for a LMC6482 op-amp follower with  $R_{load}=2\text{k}$  and  $\pm 7.5\text{ V}$  supplies.

yond) both rails (“rail-to-rail input”). Linear Technology has introduced an exotic new twist – op-amps that permit input swings well beyond the positive rail (they call them “Over-The-Top™” amplifiers).

These amplifiers can simplify single-supply circuits because you don’t need a midsupply reference, rail splitter, etc. But you have to remember that the output cannot go *below* ground – so you can’t build an audio amplifier like Figure 4.70, whose output would need to swing both sides of ground. Before looking at the characteristics of these op-amps in more detail, let’s look at a design example.

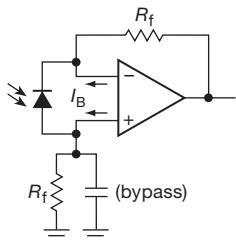
#### A. Example: single-supply photometer

Figure 4.80 shows a typical example of a circuit for which single-supply operation is convenient. We discussed a similar circuit earlier under the heading of current-to-voltage converters (and we will go further in Chapter 4x). Because a photocell circuit might well be used in a portable light-measuring instrument, and because the output is known to be positive only, this is a good candidate for a battery-operated single-supply circuit.  $R_1$  sets the full-scale output at 5 volts for an input photocurrent of  $0.5\text{ }\mu\text{A}$ . The small feedback capacitor is added to ensure stability, as we’ll explain in §4.9.3. No offset voltage trim is needed in this circuit, since the worst-case untrimmed offset of  $10\text{ mV}$  corresponds to a negligible 0.2% of full-scale meter indication. The TLC27L1 is an inexpensive micropower ( $10\text{ }\mu\text{A}$  supply current) CMOS op-amp with input and output swings to the negative rail. Its low input current ( $0.6\text{ pA}$ , typ, at room temperature<sup>37</sup>) makes it good for low-current applications like this. If you choose a bipolar op-amp for this kind of low signal-current circuit, better performance at



**Figure 4.80.** Single-supply photometer.

<sup>37</sup> Usually taken as  $25^\circ\text{C}$  on datasheets. This is a bit warmer than typical office or lab space ( $77^\circ\text{F}$ , in the King’s units), but you can rationalize that choice by saying that it allows for some heating inside the electronics enclosure.



**Figure 4.81.** Photodiode amplifier with simple bias current cancellation.

low light levels results if the photodiode is connected as in the circuit shown in Figure 4.81.

It's worth noting that the "current budget" of this circuit is dominated by the output current that drives the meter, which can go as high as  $500\ \mu\text{A}$ . It's easy to overlook a point like that, blithely assuming that the battery need provide only the op-amp's  $10\ \mu\text{A}$  quiescent current. At  $10\ \mu\text{A}$  a standard 9 V battery lasts 40,000 hours (5 years), whereas with continuous operation at  $500\ \mu\text{A}$  it would last a month.

### B. Single-supply op-amp innards

It's helpful to look at the circuitry of a typical single-supply op-amp, both to understand how these types achieve operation to one or both rails, and also to appreciate some of the subtleties and pitfalls of designing them into your circuits. Figure 4.82 is a simplified schematic of the very popular TLC270 series of CMOS single-supply op-amps. The input stage is a *p*-channel MOSFET differential amplifier with current-mirror active load. The use of enhancement-mode *p*-channel input transistors lets the inputs go to the negative rail (and a bit beyond, until the omnipresent input protection diodes begin to conduct), but prevents input operation to the positive rail (because there would be no forward gate-source voltage).

Unlike the classic conventional op-amp with its push-pull follower output stage (Figure 4.43), this output stage is unsymmetrical: an *n*-channel follower  $Q_6$  for the pullup and another *n*-channel common-source amplifier  $Q_7$  for the pull-down. That's done because a follower at  $Q_7$  (which would have to be *p*-channel) could not pull all the way down, given that its lowest gate drive voltage is ground. This unsymmetrical output requires the common-source driver  $Q_5$  for  $Q_6$ 's gate, with matching threshold voltages for  $Q_5$  and  $Q_7$  to set the output-stage quiescent current. The feedback capacitor  $C_{\text{comp}}$  is for frequency compensation (see §4.9.2). This output stage can saturate at ground, with an impedance of  $Q_7$ 's  $R_{\text{ON}}$ ; but it can't reach  $V_+$ , because  $Q_6$  is an *n*-channel MOSFET follower.

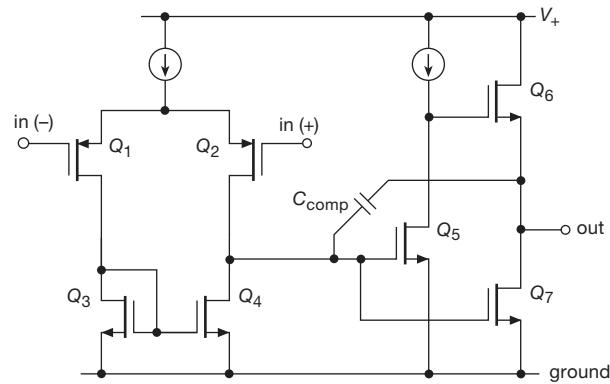
**Exercise 4.20.** What sets the source voltage of  $Q_1$  and  $Q_2$  when the inputs are approximately at ground? And what determines the high end of the input range? Why is the latter always below  $V_+$ ?

**Exercise 4.21.** What sets the maximum positive voltage to which  $Q_6$  can pull the output, assuming the op-amp is lightly loaded?

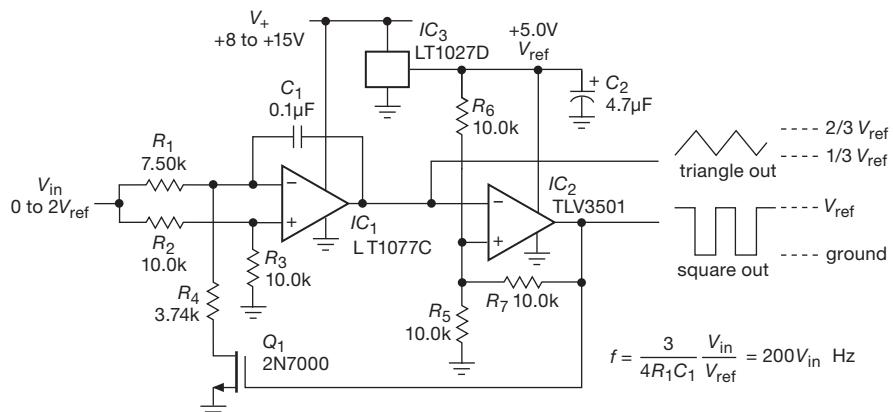
This same output-stage structure – follower pullup with amplifier pull-down – can be built with bipolar transistors; an example is the popular LT1013/LT1014 single-supply dual-quad op-amps, improved variants of the classic LM358/LM324 op-amps. A note of caution: don't make the mistake of assuming that you can make *any* op-amp's output work down to the negative rail simply by providing an external current sink. In most cases the circuitry driving the output stage does not permit that. Look for explicit permission in the datasheet!

One way to achieve rail-to-rail outputs – i.e., operation to *both* supply rails – is to replace the *n*-channel follower pullup  $Q_6$  in Figure 4.82 with a *p*-channel common-source amplifier; then each transistor can saturate to its respective rail. This requires some driver circuitry changes, of course. An analogous circuit can be built with bipolar transistors – common emitter *pnp* pullup and *npn* pull-down. Contemporary examples include the CMOS TLC2270, LMC6000, and MAX406 series, and the bipolar LM6132, LT1881, and MAX4120 series. As we'll see in Chapter 4x, there are other ways to make single-supply and rail-to-rail outputs. These amplifiers differ in important ways, and you must watch out for misleading statements about output swing to the negative rail (ground).

These output stages are pretty straightforward, and not surprising. But they don't generalize to the *input* stage. How, indeed, can you possibly achieve rail-to-rail input capability? To complete the picture without going into any



**Figure 4.82.** Simplified schematic of the TLC271-series single-supply op-amp.



**Figure 4.83.** Precision voltage-controlled waveform generator.

detail, the trick is to design an amplifier with two independent input stages, one *p*-channel (or *pnp*) and the other *n*-channel (or *npn*). We talk about them, and some other nifty tricks (such as putting on-chip voltage generators to create bias supplies beyond the rails, in combination with a conventional op-amp like Figure 4.43), in Chapter 4x. Single-supply op-amps are indispensable in battery-operated equipment.

#### 4.6.4 Example: voltage-controlled oscillator

Figure 4.83 shows a clever circuit, borrowed from the application notes of several manufacturers. IC<sub>1</sub> is an integrator, rigged up so that the capacitor current ( $V_{in}/15k$ ) changes sign, but not magnitude, when  $Q_1$  conducts. IC<sub>2</sub> is connected as a Schmitt trigger, with thresholds at one-third and two-thirds of  $V_{ref}$ . The *n*-channel MOSFET  $Q_1$  is here used as a switch, pulling the bottom side of  $R_4$  to ground when IC<sub>2</sub>'s output is HIGH and leaving it open-circuited when the output is LOW.

A nice feature of this circuit is its operation from a single positive supply. The TLV3501 is a CMOS comparator with rail-to-rail output swing, which means that the output of the Schmitt goes all the way from  $V_{\text{ref}}$  to ground; this ensures that the thresholds of the Schmitt don't drift, as they would with an op-amp of conventional output-stage design, with its ill-defined limits of output swing. In this case the result is stable frequency and amplitude of the triangle wave. Note that the frequency depends on only the ratio  $V_{\text{in}}/V_{\text{ref}}$ ; this means that if  $V_{\text{in}}$  is generated from  $V_{\text{ref}}$  by a resistive divider (made from some sort of resistive transducer, say), the output frequency won't vary with  $V_{\text{ref}}$ , only with changes in resistance. This is another example

of ratiometric techniques; circuit designers like to use this trick to minimize dependence on power-supply voltages.

Some additional points.

- Both the frequency conversion coefficient and the output swing amplitude are set by the reference voltage that powers IC<sub>2</sub> ( $V_{\text{ref}}$ ), in this case a precise and stable +5.00 V provided by the 3-terminal voltage reference IC<sub>3</sub>. This voltage could be left unregulated if the control voltage is arranged to be proportional to it, as described above. The output amplitude would, however, still be dependent on that supply rail. The solution just shown is preferable.
  - The integrator op-amp, IC<sub>1</sub>, is a “precision” op-amp, with a maximum offset voltage of 60  $\mu$ V. It was chosen to provide accurately proportional frequency down near zero volts input. You can think of this instead in terms of *dynamic range* of the frequency control: input offset voltage in the integrator op-amp produces an error in frequency equivalent to a value of  $V_{\text{in}}$  of twice that offset voltage (because of the divider  $R_2R_3$ ); to say it another way, at an input voltage  $V_{\text{in}} \approx 2V_{\text{OS}}$ , the output frequency will be in error by 100% (it could be as large as twice the programmed frequency and as low as zero). So the ratio of maximum to minimum frequency is roughly equal to  $V_{\text{ref}}/V_{\text{OS}}$ . The LT1077C shown in the figure provides a dynamic range of nearly 100,000:1 (the ratio  $V_{\text{ref}}/V_{\text{OS}} = 5\text{ V}/60\text{ }\mu\text{V}$ ).
  - The integrator op-amp must operate down to zero volts input; i.e., it must be a “single-supply” (or “ground-sensing”) op-amp, of which the LT1077C is an example.
  - Input current  $I_B$  of the integrator op-amp also causes an error, most serious when the control voltage  $V_{\text{in}}$  is near zero volts. The LT1077C has well-matched inputs with  $I_B(\text{max})=11\text{ nA}$ , which causes a worst-case error

equivalent to about  $30\ \mu\text{V}$  of unbalance when it flows through the network of unequal input resistors. This is smaller than the error contribution that is due to worst-case  $V_{OS}$ ; the combination results in a worst-case equivalent error of  $90\ \mu\text{V}$ , or a dynamic range (untrimmed) of 50,000:1. The fact that offset voltage effects dominate over bias current effects is no accident: that is why the resistor values  $R_1$ – $R_4$  were chosen as small as they are (and the capacitor value  $C_1$  was then chosen to produce the desired frequency range).

- The LT1077C could be trimmed to extend the dynamic range; ultimately it is *drift* in  $V_{OS}$  and  $I_B$  (over time and temperature) that set the circuit's overall stability near zero frequency.
- The TLV3501 is an unusually fast (4.5 ns) comparator with rail-to-rail output swing. However, its supply voltage is limited to +5.5 V maximum. If you wanted to run that portion of the circuit at a higher voltage, you could substitute a fast rail-to-rail op-amp like the CA3130. The latter part has been around a long time and is nearing extinction;<sup>38</sup> but it excels in speed for a low-power op-amp because it is *uncompensated* (see §4.9.2B). It would not be suitable for the input op-amp, however, because it is not stable as an integrator, for reasons we will see shortly. It also has large input offset voltage.
- Another possibility is to replace the Schmitt trigger circuit with a CMOS 555-type timer IC, for example, an ICL7555. These have stable input thresholds at one-third and two-thirds of the supply rail, and rail-to-rail fast output swing.
- Switch alternatives: IC switches like the SD210 or 74HC4066 (the latter belongs to the 74HC family of digital logic) could replace the discrete MOSFET  $Q_1$ ; their lower capacitance would improve operation at high frequencies.
- Another possibility, if power consumption matters more than maximum frequency or dynamic range, is to use low-power CMOS rail-to-rail op-amps for both ICs, for example, a TLC2252 dual op-amp (35  $\mu\text{A}$  per channel). In this case scale up the resistor values, particularly in the input stage, because CMOS op-amps have negligible input current for this application.
- If the use of a dual op-amp in a single package seems particularly appealing, then a good overall choice is the bipolar LM6132, with rail-to-rail inputs and outputs and a slew rate of 14  $\text{V}/\mu\text{s}$ ; in the same family you can get faster op-amps (LM6142, LM6152), at the cost of higher input and supply currents.

- An elegant single-IC solution is the use of a combination op-amp–comparator–reference IC like the MAX951. We looked around for a way to use such a chip here, but, alas, we couldn't squeeze the excellent performance of Figure 4.83 out of any of the combination chips currently available, nor from special-purpose timers like the LTC699x-series (§7.1.4B). This illustrates the circuit performance advantage you get if you can combine the best available ICs for the given task, rather than having to accept a pre-assembled combination.

**Exercise 4.22.** Derive the expression for output frequency shown in Figure 4.83. Along the way, verify that the Schmitt thresholds and integrator currents are as advertised.

#### 4.6.5 VCO implementation: through-hole versus surface-mount

Traditionally, electronic components were made with wire leads sticking out the ends (e.g., “axial-lead” resistors and capacitors), or rows of pins sticking down (e.g., ICs with DIP – “dual in-line” – packaging). Contemporary practice has shifted strongly toward “surface-mount” components, in which the connections are made directly to contacts on a ceramic or plastic package. See, for example, the photographs of resistors in Chapter 1 (Figure 1.2), of op-amps earlier in this chapter (Figure 4.1), or of small logic (Figure 10.23) in Chapter 10.

The *good news* is that surface-mount technology (SMT) lets you make smaller gadgets; and it is better *electrically* as well, because of reduced inductances in the smaller packages.

The *bad news* is that SMT makes it difficult to wire up a circuit on the spur of the moment on a prototype “breadboard” (of either the solder-in or plug-in style), an exercise that is fast and easy with through-hole components. The problem is compounded by the fact that many new high-performance components (e.g., op-amps) are available *only* in surface-mount packages.

In a nutshell, your choices boil down to (a) sticking with through-hole components (if you can get the parts you need) and enjoying the easy prototyping and ability to build a one-off gadget quickly; (b) going with the flow, and using mostly SMT components, laying out a printed circuit board for each circuit you want to build; or (c) trying to retain the best of both worlds by prototyping with through-hole components, where available, and using SMT adapters (or “carriers”) for the SMT components that you cannot get in through-hole packages. The latter are tiny circuit boards on which you solder an SMT component, whose leads connect

<sup>38</sup> See “Here Yesterday, Gone Today” on page 273.

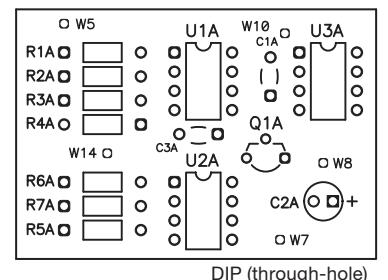
to a row of pins, producing a faux through-hole component. We've struggled with this whole business and have concluded that this last option, though attractive in principle, is fast fading away, because of the decreasing availability of through-hole components.

To give a glimpse of the tradeoffs, we laid out the voltage-controlled oscillator (VCO) circuit of Figure 4.83 on printed-circuit boards, exploring the alternatives of (a) through-hole components, (b) relatively large SMT components, and (c) small SMT components. Figure 4.84 shows them at actual size; we've shown only the component outlines and "pads" (metal foil patterns that make the connections to the components). For the through-hole board we used standard DIP op-amp and comparator and axial-lead 1/4-watt resistors; for the large SMT we used SOIC-8 op-amp and comparator and 0805 SMT resistors; and for the small SMT we used SOT-23 op-amp and comparator and the smaller 0603 resistors.<sup>39</sup> The latter is 4.5 times smaller than the through-hole board. And there is no penalty in performance; in fact, smaller components generally deliver somewhat better performance owing to smaller parasitic inductances.

#### 4.6.6 Zero-crossing detector

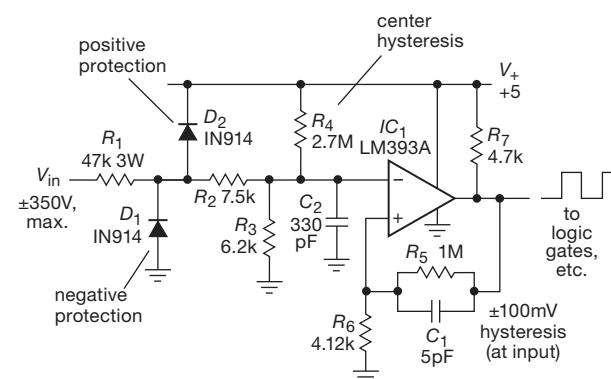
This example illustrates the use of a single-supply *comparator*, a close kin to the single-supply op-amp. Like the latter, it will operate with input signals all the way to the lower supply rail, which often is ground. The circuit, shown in Figure 4.85, generates an output square wave for use with 5 V "TTL" logic (0 to +5 V range) from an input wave of any amplitude up to 150 volts rms. The LM393 is a comparator IC (like the TLV3501 we used in the last example), specialized for this sort of application; it cannot be used as an amplifier, in the manner of an op-amp, because its internal phase shifts are not tailored ("compensated," see §4.9) to permit feedback without oscillation. It also has an "open-collector" output, which you must pull up externally to a supply rail, as shown. Its internal circuit is shown in Figure 4.86; note the overall similarity to an op-amp (Figure 4.43), with the important omission of the compensation capacitor  $C_C$ , and the lack of a "pullup" transistor at the output. We treat comparators in more detail in §12.3.

<sup>39</sup> The 4-digit designation gives the length and width, in units of 0.01" (0.25 mm); so, for example, an 0603 resistor is 0.06" × 0.03" (1.5 mm × 0.75 mm). We think that's pretty small, but the industry hasn't stopped there: standard sizes include 0402 and 0201 (the latter a mere dust mote, 0.5 mm × 0.25 mm!).



**Figure 4.84.** Printed-circuit layouts for the VCO of Figure 4.83. The use of small surface-mount components reduces the board area to 22% of the analogous board using through-hole components. Additional benefits are a greater selection of available parts, and better electrical performance.

Resistor  $R_1$ , combined with  $D_1$  and  $D_2$ , limits the input swing to  $-0.6$  volt to  $+5.6$  volts, approximately; its power rating is set by the maximum rms input voltage. Resistive divider  $R_2R_3$  is necessary to limit negative swing to less than 0.3 volt, the limit for a 393 comparator.  $R_5$  and  $R_6$  provide hysteresis for this Schmitt trigger circuit, with  $R_4$  setting the trigger points symmetrically about ground. The input impedance is nearly constant because of the large  $R_1$  value relative to the other resistors in the input attenuator. A 393 is used because its inputs can go all the way to ground, making single-supply operation simple.



**Figure 4.85.** Zero-crossing level detector with input protection.

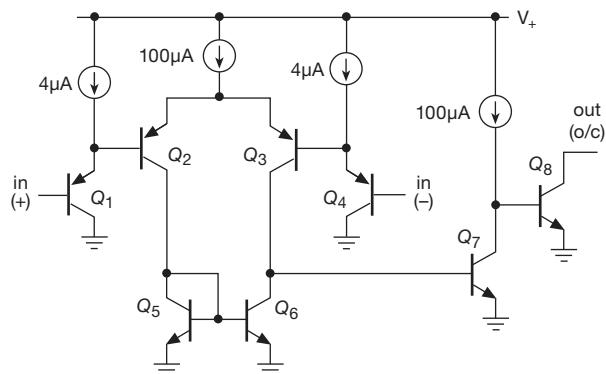


Figure 4.86. Schematic of the LM393 single-supply comparator.

**Exercise 4.23.** Verify that the trigger points are at  $\pm 100\text{ mV}$  at the input signal.

Some additional points.

- The vintage LM393 severely limits the allowable swing below ground, because the output will switch polarity if the input goes below  $-0.3\text{ V}$ , a pathology tactfully called *phase reversal* in the datasheet. That is prevented here by diode  $D_1$  and divider  $R_2R_3$ ; alternatively the low side of  $D_1$  could be biased a diode drop above ground, as in Figure 5.81. Resistor  $R_3$  could be omitted if a modern comparator like the LT1671 were used; the latter also has internal active pullup to  $+5\text{ V}$ , so you would omit pull-up resistor  $R_7$  as well.
- We intentionally set the Schmitt thresholds symmetrically around ground, but that may not be the best choice. For example, you might want output transitions accurately synchronized with the exact zero crossings of the input waveform. Omitting  $R_4$  would set the negative-going input threshold exactly at  $0\text{ V}$ ; alternatively, you could set the positive-going threshold to  $0\text{ V}$  with a properly-chosen value for  $R_4$  (test your understanding with Exercise 4.24).
- By using capacitive feedback only (omitting  $R_5$ ), you can have both thresholds at  $0\text{ V}$  with some of the benefits of hysteresis. In this case, the hysteresis is transient, with a time constant  $\tau = C_1R_6$ , by which time you are assuming the input waveform will have left the threshold region. So, for example, if you were using this circuit to sense zero crossings of a  $60\text{ Hz}$  sinewave, you might choose  $C_1 = 0.1\text{ }\mu\text{F}$  for a  $0.5\text{ ms}$  time constant (but see next item). The drawback is that you're making assumptions about the input's minimum slew rate and maximum zero-crossing frequency. You could imagine a more elaborate scheme, with additional comparators, such that the input threshold is restored to  $0\text{ V}$  after the input wave-

form passes a second higher threshold. This design challenge would yield a precise zero-crossing circuit (for both waveform slopes) with no restrictions on input speed, etc.

- Be careful if you decide to increase the value of the speed-up capacitor  $C_1$  – this capacitor causes a negative-going transient at the inverting input of the comparator, and if the capacitor is much larger than a few pico-farads you may cause phase reversal at the comparator's output (a pathology of many comparators, including the LM393). In that case it's best to use a modern comparator whose datasheet specifically brags that it is free of phase reversal; an example is the MAX989.

**Exercise 4.24.** What value of  $R_4$  in Figure 4.85 puts the positive-going input threshold at  $0\text{ V}$ ?

**Exercise 4.25.** Try designing a hysteretic circuit, with several comparators, such that *both* thresholds are precisely at  $0\text{ V}$ , under the assumption that the input waveform always travels a minimum of  $50\text{ mV}$  beyond ground before coming back.

#### 4.6.7 An op-amp table

We've collected in Table 4.2a on the facing page a representative selection of useful op-amps, including many of our favorites. You can get an idea of the price and performance of parts that are in wide use. Better yet, use this table as a starting point in your next design! More comprehensive op-amp tables are located in the chapter on precision design (Table 5.4, High-speed op-amps; Table 5.5, Precision op-amps; Table 5.6, Auto-zero op-amps), and in the chapter on noise (Table 8.3, Low-noise op-amps).

#### 4.7 Other amplifiers and op-amp types

In this first op-amp chapter we've met the “standard” split-supply op-amp, implemented variously with bipolar transistors, JFETs and MOSFETs. We've also seen examples of single-supply op-amps, some with rail-to-rail outputs (and even rail-to-rail inputs).

There are other choices, some of which we'll look at in Chapters 4x and 5. It's worth listing them here, because one or more of them may be the best solution to a design problem that looks initially like it needs an op-amp.

##### Current-feedback op-amps

These look a lot like ordinary (“voltage-feedback”) op-amps, but differ by having a low-impedance inverting input terminal that is a current summing junction. They excel in wideband circuits with moderate to high voltage gain; see the discussion in Chapter 4x.

Table 4.2a Representative Operational Amplifiers (see also Tables 5.2-5.6 and 8.3)

Part # <sup>a</sup>	#/pkg	$\frac{V_{DD}}{2}$ - $\frac{V_{SS}}{2}$ - $I_Q$ $\frac{V_{DD}}{2}$ - $\frac{V_{SS}}{2}$ - $I_Q$	cost qty 25 (\$/US)	Total Supply		$I_Q$ typ <sup>b</sup> (mA)	$f_T$ typ (MHz)	$SR$ typ	$V_{OS}$ max ( $\mu$ V)	$I_{bias}$ typ (nA)	$e_n$ typ (nV/ $\sqrt{Hz}$ )	Swing to supplies?		Comments
				min (V)	max (V)							In	Out	
LM358, 324	2,4	• • -	0.16	3	32	1	0.5	7000	45	40	-	-	-	single-supply jellybean
LT1013, 1014	2,4	• • -	1.30	4	44	0.7	0.8	0.4	40	12	22	-	-	precision single-supply
LT1077A	1	• • -	4.11	2.2	44	0.05	0.23	0.08	40	7	27	-	-	low-power bipolar, also OP193
LMC6482A, 84A	2,4	• • -	1.73	3	16	1.3	1.5	1.3	750	20fA	37	•	•	CMOS jellybean, LMC7101 SOT-23 CMOS
TLC2272A, 74	2,4	• • -	1.57	4	16	2.2	2.2	3.6	950	0.001	9	•	•	micro-power!
LMC6442A	2	• • -	2.00	2.2	16	0.002	0.01	0.004	3000	5fA	170	•	•	
LMC6041, 42, 44	1,2,4	• • -	1.48	4.5	16	0.02	0.08	0.015	3000	2fA	83	-	-	
LMC6081A, 82, 84	1,2,4	• • -	2.72	4	16	0.45	1.3	1.5	350	10fA	22	-	-	for low power, LMC6061 has $I_Q=20\mu A$
TLV2401, 02	1,2	-	1.42	2.5	16	0.0009	0.005	0.002	1200	0.1	500	•	•	pico-power, operates to $V_{CC}+5V$
LMC7101A	1	-	0.93	2.7	16	0.5	1	1	3000	0.001	37	•	•	similar to LM6482
LF411, 412C	1,2	• • -	0.72	7	36	4.5	3	13	2000	0.05	18	-	-	JFET, TI, dual cheaper than single
LF347B	4	• • -	0.58	7	36	8	3	13	5000	0.05	18	•	•	low cost JFET, 15¢ per op-amp
LT1057A, 1058	2,4	• • -	6.30	7	40	3.4	5	24	450	0.05	13	-	-	improved LF412, also see AD712
OPA727, 2727	1,2	-	2.58	4	13	4.3	20	30	150	0.085	6	•	•	e-trim CMOS
OPA376, 2376	1,2	-	2.03	2.2	7	0.76	5.5	2	25	0.2pA	7.5	-	-	e-trim CMOS
TLC227C, 274C	2,4	• • -	0.69	3	16	1.4	2.2	5.3	10mV	0.1pA	27	-	-	consider TLV27x family
OPA129B	1	• • -	10.15	10	36	1.2	1	2.5	2000	30fA	17	-	-	electrometer, mass spec, pH probe
LT1012A	1	• • -	5.11	2.4	40	0.37	0.4	0.2	25	0.025	14	-	-	low $I_B$ bipolar
LTC1050C	1	• • -	4.30	6	18	1.1	2.5	4	5	0.01	1.6 $\mu V^C$	-	-	chopper
LT1637	1	• • -	2.32	2.7	44	0.19	1	0.35	350	20	27	+	•	"over-the-top": $V_{IN}$ to $V_{EE}+44V$
LT1097	1	• • -	2.33	2.6	40	0.35	0.7	0.2	50	0.04	14	-	-	CLOAD stable, comp pin
OPA177	1	• • -	2.30	6	44	1.5	0.6	0.3	60	0.5	7	-	-	improved OP-07
OPA277, 2277, 4227	1,2,4	• • -	3.60	4	36	0.8	1	0.8	20	0.5	8	-	-	improved OP-07, see also OPA227
LM6132A, 34	2,4	• • -	2.92	2.7	35	0.8	11	14	2000	110	27	•	•	early RIO
AD797A	1	• • -	8.36	10	36	8.2	80	20	80	250	0.9	-	-	low distortion, low noise
ADA4000-1, 2, 4	1,2,4	-	1.46	8	36	1.3	5	20	1700	0.005	16	•	•	JFET
LT6220, 21, 22	1,2,4	-	2.61	2.5	12.6	0.9	60	20	200	15 <sup>h</sup>	10	•	•	low-noise JFET
OPA627A	1	• • -	24.50	9	36	7	16	55	250	0.002	5.6	-	-	fast JFET
OPA657	1	-	12.90	8	13	14	1600	700	1800	0.002	4.8	-	-	high voltage, also OPA445 miniDIP
OPA454	1	-	4.88	10	120	3.2	2.5	13	4000	0.002	35	-	-	fast VFB <sup>e</sup> , THS4021/22 decomp, G>10
THS4011, 12	1,2	• • -	5.60	9	33	7.8	290	310	6000	2000	7.5	-	-	CFB <sup>e</sup> , 100mA output current
LM771A	1	-	3.60	8	36	6.5	200	4100	4000	2700	14	-	-	CFB <sup>e</sup>
EL516	1	-	2.32	5	12.6	5	1400	6000	5000	8500	1.7	-	-	
AD8011	1	• • -	3.98	3	12.6	1	570 <sup>d</sup>	3500	5000	5000	2	-	-	low-power two-stage CFB

Notes: (a) *Italicized* part numbers have corresponding number of op-amps per package. (b) quiescent current per package, for the **boldface** part number (that with the least number of op-amps; e.g., 1mA for the LM358). (c) peak-to-peak noise voltage. (d) GBW for  $G_V=10$ . (e) VFB=voltage feedback. (h) rises to 250nA at the negative rail.

Table 4.2b Monolithic Power and High-Voltage Op-Amps<sup>a</sup>

Type	Mfg	Total supply		$I_Q$	Diff'l input <sup>b</sup>	FET	$f_T$	Slew rate	$I_{out(max)}$	$P_{diss}$	Therm lim	Prog. curr lim	Package	Cost qty 25 (\$US)	Comments
<i>low power</i>															
LME49726	TI	2.5	6	0.7	full	• -	6.8	3.7	0.35	1 <sup>u</sup>	- - -	- - -	MSOP	1.29	A
OPA567	TI	2.7	7.5	3.4	full	• -	1.2	1.2	2.2	12.5 <sup>n</sup>	• • •	-	QFN	5.53	B
OPA569	TI	2.7	7.5	3.4	full	• -	1.2	1.2	2.2	25 <sup>n</sup>	• • •	-	SO-20	7.41	B,C
AD8010	Analog	10	12.6	16	1.2	- -	230	800	0.2	1.3 <sup>u</sup>	- - -	-	SO-16	6.69	D
LM6171	TI	5	36	2.5	10	- -	100	3000	0.12	0.7	- - -	-	DIP, SO	4.27	E
LTC2057HV	LTC	4.8	65	0.8	full	• -	1.5	0.45	0.02	low	- - -	-	SO-8	3.32	F
ADA4700	Analog	10	100	1.7	full	- -	3.5	20	0.03	2.5	• - -	-	SO-8	6.00	
OPA445	TI	20	100	4.2	80	- -	2	10	0.015	0.6	- - -	-	DIP, SO-8	10.07	G
OPA454	TI	10	100	3.2	full <sup>g</sup>	• -	2.5	13	0.12	7.5 <sup>n</sup>	• - •	-	SO-8	6.09	
LTC6090	LTC	9.5	140	2.8	full	• -	12	21	0.05	15 <sup>n</sup>	• - •	-	SO, TTSOP	4.87	H
<i>medium power</i>															
L2720W	ST	4	28	10	full	- -	1.2	2	1	5	• - -	-	SO-16	1.02	I
ISL1532A	Intersil	10	30	3.5 <sup>o</sup>	full	- -	50	400	1	1	- - •	-	SSOP-20	1.43 <sup>r</sup>	J
THS3120	TI	9	33	7	4	- -	130	900	0.47	15 <sup>n</sup>	- - •	-	MSOP-8	5.57	K
LT1794	LTC	10	36	26	full	- -	200	600	0.72	25 <sup>n</sup>	• - •	-	SO-20	8.09	J
LT1206	LTC	10	36	12 <sup>o</sup>	full	- •	60	900	0.5	15 <sup>p</sup>	• - •	-	DIP, TO-220	5.88	K,L
LT1210	LTC	8	36	35	full	•	35	900	2	15 <sup>p</sup>	• - •	-	TO-220-7	8.75	K,M
L272	FSC	4	40	8	full	- -	0.35	1	1	5	• - -	-	DIP, SO-16	2.08	N
PA75	Apex	5	40	8	full	- -	1.4	1.4	2.5	19	- - -	-	TO-220-7	28.88	O
TDA7256	ST	10	50	80	full	- -	9	10	3	35	• - •	-	TO-220-11	3.42	P
LM1875	TI	16	60	70	full	- -	5.5	8	4	25	• - -	-	TO-220-5	2.75	P
OPA552	TI	8	60	7	full	• -	12	24 <sup>d</sup>	0.2	25 <sup>p</sup>	• - -	-	DIP, DDPak	5.70	Q
LM675	TI	20	60	18	full	- -	5.5	8	3	40	• - -	-	TO-220-5	4.82	R
OPA547	TI	8	60	10	full	- -	1	6	0.5	25	• • •	-	TO-220-7	9.57	S
OPA548	TI	8	60	17	full	- -	1	10	3	30	• • •	-	TO-220-11	13.22	S
OPA549	TI	8	60	26	full	- -	0.9	9	8	53	• • •	-	TO-220-11	20.65	S
OPA453	TI	20	80	4.5	full	• -	7.5	23 <sup>d</sup>	0.05	25	• - -	-	TO-220-7	5.50	T
OPA541	TI	20	80	20	full	• -	2	10	10	90	- • -	-	TO-3, SIP-11	19.28	
LM3886	TI	18	84 <sup>s</sup>	50	60	- -	8	19	11.5	75	• - •	-	TO-220-11	5.94	P
TDA7293	ST	24	120	30	30	- -	-	15	6.5	75	• - •	-	TO-220-15	5.49	P
<i>higher voltage</i>															
PA340	Apex	20	350	2.2	16	• •	10	32 <sup>k</sup>	0.06	16	• - -	-	DDPak-7	21.45	U
PA90	Apex	30	400	10	20	• •	100	300 <sup>e</sup>	0.2	18	- • -	-	SIP-12 +tab	188.00 <sup>v</sup>	U
PA15	Apex	100	450	2.0	25	• •	5.8	20 <sup>d</sup>	0.2	18 <sup>n</sup>	- - -	-	SIP-10	185.00 <sup>v</sup>	U
PA98	Apex	30	450	21	25	• •	100	1000 <sup>e</sup>	0.2	18	- • -	-	SIP-12 +tab	272.00 <sup>v</sup>	U
PA97	Apex	100	900	0.6	20	• •	1	8 <sup>e</sup>	0.01	3 <sup>n</sup>	- - -	-	SIP-12	176.00 <sup>v</sup>	V

**Notes:** (a) within categories, sorted by maximum voltage, then output current; the Apex parts are hybrid, and neither PCB nor instrument-box types are listed. (b) not to exceed total supply voltage. (c)  $P_{diss}$  with case at  $T_c = 50^\circ\text{C}$ , based on  $R_{\theta\text{JC}}$ . (d) when comp for  $G > 10$ . (e) when comp for  $G > 100$ . (g) internal JFETs limit current to 4mA. (h) see notes. (k) for  $C_C=4.7\text{pF}$ ,  $G \geq 10$ . (n) provided you can get the heat out of the package! (o) adjustable. (p) power package. (r) qty 1k. (s) 94V w/o signal. (u) to ambient. (v) unit qty; see distributor prices (and your banker) for larger qty.

**Comments:** (A) dual, RRO. (B) RRIO. (C) with current monitor. (D) video. (E) VFB with CFB. (F) auto-zero, 4 $\mu\text{V}$ . (G) has  $V_{OS}$  trim; also in TO-99. (H) dual=6091. (I) update of L272. (J) dual, ADSL driver. (K) current-feedback, CFB. (L) can drive 10nF capacitive loads. (M) 1.1A min.

(N) Fairchild's version. (O) amp+buffer. (P) audio amplifier. (Q) slower OPA551 for  $G=1$ . (R) classic workhorse. (S) current-limit adjustment with resistor or external current. (T) slower OPA452 for  $G=1$ . (U) MOSFET output. (V) "inexpensive."

### "Zero-drift" op-amps

These unusual op-amps, which include auto-zero and chopper-stabilized amplifiers, are tailored for precision ( $V_{OS}$ ) applications. They use internal MOS switches to measure and correct for input offset error. These are the only amplifiers with values of

untrimmed  $V_{OS}$  down to 5 $\mu\text{V}$  or less. See Table 5.6 on page 335.

### High-voltage, high-power op-amps

You can get op-amps with maximum output currents of 25 amps or more, or with power supply voltages to 1 kV or

### “Here Yesterday, Gone Today”

In its untiring quest for better and fancier chips, the semiconductor industry can sometimes cause you great pain. It might go something like this: you’ve designed and prototyped a wonderful new gadget; debugging is complete, and you’re ready to go into production. When you try to order the parts, you discover that a crucial IC has been discontinued by the manufacturer! An even worse nightmare goes like this: customers have been complaining about late delivery on some instrument that you’ve been manufacturing for many years. When you go to the assembly area to find out what’s wrong, you discover that a whole production run of boards is built, except for one IC that “hasn’t come in yet.” You then ask purchasing why they haven’t expedited the order; turns out they have, just haven’t received it. Then you learn from the distributor that the part was discontinued six months ago and that none is available!

Why does this happen, and what do you do about it? We’ve generally found four reasons that ICs are discontinued.

1. *Obsolescence*: Much better parts come along, and it doesn’t make much sense to keep making the old ones. This has been particularly true with digital memory chips (e.g., small static RAMs and EPROMs, which are superseded by denser and faster versions each year), though linear ICs have not entirely escaped the purge. In these cases there is often a pin-compatible improved version that you can plug into the old socket.
2. *Not selling enough*: Perfectly good ICs sometimes disappear. If you are persistent enough, you may get an explanation from the manufacturer – “there wasn’t enough demand,” or some such story. You might characterize this as a case of “discontinued for the convenience of the manufacturer.” We’ve been particularly inconvenienced by Harris’s discontinuation of their splendid

HA4925 – a fine chip, the fastest quad comparator, now gone, with no replacement anything like it. In our first edition we reported that Harris also discontinued the HA2705 – another great chip, the world’s fastest low-power op-amp, gone without a trace! Since that time, Maxim came out with the MAX402, similarly a fast low-power op-amp. Lots of us used it; then – whammo – can’t get it! Sometimes a good chip is discontinued when the wafer fabrication line changes over to a larger wafer size (e.g., from the original 3" diameter wafer to a 5" or 6" wafer).

3. *Lost schematics*: You might not believe it, but sometimes the semiconductor house loses track of the schematic diagram of some chip and can’t make any more! This apparently happened with the Solid State Systems SSS-4404 CMOS 8-stage divider chip.
4. *Upgraded* production line: Sometimes a manufacturer will replace older test equipment (which may have been working just fine) with the latest and greatest new stuff. Problem is, the programs to run the new testers aren’t finished yet. So, the wafer line *could* be making lots of chips … but there’s no way to test them. This scenario appears to have played out in the case of the magnificent OPA627, one of our all-time favorites (there was nearly a year in which you couldn’t get these puppies, but, thankfully, it’s back in production).
5. *Manufacturer out of business*: This also happened to the SSS-4404! If you’re stuck with a board and no available IC, you’ve got several choices. You can redesign the board (and perhaps the circuit) to use something that is available. This is probably best if you’re going into production with a new design or if you are running a large production of an existing board. A cheap and dirty solution is to make a little “daughterboard” that plugs into the empty IC socket and includes whatever it takes to emulate the nonexistent chip. Although this latter solution isn’t terribly elegant, it gets the job done.

more! These are specialized (and expensive) devices, extremely useful for applications such as piezo drivers, servo drivers, and so on. See Table 4.2b on the facing page for some favorites.

### Micropower op-amps

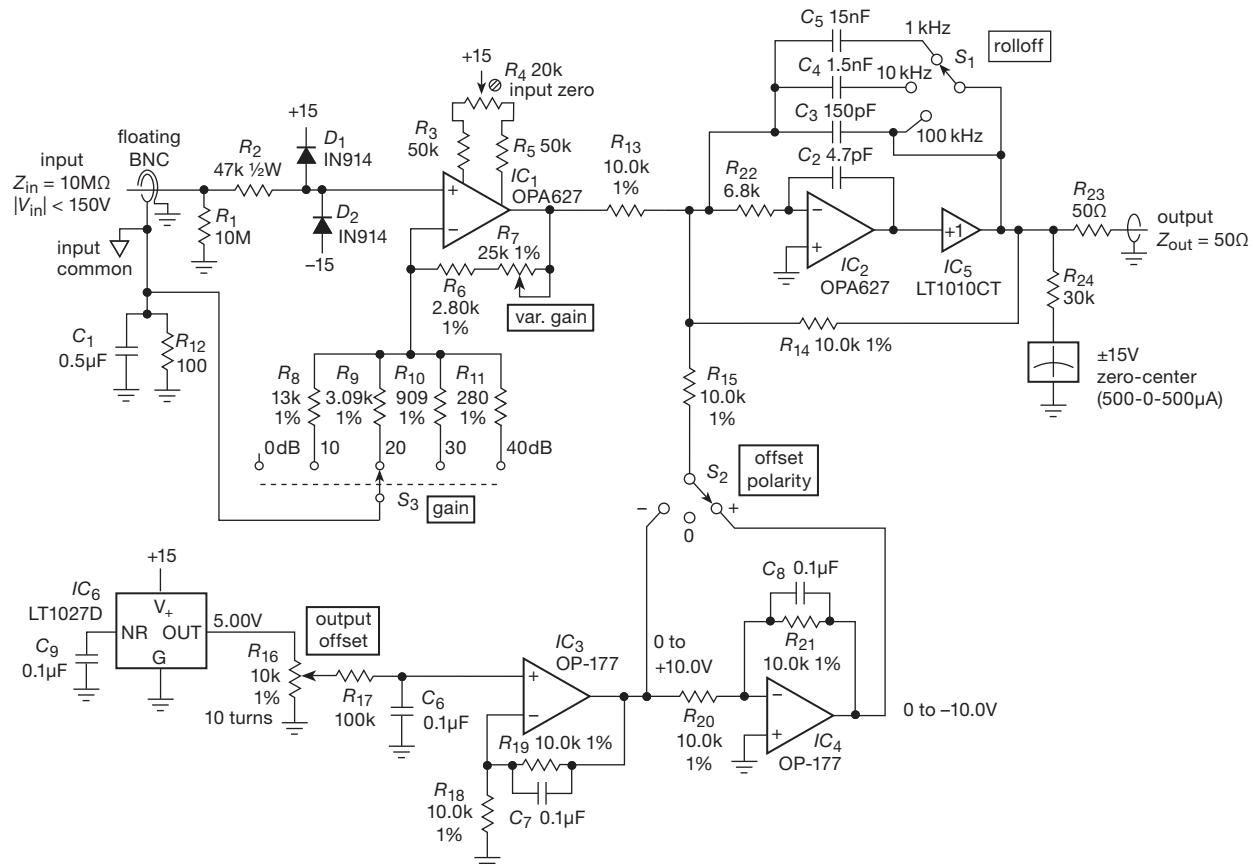
At the other end of the spectrum, you can get op-amps with quiescent currents as low as a microamp or less. These things aren’t blazingly fast – the LMC6442, with

$I_Q = 10\ \mu A$  per amplifier, has an  $f_T$  of 10 kilohertz, and a slew rate<sup>40</sup> of 0.004 V/ $\mu s$  – but they do let you run a portable instrument just about forever on a single battery.

### Instrumentation amplifiers

These are integrated differential amplifiers with settable voltage gain. They contain several op-amps internally and

<sup>40</sup> The manufacturer would *never* use “V/ $\mu s$ ” on a datasheet of such a sluggish op-amp – look for V/millisecond instead.



**Figure 4.87.** Laboratory dc amplifier with output offset. Op-amp power supply connections and bypass capacitors are not shown explicitly, a common practice in circuit schematics.

excel in stability and common-mode rejection. Instrumentation amplifiers are discussed in §5.15.

#### Video and radiofrequency amplifiers

Specialized amplifiers for use with video signals, or with communications signals at frequencies from 10 MHz to 10 GHz, are widely available as fixed-gain amplifier modules. At these frequencies you generally don't use op-amps.

#### Dedicated amplifier variants

Microphone preamps, speaker amplifiers, stepping motor drivers, and the like are available as customized ICs with superior characteristics and ease of use.

### 4.8 Some typical op-amp circuits

#### 4.8.1 General-purpose lab amplifier

Figure 4.87 shows a dc-coupled “decade amplifier” with settleable gain, bandwidth, and wide-range dc output offset.

IC<sub>1</sub> is a low-noise JFET-input op-amp with noninverting gain from unity (0 dB) to  $\times 100$  (40 dB) in accurately calibrated 10 dB steps; a vernier is provided for variable gain. IC<sub>2</sub> is an inverting amplifier; it allows offsetting the output over a range of  $\pm 10$  volts, accurately calibrated by the 10-turn pot R<sub>16</sub> by injecting current into the summing junction. C<sub>3</sub>–C<sub>5</sub> set the high-frequency rolloff, because it is often a nuisance to have excessive bandwidth (and noise). IC<sub>5</sub> is a power booster for driving low-impedance loads or cables; it can provide  $\pm 150$  mA output current.

Some interesting details: a 10 MΩ input resistor is small enough, since the bias current of the OPA627 is 10 pA (maximum, at room temperature), thus producing a 0.1 mV error with open input. R<sub>2</sub>, in combination with clamp diodes D<sub>1</sub> and D<sub>2</sub>, limits the input voltage at the op-amp to the range  $V_- - 0.6$  V to  $V_+ + 0.6$  V. With the protection components shown, the input can go to  $\pm 150$  volts without damage. The JFET-input OPA627 was chosen for its combination of low input current ( $I_B=1$  pA, typ), modest pre-

cision ( $V_{OS}=100\ \mu V$ , max), low noise ( $e_n=5\ nV/\sqrt{Hz}$ , typ), and wide bandwidth ( $f_T=16\ MHz$ , typ); the latter is needed to preserve some loop gain at the high-frequency end of the instrument (100 kHz) when running at full gain (40 dB).

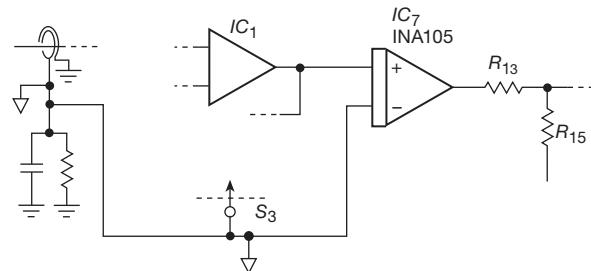
The output stage is an inverter with a unity-gain power buffer inside the feedback loop. The vintage LT1010 has plenty of slew rate, bandwidth, and muscle, with less than  $10\ \Omega$  open-loop output impedance (which of course is lowered by feedback; see §2.5.3C). Both it and the OPA627 have enough slew rate (75 V/ $\mu s$  and 55 V/ $\mu s$ , respectively) to generate a full  $\pm 15\ V$  output swing at the full 100 kHz bandwidth of the instrument. A power buffer like this is good for isolating capacitive loads from the op-amp (more on this in Chapter 4x; see also §§4.6.1B and 4.6.2); furthermore, it takes the heat when driving a hefty load, which keeps  $IC_2$  cool, an important consideration with precision (low- $V_{OS}$ ) op-amps. It takes lots of drive compared with an op-amp – up to 0.5 milliamp – but that's no problem when you're driving it with an op-amp.

The offset circuit consists of a precision LT1027 3-terminal voltage reference IC. We'll learn more about these in Chapter 9; they generate a highly stable voltage output when powered from a noncritical dc rail that is at least 2 volts higher than their specified output voltage. This particular part comes in several grades, the best of which (LT1027A) has a maximum error of 1 mV, and is guaranteed to drift less than 2 ppm/ $^{\circ}C$ ; for this application we'd save some money by choosing the inexpensive "D" grade ( $5.0\ V \pm 2.5\ mV$ ; 5 ppm/ $^{\circ}C$ ). The OP177 is a highly stable precision op-amp ( $V_{os} < 10\ \mu V$ ,  $TCV_{os} < 0.1\ \mu V/^{\circ}C$  in its best grade) that provides a stable offsetting voltage. Capacitor  $C_6$  bypasses noise on the reference voltage, and  $C_7$  and  $C_8$  reduce amplifier noise by limiting the bandwidth of the amplifiers. For a dc application like this you don't need, and don't want, lots of bandwidth. We'll talk in detail about this sort of precision design in Chapter 5.

Some additional points.

- In a circuit like this the input protection network can limit the ultimate bandwidth, because  $R_2$  forms a lowpass filter in combination with the combined input capacitance of  $IC_1$ , diode capacitance, and associated wiring capacitance. In this case the total capacitance is approximately 12 pF, which puts the 3 dB point at 300 kHz, well above the 100 kHz high-frequency limit of the instrument. To use a similar protection circuit in a wideband amplifier, you could reduce the value of  $R_2$ , put a small capacitor (47 pF, say) across it, or both. You could also use clamp diodes with lower capacitance, for example a 1N3595 or a PAD5 (see Chapter 1x).

- A really useful general-purpose laboratory amplifier should have true *differential* inputs. This is best done with an instrumentation amplifier, rather than an op-amp; see §5.15. Here we've compromised with a "pseudo-differential" configuration, in which the input common terminal (which is the return path for feedback), floating from circuit ground with a  $100\ \Omega$  resistor, is allowed to accommodate a small amount of signal from the input source. A better arrangement, though still not symmetrically differential, is shown in Figure 4.88, where a difference amplifier ( $IC_7$ ) uses the floating input common as a reference. Note the use of a chassis-isolated BNC panel connector.



**Figure 4.88.** Difference amp cancels error from signal on input common.

- In many situations it is preferable to introduce the dc offset at the input rather than at the output. Then you can change gain, without adjusting the offset, to zoom in on a portion of the input signal. This requires a much larger range of offset voltage, and other circuit changes as well. We'll see an example in Chapter 5.
- Watch out for op-amps that exhibit phase reversal when their inputs go more than 0.3 V below  $V_-$ ; in such cases a restrictive input clamp must be used to prevent negative swings below that limit. This is a common defect of many op-amps, which the excellent OPA627 does not share.
- Contemporary instrumentation usually provides for remote operation, with digital control from a computer. This circuit, however, uses mechanical controls for gain, bandwidth, and offset. You could replace the mechanical switches with analog switches, and use a DAC to generate the offset, to adapt this instrument to digital control.
- The rolloff capacitors  $C_3-C_5$  close the loop around the output amplifier pair ( $IC_2+IC_5$ ) at high frequencies, which is beneficial in terms of reducing noise. But it also promotes instability, owing to the combined phase shifts of the two amplifiers. This arrangement is still OK, though, as long as the bandwidth of buffer  $IC_5$  is much greater than that of amplifier  $IC_2$ .

But that is not the case here: the OPA627 op-amp has a unity gain bandwidth of  $f_T=16\text{ MHz}$ , at which it specifies a  $75^\circ$  phase margin. But the LT1010 buffer adds about  $50^\circ$  of additional lagging phase shift, pushing the amplifier close to instability (see §4.9 for an explanation of phase margin and stability). The solution here is to use a small feedback capacitor around the op-amp ( $4.7\text{ pF}$ ,  $C_2$ ), which closes the high-frequency feedback path directly. This rolls off its gain to unity at about  $1\text{ MHz}$ , at which frequency the buffer contributes less than  $5^\circ$  additional lagging phase shift.

**Exercise 4.26.** Check that the gain is as advertised. How does the variable-offset circuitry work? At what frequency would the slew-rate-limited output swing drop below  $\pm 15\text{ V}$ ?

#### 4.8.2 Stuck-node tracer

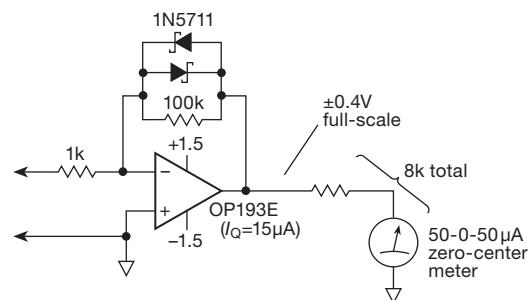
Here's a nice example of an op-amp circuit with *nonlinear* feedback. A tricky troubleshooting problem is a so-called *stuck node*, in which there is a short somewhere on a circuit board. It may be an actual short-circuit in the wiring itself, or it may be that the output of some device (for example a digital logic gate, see Chapter 10) is held in a fixed state. It's hard to find, because anywhere you look on that line, you measure zero volts to ground.

A technique that does work, however, is to use a sensitive voltmeter to measure voltage drops *along* the stuck trace. A typical signal trace on a printed-circuit board might be  $0.010''$  wide and  $0.0013''$  thick (1 ounce per square foot), which has a resistance along the trace of  $53\text{ m}\Omega/\text{in}$ . So if there's a device holding the line to ground somewhere and you inject a diagnostic current of  $10\text{ mA}$  dc somewhere else, there will be a voltage drop of  $530\text{ }\mu\text{V}$  per inch in the direction of the stuck node.

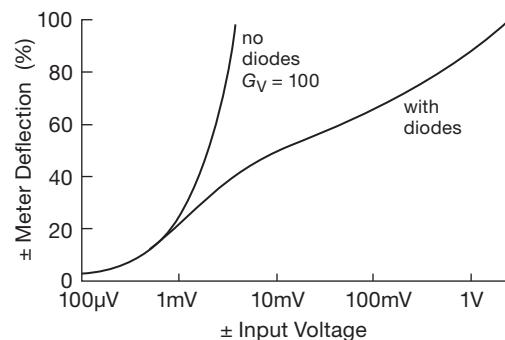
Let's design a stuck-node tracer. It should be battery powered so that it can float anywhere on the powered circuit under test. It should be sensitive enough to indicate a drop of as little as  $\pm 100\text{ }\mu\text{V}$  on its zero-center meter, with larger meter deflections for larger drops. Ideally it should have a nonlinear scale, so that even for voltage drops of tens of millivolts the meter will not go off scale. And with some care it should be possible to design a circuit that draws so little battery current that we can omit the on/off switch: 9 V batteries or AA-size cells give nearly their full shelf life of several years at continuous drain currents of less than  $20\text{ }\mu\text{A}$  (they have capacities of about  $500\text{ mAh}$  and  $2500\text{ mAh}$ , respectively).

With a floating supply provided by batteries, the simplest circuit is a high gain inverting amplifier driving a

zero-center meter (Figure 4.89). Because the input and output are both intrinsically bipolar, it's probably best to use a pair of AA cells, running the op-amp from  $\pm 1.5$  volt unregulated supplies. The back-to-back Schottky diodes reduce the gain gracefully at large output swings and prevent pegging; Figure 4.90 plots the resulting meter deflection versus  $V_{\text{in}}$ .



**Figure 4.89.** Stuck-node tracer: high-gain floating dc amplifier with nonlinear feedback.



**Figure 4.90.** Stuck-node tracer achieves large dynamic range through nonlinear feedback.

The major difficulty in this design is in achieving an input offset of less than  $100\text{ }\mu\text{V}$  while maintaining micropower current drain, all with supply voltages of just  $\pm 1.5$  volts. The OP193 is specified to operate down to  $2\text{ V}$  total supply voltage, and its output stage swings to the negative rail and to within a volt of the positive rail. In its best grade ("E" suffix) its offset voltage is  $75\text{ }\mu\text{V}$ , maximum. Its quiescent current of just  $15\text{ }\mu\text{A}$  ensures that the batteries will last their full shelf life, since that current would provide continuous operation for over 150,000 hours from a  $2500\text{ mAh}$  battery.

Some additional points.

- One subtle problem with this circuit is that an alkaline battery at the end of its life is down to about  $1.0\text{ V}$

terminal voltage; so you would have insufficient headroom to provide full-scale positive output voltage (+0.5 V), given the all-*nPN* output stage. A solution is to use a higher battery voltage (e.g., 3 V lithium cells, or multiple alkaline 1.5 V AA cells). But operation from a single pair of AA cells is an elegance worth preserving. In this case you would do better to use an op-amp with true rail-to-rail output, for example, the CMOS OPA336. The latter has a quiescent current of 20  $\mu$ A, operates down to 2.3 V total supply voltage, and has an untrimmed offset voltage of 125  $\mu$ V. Its input voltage range goes to the negative rail and to within 1 volt of the positive rail; the latter is fine here, because we have chosen an inverting amplifier configuration with both inputs at 0 V.

- We rather artificially constrained the circuit design by choosing a zero-center analog meter and then insisting on using just a pair of AA alkaline cells. In real life you'd probably be happier with an *audio* output, with the pitch increasing with input voltage drop; then you could keep your eye on the circuit as you probe around. For this job you'd probably use a simple current-controlled oscillator, built with an op-amp relaxation oscillator or a 555-type timer IC (Chapter 7); for a noncritical application you don't need the linearity and stability of the VCO we designed in Figure 4.83.
- Don't forget the "rail-splitter" techniques we discussed in §4.6.1B; you can always use those tricks to create a split plus and minus rail, for example, from a single 9 V battery. With  $\pm 4.5$  V rails, you have a much wider range of op-amps to choose from. We were forced to choose from a rather small selection that run on 2 V total supply, draw only tens of microamps supply current, and have "precision" low input offset voltage. Once you have 5 V total supply available (a 9 V battery is down to 6 V at the end of its life), there are literally hundreds of available op-amps, dozens of which run on micropower current drain and have precision low offsets. See, for example, Table 5.5 on page 320.

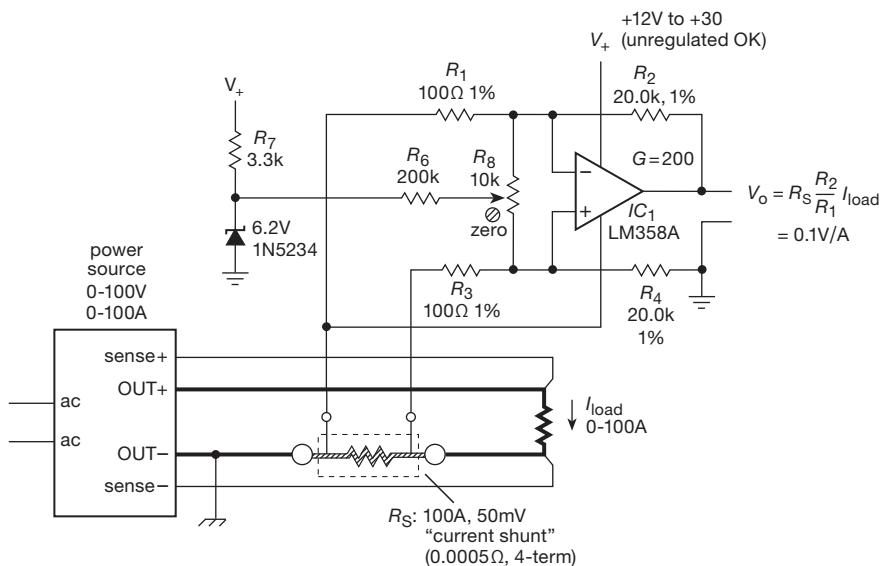
#### 4.8.3 Load-current-sensing circuit

Figure 4.91 shows a hefty (10 kW!) power supply driving a 100 amp load; the illustrated circuit provides a voltage output proportional to load current, for use with a current regulator, metering circuit, or whatever. The output current is sensed with a *current shunt*, a calibrated manganin-metal 4-terminal power resistor  $R_S$ , of resistance 0.0005  $\Omega$ , whose "Kelvin connection" of four leads ensures that the sensed voltage does not depend on a low-resistance bond to the sensing terminals (as would be the case if you tried

to do the same thing with a conventional 2-terminal resistor). The voltage drop goes from 0 to 50 mV, with probable common-mode offset caused by the effects of resistance in the ground lead (note that the power supply is connected to chassis ground at the output). For that reason the op-amp is wired as a differential amplifier, with a gain of 200. Voltage offset is trimmed externally with  $R_8$ , as the venerable LM358A doesn't have internal trimming circuitry. A zener reference with a few percent stability is adequate for trimming, because the trimming is itself a small correction (you hope!). The supply voltage,  $V_+$ , could be unregulated, since the power-supply rejection of the op-amp is more than adequate, 85 dB (typ) in this case.

Some additional points.

- Chassis ground and circuit ground would be connected together, somewhere. But there could easily be a volt or so separating circuit ground from the sensing point along the high-current negative return, because of the very large currents flowing. For that reason we connected the negative supply lead of the op-amp to the more negative end of the current-shunt output. This ensures that the common-mode voltage appearing at the op-amp input never goes below its negative rail; it is a "single-supply" op-amp, with operating common-mode range to its negative rail.
- Low offset voltage is important in this application; for example, to achieve 1% accuracy in a current measurement made at 10% of full-scale load current (i.e., a 10 A load, producing a sense voltage of 5 mV) requires an offset voltage no greater than 50  $\mu$ V! We chose the vintage LM358A for our initial design, because it costs only 20 cents. But its poor untrimmed offset (3 mV, max) necessitates external manual trimming; and its lack of external trim terminals forced us to use lots of components. The need for manual trimming might not seem important if you're just building one of these for your lab; but in production it's an extra step, requiring a test setup and procedure, as well as additional parts inventory, etc.
- So, you might choose instead the LT1006, a single op-amp that lets you trim externally with a single 10k pot. However, its improved performance ( $V_{OS}=80 \mu$ V, max, untrimmed) in the least expensive grade – 40 times better than the LM358A – means that you hardly need to trim at all. Carrying this idea further, you could choose instead the LT1077A, a single-supply op-amp with 40  $\mu$ V maximum untrimmed offset; it too can be trimmed externally.
- For the utmost in accuracy you should use a chopper-stabilized ("zero-drift") op-amp, for example the LTC1050C. It has 5  $\mu$ V maximum offset voltage in the cheapest grade (combined with sub-nanoamp input



**Figure 4.91.** High-power current-sensing amplifier.

bias current, which doesn't matter here). This op-amp includes on-chip capacitors for its chopper, and operates from a single-supply (with input common-mode range to the negative rail), just like the LM358. Its  $5\text{ }\mu\text{V}$  offset voltage corresponds to 1% accuracy at 1% of full scale; that's a dynamic range of 10,000:1, not bad for a simple circuit. See Table 5.6 on page 335 for auto-zero op-amp choices.

- Finally, an interesting design alternative is to do *high-side* current sensing. That is, the shunt is connected instead to the OUT+ power terminal. This has the advantage of keeping all the circuit grounds (power supply, and load) connected together. We'll see how to do this in the advanced Chapter 9x.

#### 4.8.4 Integrating suntan monitor

We nerds don't ordinarily go to the beach. But when we do, we like to rely on some electronics to tell us when to turn over. What we want to monitor, of course, is the *integrated* dose of tan-producing (UV-rich) sunlight.

There are many ways to accomplish this; in fact, we'll revisit this task when we turn to mixed-signal (analog + digital) electronics (in Chapter 13), and again when we're looking for nifty things to do with microcontrollers (in Chapter 15). Here we want to show how an op-amp integrator can be used to build a suntan monitor circuit.

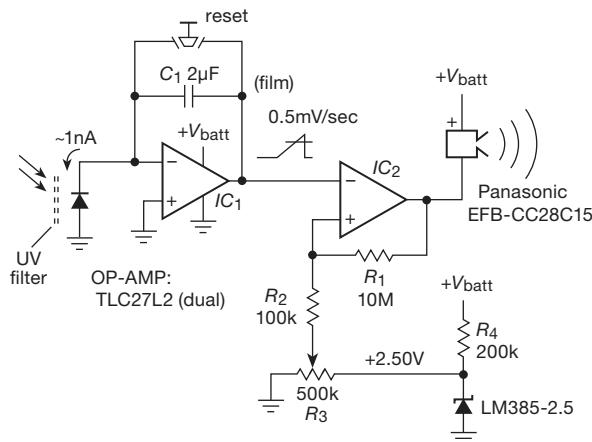
The idea is to integrate (accumulate) the photocurrent from a sensor whose output is proportional to the intensity

of tanning sunlight. We'll imagine that we have a photodiode, optically filtered to pass only the UV rays of interest, with an output short-circuit current of  $\sim 1\text{ nA}$  (nominal) in full sunlight; we'll assume that the photocurrent might range down to a tenth of this value, or so, in hazy sun.

#### A. First try: direct integration

The circuit in Figure 4.92 is a reasonable first try. It uses a single-supply CMOS micropower op-amp ( $10\text{ }\mu\text{A}$  per amplifier), powered from a 9 V battery, to integrate the (negative) photocurrent. A nanoamp produces a positive-going ramp of  $0.5\text{ mV/s}$  at the op-amp's output, which we connect to a Schmitt trigger comparator with settable positive threshold. The LM385-2.5Z micropower two-terminal (zener-type) voltage reference then gives us a range of 0 to 1.5 hours ( $\sim 5000\text{s}$ ) full sunlight equivalent (let's call it "FSE"), at which point the comparator output goes to ground, driving the piezo alarm. The latter draws  $15\text{ mA}$ , a substantial battery load, but it is *very* loud, so even a dozing nerd will quickly enough shut the thing off (via the "reset" button). This circuit draws about  $50\text{ }\mu\text{A}$  when integrating, good for about 8000 hours of operation (a 9 V battery has a capacity of  $500\text{ mAh}$  at low drain). 8000 hours is about a year, so that's a lot of tanning; the battery will die of old age first.

**Exercise 4.27.** The LM385 requires a minimum of  $10\text{ }\mu\text{A}$  of current for proper operation. What does the circuit provide, at the end of battery life (6 V)?



**Figure 4.92.** Integrating suntan monitor, first try.

**Exercise 4.28.** How much hysteresis does Schmitt trigger IC<sub>2</sub> provide? How will that affect the operation?

### B. Second try: two-step conversion

One problem with the last circuit is that the unfiltered photodiode current is at least a few microamps, in direct sunlight. Trying to cut down the light by a factor of a thousand is risky, because you get light leaks, etc., that cause large errors.

The circuit in Figure 4.93 fixes that, by first converting the photocurrent (however large its magnitude) to a voltage, then integrating that in a second stage where we can choose an input resistor to generate a current in the nanoamp range. Now, however, we've got to use split supplies. That's because whichever polarity we choose for the transresistance (current to voltage) amplifier's output (by connecting the photodiode appropriately), the subsequent integrator's output will be the opposite polarity; integrators invert. In our circuit we used a 2.5 V reference to split the 9 V battery; most of the current in the circuit is between the positive and negative rails, so the reference needs less than 20 μA of bias.<sup>41</sup> In this circuit we've shown a two-pole power switch, wired so the integrating capacitor is held reset until power is turned on.

The integrator output triggers a Schmitt comparator, as before, driving the mighty-lunged piezo screamer. Note that its large drive current is rail-to-rail; it does not pass through our ground reference. The circuit's operating cur-

rent is about 60 μA, good for nearly a year of continuous operation.

A final note: the LMC6044 is a quad, rail-to-rail output, micropower op-amp (10 μA/amplifier). So if a stiff ground reference were needed, the unused op-amp section could be configured as in Figure 4.73, with the stabilizing trick of Figure 4.76A.

### C. The “Mark-III” suntan integrator

It's always fun to see how elegantly you can shrink down circuit complexity. In this case there's a nice trick you can use to eliminate the two-stage integration, namely, a “current divider.” Figure 4.94 shows how it's done: the photocurrent drives a pair of resistors, bridging the same voltage (because the inverting input is a virtual ground); the current divides proportional to the relative conductance, in this case in the ratio of 1000:1 if pot R<sub>2</sub> is turned to minimum resistance. That means that a photocurrent of 1 μA would inject a current of 1 nA into the integrator. If you prefer, you can think of the circuit as a resistive load (R<sub>1</sub>) in series with R<sub>2</sub>, which easily dominates R<sub>3</sub>), which develops a voltage  $V_{in} = I_{diode}(R_1 + R_2)$ ; that voltage is the input to the integrator, via R<sub>3</sub>. Because the voltage developed by the photocurrent can range up to nearly a volt, it's necessary to back-bias the detector diode, in this case with a forward-biased diode D<sub>2</sub>, which generates a -0.4 V rail.

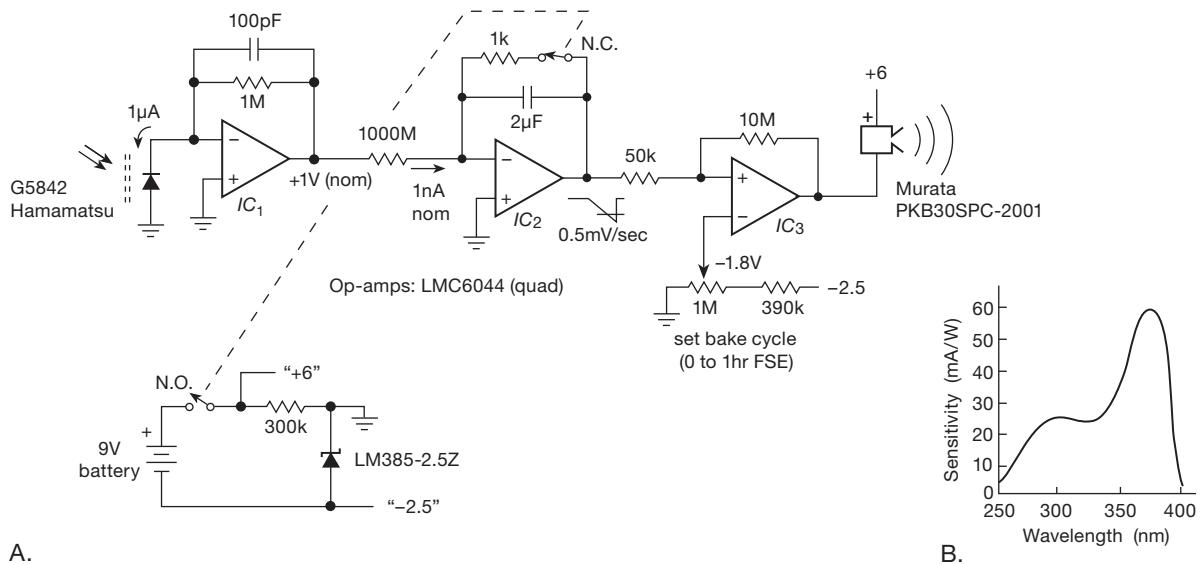
The integrator's positive-going output ramp drives Schmitt comparator A<sub>2</sub>, with fixed comparison voltage provided by reference D<sub>1</sub>. Its output drives the by-now usual piezo alarm.

Now for the elegance: it turns out you can get, packaged in a single small IC, a combination op-amp, comparator, and voltage reference. The MAX951 shown is just one of several such offerings, and it fills the bill here. It is because of the internal connection of D<sub>1</sub> and IC<sub>2</sub>'s inverting input that we were forced to put the suntan control at the input, rather than at the comparator.

A few additional comments.

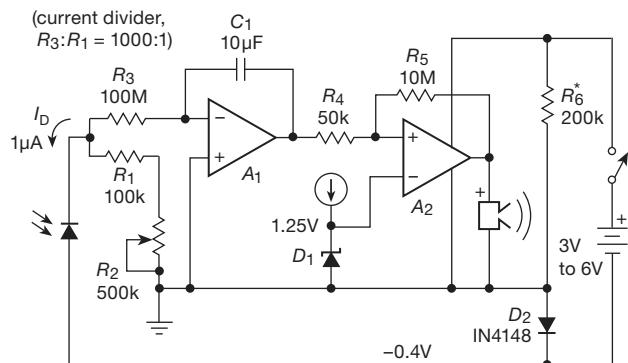
- The accuracy of the current divider depends on the accuracy of the virtual ground. The op-amp shown has a maximum offset voltage of 3 mV, so at 10% full sunlight and with the control set to minimum resistance (maximum bake cycle), the error is about 30% (10 mV signal, 3 mV offset). In other words, the circuit elegance involves a compromise in performance, relative to the more straightforward (some might say heavy-handed) approach in Figure 4.93, where the error is about 3% at minimum sunlight.
- Diode D<sub>2</sub> will be forward biased by the IC's quiescent

<sup>41</sup> Alternatively, we could have used a TLE2425 3-terminal “rail splitter” (§4.6.1B), which, however, would consume 170 μA. Although that would dominate the power budget, the thing would still run for 2000 hours (about 3 months) of continuous operation.



**Figure 4.93.** Integrating suntan monitor, second try; FSE is “full sunlight equivalent.” A. Schematic. B. Spectral response of the Hamamatsu G5842 photodiode, whose short-circuit photocurrent in sunlight is about  $1\mu\text{A}$ .

current of  $7\mu\text{A}$  as long as the photocurrent is less than this value. Thus the biasing resistor  $R_6$  can be omitted unless a maximum photocurrent of more than about  $5\mu\text{A}$  is anticipated.



**Figure 4.94.** Integrating suntan monitor, third try.  $A_1$ ,  $A_2$ , and  $D_1$  all live inside a MAX951 multifunction chip.  $A_2$  is a comparator.

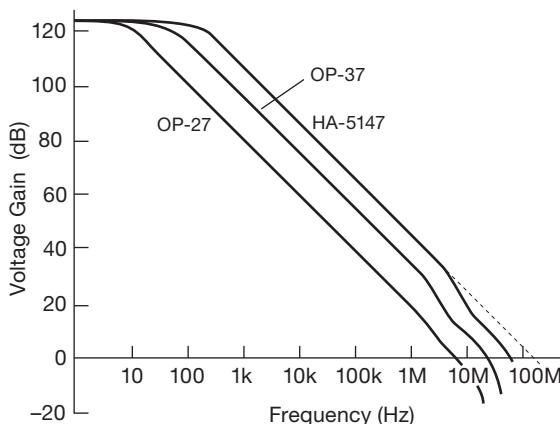
- The MAX951 has a specified operating voltage range of 2.7 V to 7 V. The low-voltage operation is a pleasant bonus; but for this particular IC it also means that we cannot run directly from a 9 V battery, unless we use a voltage regulator (see Chapter 9) to reduce the supply voltage to 7 V or less. This illustrates an important lesson, namely, that you have to watch out for low maximum supply-voltage ratings when using ICs intended for low-voltage operation. It also illustrates the trend of IC

manufacturers toward lower supply voltages for their new product designs.

#### 4.9 Feedback amplifier frequency compensation

We first met feedback in Chapter 2 (§2.5), where we saw its beneficial effects on stability and predictability of amplifier gain, and the reduction of an amplifier’s inherent nonlinearities. We saw, also, how it affects input and output impedances of amplifiers: for example, by sensing output voltage, and using series feedback at the input, the input impedance is raised and the output impedance is lowered, both by a factor of the loop gain. All is not rosy, though: the combination of gain with feedback creates the possibility of oscillation. Here, in the context of op-amps, we continue the treatment of negative feedback, looking at the important subject of *frequency compensation* – the business of preventing oscillation in amplifiers with negative feedback. The material in §2.5 is a necessary background for the sections that follow.

Let’s begin by looking at a graph of open-loop voltage gain versus frequency for several op-amps: you’ll typically see something like the curves in Figure 4.95. From a superficial look at such a *Bode plot* (a log–log plot of open-loop gain and phase versus frequency) you might conclude that the OP27 is an inferior op-amp, since its open-loop gain drops off so rapidly with increasing frequency. In fact, that rolloff is built into the op-amp intentionally and is



**Figure 4.95.** Open-loop gain versus frequency for three similar op-amps.

recognizable as the same  $-6 \text{ dB/octave}$  curve characteristic of an  $RC$  lowpass filter. The OP37, by comparison, is identical to the OP27 except that it is *decompensated* (and similarly for the discontinued<sup>42</sup> HA-5147). Op-amps are most often internally compensated, with decompensated and uncompensated varieties sometimes available. Let's take a look at this business of frequency compensation.

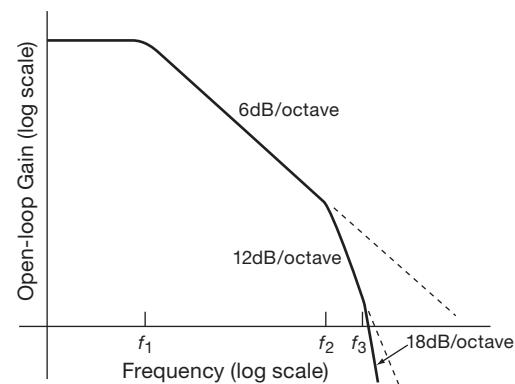
#### 4.9.1 Gain and phase shift versus frequency

An op-amp (or, in general, any multistage amplifier) will begin to roll off at some frequency because of the lowpass filters formed by signals of finite source impedance driving capacitive loads within the amplifier stages. For instance, it is common to have an input stage consisting of a differential amplifier, perhaps with current-mirror load (see the LF411 schematic in Figure 4.43), driving a common-emitter second stage. For now, imagine that the capacitor labeled  $C_C$  in that circuit is removed. The high output impedance of the input stage  $Q_2$ , in combination with the combined capacitance seen at its output, forms a lowpass filter whose 3 dB point might fall somewhere in the range of 100 Hz to 10 kHz.

The decreasing reactance of this capacitance with increasing frequency gives rise to the characteristic  $6 \text{ dB/octave}$  rolloff: at sufficiently high frequencies (which may be below 1 kHz), the capacitive loading dominates the collector load impedance, resulting in a voltage gain  $G_V = g_m X_C$ , i.e., the gain drops off as  $1/f$ . It also produces a  $90^\circ$  lagging phase shift at the output relative to the input signal. (You can think of this as the tail of an  $RC$  low-

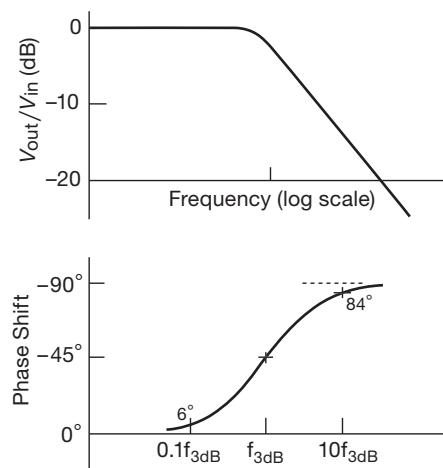
pass filter characteristic, where  $R$  represents the equivalent source impedance driving the capacitive load. However, it is not necessary to have any actual resistors in the circuit.)

In a multistage amplifier there will be additional rolloffs at higher frequencies, caused by lowpass filter characteristics in the other amplifier stages, and the overall open-loop gain will look something like that shown in Figure 4.96. The open-loop gain begins dropping at  $6 \text{ dB/octave}$  at some low frequency  $f_1$ , because of capacitive loading of the first-stage output. It continues dropping off with that slope until an internal  $RC$  of another stage rears its ugly head at frequency  $f_2$ , beyond which the rolloff goes at  $12 \text{ dB/octave}$ , and so on.



**Figure 4.96.** Multistage amplifier: gain versus frequency.

What is the significance of all this? Remember that an  $RC$  lowpass filter has a phase shift that looks as shown in Figure 4.97. Each lowpass filter within the amplifier has a similar phase-shift characteristic, so the overall phase



**Figure 4.97.** Bode plot: gain and phase versus frequency.

<sup>42</sup> See “Here Yesterday, Gone Today” on page 273.

shift of the hypothetical amplifier will be as shown in Figure 4.98.

Now here's the problem: if you were to connect this amplifier as an op-amp follower, for instance, it would oscillate. That's because the open-loop phase shift reaches  $180^\circ$  at some frequency at which the gain is still greater than 1 (negative feedback becomes positive feedback at that frequency). That's all you need to generate an oscillation, as any signal whatsoever at that frequency builds up each time around the feedback loop, just like a public address system with the gain turned up too far.

### A. Stability criterion

The criterion for stability against oscillation for a feedback amplifier is that its open-loop phase shift must be less than  $180^\circ$  at the frequency at which the loop gain is unity. This criterion is hardest to satisfy when the amplifier is connected as a follower, since the loop gain then equals the open-loop gain, the highest it can be. Internally compensated op-amps are designed to satisfy the stability criterion even when connected as followers; thus they are stable when connected for any closed-loop gain with a simple resistive feedback network. As we hinted earlier, this is accomplished by deliberately modifying an existing internal rolloff in order to put the 3 dB point at some low frequency, typically 1 Hz to 20 Hz. Let's see how that works.

## 4.9.2 Amplifier compensation methods

### A. Dominant-pole compensation

The goal is to keep the open-loop phase shift much less than  $180^\circ$  at all frequencies for which the loop gain is greater than 1. Assuming that the op-amp may be used as a follower, the words "loop gain" in the last sentence can be replaced with "open-loop gain." The easiest way to do this is to add enough *additional* capacitance at the

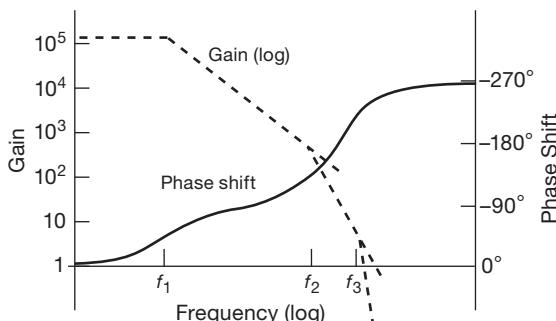


Figure 4.98. Gain and phase in a multistage amplifier.

point in the circuit that produces the initial 6 dB/octave rolloff, so that the open-loop gain drops to unity at about the 3 dB frequency of the next "natural"  $RC$  filter. In this way the open-loop phase shift is held at a constant  $90^\circ$  over most of the passband, increasing toward  $180^\circ$  only as the gain approaches unity. Figure 4.99 shows the idea. Without compensation, the open-loop gain drops toward 1, first at 6 dB/octave, then at 12 dB/octave, etc., resulting in phase shifts of  $180^\circ$  or more before the gain has reached 1. By moving the first rolloff down in frequency (forming a "dominant pole"), the rolloff is controlled so that the phase shift begins to rise above  $90^\circ$  only as the open-loop gain approaches unity. Thus, by sacrificing open-loop gain, you buy stability. Because the natural rolloff of lowest frequency is usually caused by the Miller effect in the stage driven by the input differential amplifier, the usual method of dominant-pole compensation consists simply of adding additional feedback capacitance around the second-stage transistor, so the combined voltage gain of the two stages is  $g_m X_C$  or  $g_m / 2\pi f C_{\text{comp}}$  over the compensated region of the amplifier's frequency response (Figure 4.100). In practice, Darlington-connected transistors would probably be used for both stages.

By putting the dominant-pole unity-gain crossing at the 3 dB point of the next rolloff, you get a phase margin of about  $45^\circ$  in the worst case (follower), since a single  $RC$  filter has a  $45^\circ$  lagging phase shift at its 3 dB frequency, i.e., the phase margin equals  $180^\circ - (90^\circ + 45^\circ)$ , with the  $90^\circ$  coming from the dominant pole.

An additional advantage of using a Miller-effect pole for

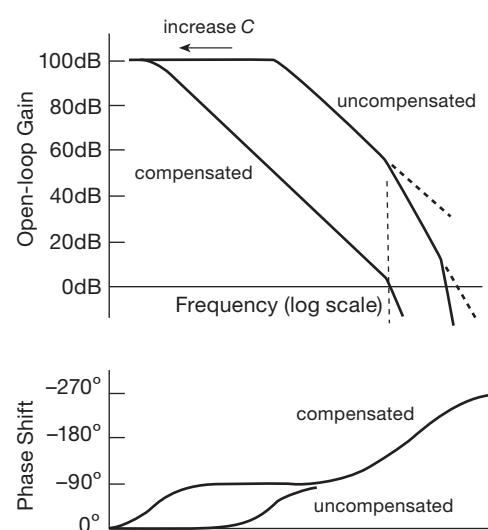


Figure 4.99. "Dominant-pole" compensation.

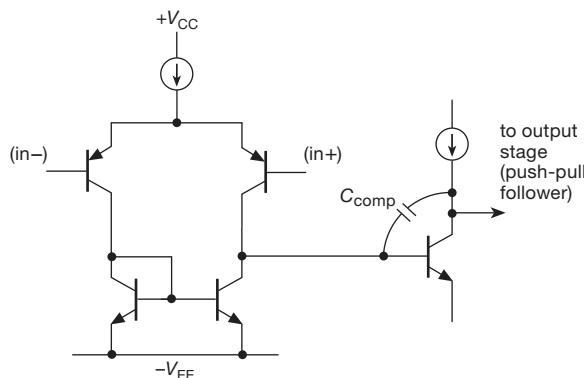


Figure 4.100. Classic op-amp input stage with compensation.

compensation is that the compensation is inherently insensitive to changes in voltage gain with temperature or manufacturing spread of gain: higher gain causes the feedback capacitance to look larger, moving the pole downward in frequency in exactly the right way to keep the unity-gain crossing frequency unchanged. In fact, the actual 3 dB frequency of the compensation pole is quite irrelevant; what matters is the point at which it intersects the unity-gain axis (Figure 4.101).

### B. Decompensated and uncompensated op-amps

If an op-amp is used in a circuit with closed-loop gain greater than unity (i.e., not a follower), it is not necessary to put the pole (the term for the “corner frequency” of a lowpass filter, see Chapter 1x) at such a low frequency as the stability criterion is relaxed because of the lower loop gain. Figure 4.102 shows the situation graphically.

For a closed-loop gain of 30 dB, the loop gain (which is the ratio of the open-loop gain to the closed-loop gain) is less than for a follower, so the dominant pole can be placed at a higher frequency. It is chosen so that the open-loop gain reaches 30 dB (rather than 0 dB) at the fre-

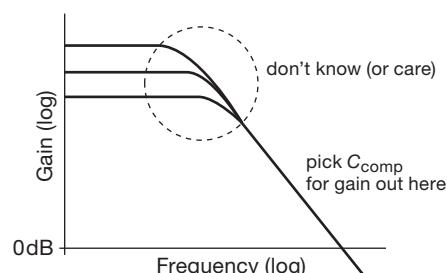


Figure 4.101. The compensation capacitor is chosen to set the open-loop unity-gain frequency; the low-frequency gain is unimportant.

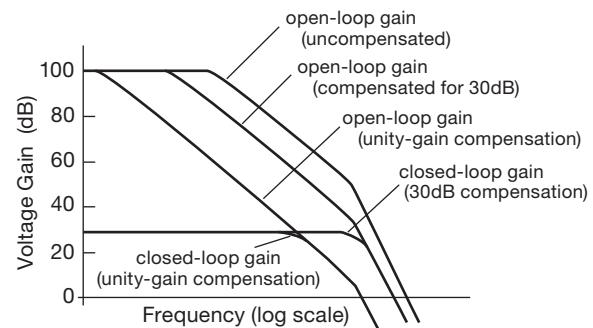


Figure 4.102. Stability is easier to achieve with larger closed-loop gain.

quency of the next natural pole of the op-amp. As the graph shows, this means that the open-loop gain is higher over most of the frequency range, and the resultant amplifier will work at higher frequencies. Some op-amps are available in “decompensated” (a better word might be “undercompensated”) versions, which are internally compensated for closed-loop gains greater than some minimum ( $A_v > 5$  in the case of the OP37); these specify a minimum closed-loop gain, and require no external capacitor. Another example is the THS4021/2, a decompensated version ( $G_v \geq 10$ ) of the unity-gain stable THS4011/2. These are really speedy op-amps, with an  $f_T$  of 300 MHz (for the “slow” THS4011/2), and greater than 1 GHz for the THS4021/2. For the decompensated versions the manufacturer (TI) supplies recommended external capacitance values (sometimes in combination with a resistor; see below) for a selection of minimum closed-loop gains.<sup>43</sup> Decompensated or uncompensated op-amps are worth using if you need the added bandwidth and your circuit operates at high gain; see Chapter 4x for further discussion.

*Some intuition:* it may at first seem paradoxical that an op-amp circuit configured for a *low*-gain circuit is more prone to oscillation than one configured for *high*-gain. But it makes sense: the better stability of an op-amp connected for a closed-loop gain of  $G_{CL}=100$  (40 dB), say, comes about because the feedback network (resistive divider) attenuates signals by a factor of 100. So it’s harder to sustain an oscillation going around the loop, compared with a follower (in which the feedback has unity gain).

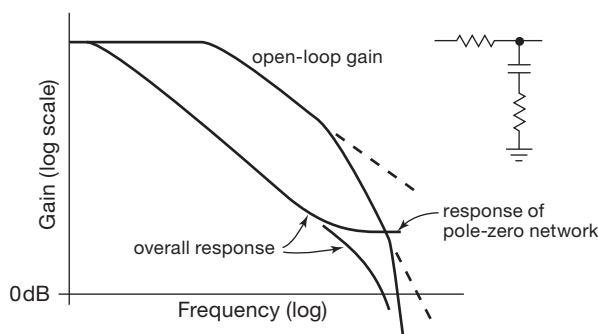
<sup>43</sup> In some cases external compensation components are required for any plausible closed-loop gain; these are properly called “uncompensated” op-amps.

### C. Pole-zero compensation

It is possible to do a bit better than with dominant-pole compensation by using a compensation network that begins dropping (6 dB/octave, a “pole”) at some low frequency, then flattens out again (it has a “zero”) at the frequency of the second natural pole of the op-amp. In this way the amplifier’s second pole is “cancelled,” giving a smooth 6 dB/octave rolloff up to the amplifier’s third pole. Figure 4.103 shows a frequency-response plot. In practice, the zero is chosen to cancel the amplifier’s second pole; then the position of the first pole is adjusted so that the overall response reaches unity gain at the frequency of the amplifier’s third pole. A good set of datasheets for an op-amp with external compensation will often give suggested component values (an  $R$  and a  $C$ ) for pole-zero compensation, as well as the usual capacitor values for dominant-pole compensation. Moving the dominant pole downward in frequency actually causes the second pole of the amplifier to move upward somewhat in frequency, an effect known as “pole splitting.” The frequency of the cancelling zero is then chosen accordingly.

#### 4.9.3 Frequency response of the feedback network

In all of the discussion thus far we have assumed that the feedback network has a flat frequency response; this is usually the case, with the standard resistive voltage divider as a feedback network. However, there are occasions when some sort of equalization amplifier is desired (integrators and differentiators are in this category) or when the frequency response of the feedback network is modified to improve amplifier stability. In such cases it is important to remember that the Bode plot of *loop* gain versus frequency is what matters, rather than the curve of open-loop gain. To make a long story short, the curve of ideal closed-loop



**Figure 4.103.** Cancelling the amplifier’s second pole in “pole-zero” compensation.

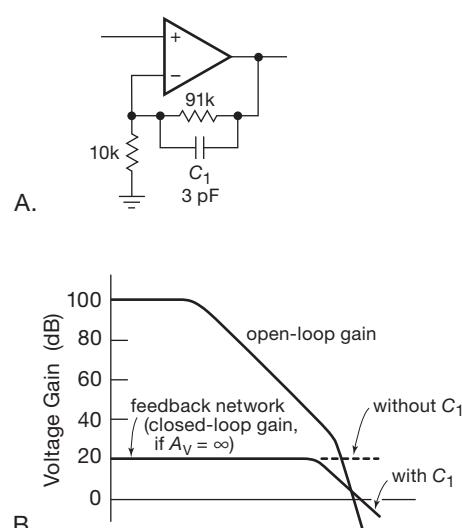
gain versus frequency should intersect the curve of open-loop gain, with a *difference* in slopes of 6 dB/octave. As an example, it is common practice to put a small capacitor (a few picofarads) across the feedback resistor in the usual inverting or noninverting amplifier. Figure 4.104 shows the circuit and Bode plot.

The amplifier would have been close to instability with a flat feedback network since the loop gain would have been dropping at nearly 12 dB/octave where the curves meet. The capacitor causes the loop gain to drop at 6 dB/octave near the crossing, guaranteeing stability. This sort of consideration is very important when designing differentiators because an ideal differentiator has a closed-loop gain that *rises* at 6 dB/octave; it is necessary to roll off the differentiator action at some moderate frequency, preferably going over to a 6 dB/octave rolloff at high frequencies. Integrators, by comparison, are very friendly in this respect, owing to their 6 dB/octave closed-loop rolloff. It takes real talent to make a low-frequency integrator oscillate!

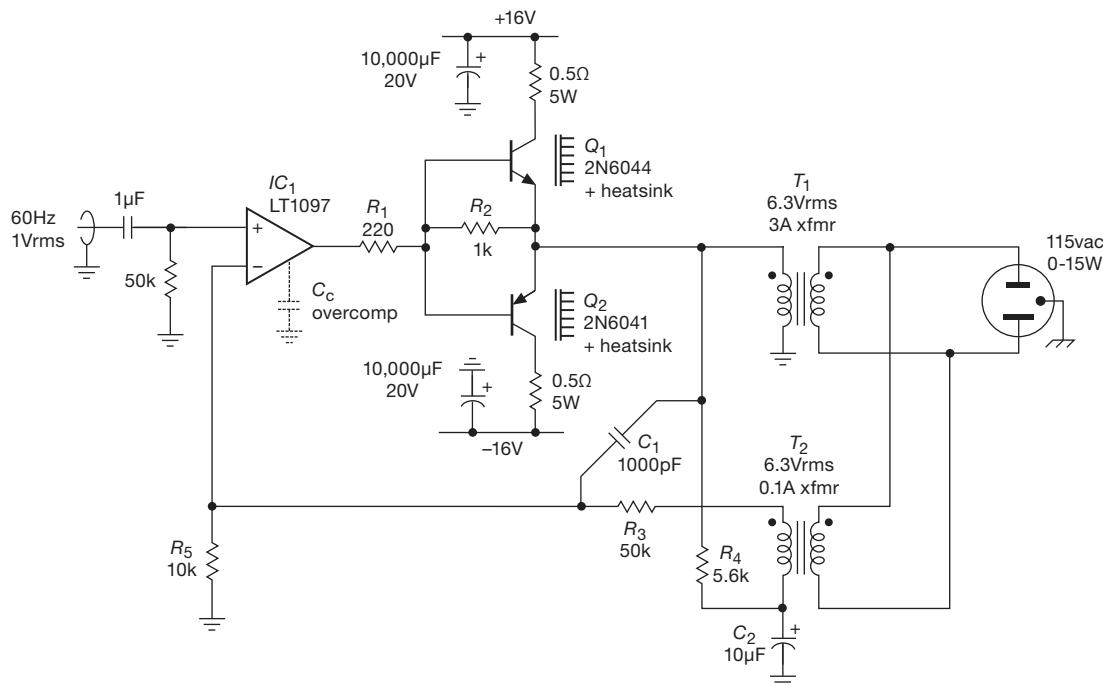
**Exercise 4.29.** Show on a Bode plot that the value of stabilizing resistor  $R_1$  in Figure 4.69 stops the differentiator action (i.e., flattens the curve of closed-loop gain) before the crossing point of open-loop and closed-loop gains. Explain our value of minimum recommended resistance  $R_1$ .

#### A. What to do

In summary, you are generally faced with the choice of internally compensated or uncompensated op-amps. It is simplest to use the compensated variety, and that’s the usual choice. You might begin by considering the



**Figure 4.104.** A small feedback capacitor enhances stability.



**Figure 4.105.** Output amplifier for 60 Hz power source. The push-pull output transistors  $Q_1$  and  $Q_2$  are power Darlintons in a plastic power package.

conventional LF411 (JFET,  $\pm 5$  V to  $\pm 15$  V supply) or an improved version (the LT1057), or the rail-to-rail input and output LMC6482 (CMOS, +3 V to +15 V supply), or perhaps the accurate and quiet LT1012, all internally compensated for unity gain. If you need greater bandwidth or slew rate, look for a faster compensated op-amp (see Table 4.2a on 271 for some choices). If it turns out that nothing is suitable and the closed-loop gain is greater than unity (as it usually is), you can use a uncompensated (or uncompensated) op-amp, perhaps with an external capacitor as specified by the manufacturer for the gain you are using. Using our previous example, the popular OP27 low-noise precision op-amp (unity-gain-compensated) has  $f_T = 8$  MHz and a slew rate of  $2.8$  V/ $\mu$ s; it is available as the uncompensated OP37 (minimum gain of 5), with  $f_T = 63$  MHz and a slew rate of  $17$  V/ $\mu$ s.<sup>44</sup>

### B. Example: precision 60 Hz power source

Uncompensated op-amps, or op-amps with a compensation pin, also give you the flexibility of *overcompensating*, a

simple solution to the problem of additional phase shifts introduced by other stuff in the feedback loop. Figure 4.105 shows an example. This is a low-frequency amplifier designed to generate a precise and stable 115 volt ac power output from a variable frequency 60 Hz low-level sinewave input.<sup>45</sup> The op-amp is wired as an ac-coupled noninverting amplifier, with its output driving a Darlington push-pull emitter-follower output stage  $Q_1Q_2$ , which in turn drives the low-voltage winding of a small power transformer,  $T_1$ , whose windings are in the ratio of 6.3 V:115 V. In this way we generate 115 V ac output without high-voltage op-amps or transistors. Of course, we pay the price in proportionally higher drive current; here the transistors need to supply about 3 A (rms) to produce a 15 W output.

To generate low distortion and a stable output voltage under load variations, we want to take feedback from the actual 115 V output sinewave. It is highly desirable, however, to keep the output fully isolated from circuit ground. So we use a second transformer  $T_2$  to produce a low-voltage replica of the 115 V output waveform, which is

<sup>44</sup> And, before it was discontinued, the similar “more-decompensated” HA-5147 (minimum gain of 10), with  $f_T = 120$  MHz and a slew rate of 35 V/ $\mu$ s.

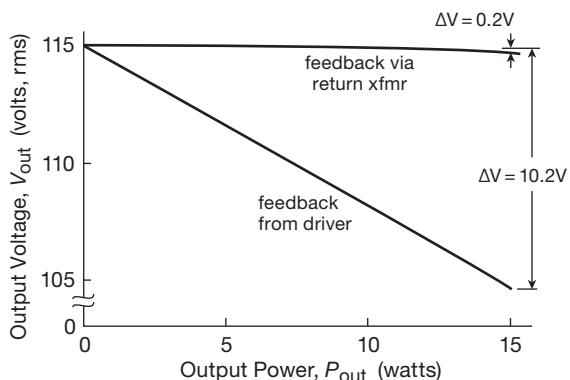
<sup>45</sup> The original design was used to drive an astronomical telescope at sidereal (star-tracking) rate. Interesting trivia: contrary to popular belief, the Earth turns on its axis once every 23 hours, 56 minutes, and 4.1 seconds; figure out why it isn’t 24:00:00!

then fed back via  $R_3$  to give the needed voltage gain of 6. Because of the unacceptably large phase shifts of the transformers at high frequencies, the circuit is rigged up so that at higher frequencies – above  $\sim 3$  kHz – the feedback comes from the low-voltage input to the transformer (via  $C_1$ ). Even though high-frequency feedback is taken directly from the push-pull output, there are still phase shifts associated with the reactive load (the transformer primary, a motor attached to the output, etc.) seen by the transistors. To ensure good stability, even with reactive loads at the 115 volt output, the op-amp can be overcompensated with a small capacitor, as shown. (The unusual LT1097 obligingly provides a pin for overcompensation.) The loss of bandwidth that results is unimportant in a low-frequency application like this.

The function of  $R_4$  and  $C_2$  may be puzzling: this bit of circuitry provides a dc feedback path for the op-amp, by averaging (lowpass filtering) the dc level applied to  $T_1$ , which then feeds back via the floating winding of  $T_2$ . We chose  $C_2$  large enough so its impedance at 60 Hz is small compared with the 50k feedback resistor; then we chose  $R_4$  for adequate smoothing consistent with stability.

The performance of this amplifier is quite satisfying. Figure 4.106 shows the output regulation, i.e., the change of rms output amplitude versus load. For comparison we show the comparable curve when feedback is taken exclusively from the driving winding of  $T_1$ , from which you can see that the desired feedback path improves output amplitude regulation, under load variations from zero to full power, from a mediocre 10% to just 0.2%. The output sinewave is very clean, with measured distortion well below 1% under all load conditions, including driving a synchronous motor (which represents a reactive load).

An application such as this represents a compromise,



**Figure 4.106.** Measured output voltage versus load for 60 Hz power source.

because ideally you would like to have plenty of loop gain to stabilize the output voltage against variations in load current. But a large loop gain increases the tendency of the amplifier to oscillate, especially if a reactive load is attached. This is because the reactive load, in combination with the transformer's finite output impedance, causes additional phase shifts within the low-frequency feedback loop. Because this circuit was built to power a telescope's synchronous driving motors (highly inductive loads), the loop gain was intentionally kept low.

Some additional points.

- With power electronics you should design conservatively so that a fault condition (e.g., too heavy a load, or even a short-circuit) doesn't destroy the device. Here we've used the simplest current-limiting method – a pair of small resistors in the collectors of the drive stage – because we didn't want to clutter the diagram (and it worked well enough, anyway!). There are better ways, though, for example, by adding a pair of transistors to rob base current when the output current (as sensed by a series resistor) exceeds a preset limit; such a scheme is commonly used within the integrated circuitry of op-amps themselves – see Figure 4.43. As we'll explain in §9.13.3, there are still better protective circuits. The problem with simple current-limiting protection is that a short-circuit load would cause the transistors to experience the limit current with the full supply voltage across them; the resulting power dissipation is far greater than the maximum under ordinary operation, which requires conservative heatsinking and component selection. *Foldback* current limiting would be better, though a bit more complicated.
- A push-pull follower with the bases tied together has a crossover region in which the feedback loop is effectively broken (see §2.4.1A). With Darlington transistors the crossover region is four  $V_{BE}$ 's, about 2.5 V. The resistor  $R_2$  in Figure 4.105 ensures that there is always some linear coupling from the op-amp to  $T_1$ , to prevent the feedback loop from rattling around under light load. Better still would be diode biasing, in the manner of Figures 2.71 or 2.72; see also the discussion of push-pull follower output stages in Chapter 2x.
- There is an elegant way to use a normal  $\pm 15$  V op-amp to generate larger voltage swings, by replacing the emitter followers in Figure 4.105 with a “pseudo-Darlington” configuration with modest noninverting gain (also known as a “series feedback pair,” see §2.5.5C), say a factor of 5. Then you can run the power output stage from a  $\pm 75$  V supply while powering the op-amp from conventional  $\pm 15$  V.

### C. Motorboating

In ac-coupled feedback amplifiers, stability problems can also crop up at very low frequencies, because of the accumulated *leading* phase shifts caused by several capacitively coupled stages. Each blocking capacitor, in combination with the input resistance (from bias strings and the like), causes a leading phase shift that equals  $45^\circ$  at the low-frequency 3 dB point and approaches  $90^\circ$  at lower frequencies. If there is enough loop gain, the system can go into a low-frequency oscillation picturesquely known as “motorboating.” With the widespread use of dc-coupled amplifiers, motorboating is almost extinct. However, old-timers can tell you some good stories about it.

### Additional Exercises for Chapter 4

**Exercise 4.30.** Design a “sensitive voltmeter” to have  $Z_{in} = 1 \text{ M}\Omega$  and full-scale sensitivities of 10 mV to 10 V in four ranges. Use a 1 mA meter movement and an op-amp. Trim voltage offsets if necessary, and calculate what the meter will read with input open, assuming (a)  $I_B = 25 \text{ pA}$  (typical for a 411) and (b)  $I_B = 80 \text{ nA}$  (typical for a 741). Use some form of meter protection (e.g., keep its current less than 200% of full scale), and protect the amplifier inputs from voltages outside the supply voltages. What do you conclude about the suitability of the 741 for low-level high-impedance measurements?

**Exercise 4.31.** Design an audio amplifier, using an OP27 op-amp (low noise, good for audio), with the following characteristics: gain = 20 dB,  $Z_{in} = 10\text{k}$ ,  $-3 \text{ dB}$  point = 20 Hz. Use the non-inverting configuration, and roll off the gain at low frequencies in such a way as to reduce the effects of input offset voltage. Use proper design to minimize the effects of input bias current on output offset. Assume that the signal source is capacitively coupled.

**Exercise 4.32.** Design a unity-gain phase splitter (see §2.2.8 in Chapter 2) using 411s. Strive for high input impedance and low output impedances. The circuit should be dc-coupled. At roughly what maximum frequency can you obtain full swing (27 V pp, with  $\pm 15 \text{ V}$  supplies), owing to slew-rate limitations?

**Exercise 4.33.** El Cheapo brand loudspeakers are found to have a treble boost, beginning at 2 kHz ( $+3 \text{ dB}$  point) and rising 6 dB/octave. Design a simple  $RC$  filter, buffered with AD611 op-amps (another good audio chip) as necessary, to be placed be-

tween preamp and amplifier to compensate for this rise. Assume that the preamp has  $Z_{out} = 50\text{k}$  and that the amplifier has  $Z_{in} = 10\text{k}$ , approximately.

**Exercise 4.34.** A 741 is used as a simple comparator, with one input grounded; i.e., it is a zero-crossing detector. A 1 volt amplitude sine wave is fed into the other input (frequency=1 kHz). What voltage(s) will the input be when the output passes through zero volts? Assume that the slew rate is  $0.5 \text{ V}/\mu\text{s}$  and that the op-amp’s saturated output is  $\pm 13 \text{ V}$ .

**Exercise 4.35.** The circuit in Figure 4.107 is an example of a “negative-impedance converter.” (a) What is its input impedance? (b) If the op-amp’s output range goes from  $V_+$  to  $V_-$ , what range of input voltages will this circuit accommodate without saturation?

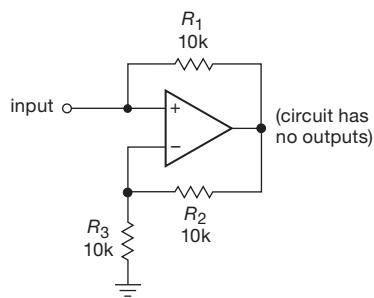


Figure 4.107. Negative-impedance converter.

**Exercise 4.36.** Consider the circuit in the preceding problem as the 2-terminal black box (Figure 4.108). Show how to make a dc amplifier with a gain of  $-10$ . Why can’t you make a dc amplifier with a gain of  $+10$ ? (Hint: the circuit is susceptible to a latchup condition for a certain range of source resistances. What is that range? Can you think of a remedy?)

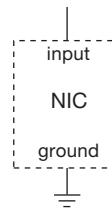


Figure 4.108. Negative-impedance connector as a 2-terminal device

## Review of Chapter 4

An A-to-O summary of what we have learned in Chapter 4. This summary reviews basic principles and facts in Chapter 4, but it does not cover application circuit diagrams and practical engineering advice presented there.

### ¶A. The Ideal Op-amp.

In Chapter 4 we explored the world of Operational Amplifiers (“Op-amps”), universal building blocks of analog circuits. A good op-amp approaches the ideal of an infinite-gain wideband noiseless dc-coupled difference amplifier with zero input current and zero offset voltage. Op-amps are intended for use in circuits with negative feedback, where the feedback network determines the circuit’s behavior. Op-amps figure importantly in the topics of Chapter 5 (Precision Circuits), Chapter 6 (Filters), Chapter 7 (Oscillators and Timers), Chapter 8 (Low-Noise Techniques), Chapter 9 (Voltage Regulation and Power Conversion), and Chapter 13 (Digital Meets Analog).

### ¶B. The “Golden Rules.”

At a basic level (and ignoring imperfections, see ¶¶K–M below), an op-amp circuit with feedback can be simply understood by recognizing that feedback from the output operates to (I) make the voltage difference between the inputs zero; and, at the same level of ignorance, (II) the inputs draw no current. These rules are quite helpful, and for dc (or low-frequency circuits) they are in error only by typical offset voltages of a millivolt or less (rule I), and by typical input currents of order a picoamp for FET types or tens of nanoamps for BJT types (rule II).

### ¶C. Basic Op-amp Configurations.

In §4.2 and §4.3 we met the basic linear circuits (detailed in ¶¶D–F below): inverting amplifier, non-inverting amplifier (and follower), difference amplifier, current source (transconductance, i.e., voltage-to-current), transresistance amplifier (i.e., current-to-voltage), and integrator. We saw also two important *non-linear* circuits: the Schmitt trigger, and the active rectifier. And in §4.5 we saw additional circuit building blocks: peak detector, sample-and-hold, active clamp, active full-wave rectifier (absolute-value circuit), and differentiator.

### ¶D. Voltage Amplifiers.

The *inverting amplifier* (Figure 4.5) combines input current  $V_{in}/R_1$  and feedback current  $V_{out}/R_2$  into a summing junction; it has voltage gain  $G_V = -R_2/R_1$  and input impedance  $R_1$ . In the *noninverting amplifier* (Figure 4.6) a fraction of

the output is fed back to the inverting input; it has voltage gain  $G_V = 1 + R_2/R_1$  and near-infinite input impedance. For the *follower* (Figure 4.8) the feedback gain is unity, i.e., the resistive divider is replaced by a connection from output to inverting input. The *difference amplifier* (Figure 4.9) uses a pair of matched resistive dividers to generate an output  $V_{out} = (R_2/R_1)\Delta V_{in}$ ; its input impedance is  $R_1 + R_2$ , and its common-mode rejection depends directly on the accuracy of the resistor matching (e.g.,  $\sim 60$  dB with  $\pm 0.1\%$  resistor tolerance). Difference amplifiers are treated in greater detail in §5.14. A pair of input followers can be used to achieve high input impedance, but a better 3-op-amp configuration is the *instrumentation amplifier*, see §5.15.

### ¶E. Integrator and Differentiator.

The *integrator* (Figure 4.16) looks like an inverting amplifier in which the feedback resistor is replaced by a capacitor; thus the input current  $V_{in}/R_1$  and feedback current  $CdV_{out}/dt$  are combined at the summing junction. Ignoring the imperfections in ¶K below, the integrator is “perfect,” thus any non-zero average dc input voltage will cause the output to grow and eventually saturate. The integrator can be reset with a transistor switch across the feedback capacitor (Figure 4.18); alternatively you can use a large shunt resistor to limit the dc gain, but this defeats the integrator operation at low frequencies ( $f \lesssim 1/R_f C$ ). The integrator’s input impedance is  $R_1$ .

The op-amp *differentiator* (Figure 4.68) is a similar configuration, but with  $R$  and  $C$  interchanged. Without additional components (Figure 4.69) this configuration is unstable (see ¶O, below).

### ¶F. Transresistance and Transconductance Amplifiers.

By omitting the input resistor, an inverting voltage amplifier becomes a *transresistance amplifier*<sup>46</sup>, i.e., a current-to-voltage converter (Figure 4.22). Its gain is  $V_{out}/I_{in} = -R_f$ , and (ignoring imperfections) the impedance at its input (which drives the summing junction) is zero. Capacitance at the input creates issues of stability, bandwidth, and noise; see §8.11 and the discussion in Chapter 4x. Transresistance amplifiers are widely used in photodiode applications.

A *transconductance amplifier* (Figures 4.10–4.15) converts a voltage input to a current output; it is a voltage-controlled current source. The simplest form uses an op-amp and one resistor (Figure 4.10), but works only with a floating load. The Howland circuit and its variations (Figures 4.14 and 4.15) drive a load returned to ground, but

<sup>46</sup> Or *transimpedance* amplifier.

their accuracy depends on resistor matching. Circuits with an external transistor (Figures 4.12 and 4.13) drive loads returned to ground, do not require resistor matching, and, in contrast to the other circuits, benefit from the intrinsically high output impedance of the transistor. In Chapter 4x we describe a nice variation on the transistor-assisted current source that achieves both high speed and bipolarity output (i.e., sinking and sourcing)

#### ¶G. Nonlinear Circuits: Peak Detector, S/H, Clamp, Rectifier.

Because of their high gain, op-amps provide accuracy to nonlinear functions that can be performed with passive components alone; in these circuits one or more diodes select the regions in which feedback acts. The *peak detector* (Figure 4.58) captures and holds the highest (or lowest) voltage since the last reset; the *sample-and-hold* (S/H) circuit (Figure 4.60) responds to an input pulse by capturing and holding the value of an input signal voltage; the *active clamp* (Figure 4.61) bounds a signal to a maximum (or minimum) voltage; the *active rectifier* creates accurate half-wave (Figures 4.36 and 4.38) or full-wave (Figures 4.63 and 4.64) outputs. In practice the performance of these circuits is limited by the finite slew rate and output current of real op-amps (see ¶M, below).

#### ¶H. Positive Feedback: Comparator, Schmitt Trigger, and Oscillator.

If the feedback path is removed, an op-amp acts as a *comparator*, with the output responding (by saturating near the corresponding supply rail) to a reversal of differential input voltage of a millivolt or less (Figure 4.32A). Adding some positive feedback (Figure 4.32B) creates a *Schmitt trigger*, which both speeds up the response and also suppresses noise-induced multiple transitions. Op-amps are optimized for use with negative feedback in linear applications (notably by a deliberate internal  $-6\text{ dB/octave}$  rolloff “compensation,” see ¶O below), so special comparator ICs (lacking compensation) are preferred, see §12.3 and Tables 12.1 and 12.6. A combination of positive feedback (Schmitt trigger) and negative feedback (with an integrator) creates an *oscillator* (Figure 4.39), a subject treated in detail in Chapter 7.

#### ¶I. Single-Supply and Rail-to-rail Op-amps.

For some op-amps both the input common-mode range and the output swing extend all the way down to the negative rail, making them particularly suited for operation with a single positive supply. Rail-to-rail op-amps allow input swings to both supply rails, or output swings to both rails,

or both; see Table 4.2a. The latter are especially useful in circuits with low supply voltages.

#### ¶J. Some Cautions.

In linear op-amp circuits, the Golden Rules (see ¶B, above) will be obeyed only if (a) feedback is negative and (b) the op-amp stays in the active region (i.e., not saturated). There must be feedback at dc, or the op-amp will saturate. Power supplies should be bypassed. Stability is degraded with capacitive loads, and by lagging phase shifts in the feedback path (e.g., by capacitance at the inverting terminal). And, most important, real op-amps have a host of limitations (¶K–N, below) that bound attainable circuit performance.

#### ¶K. Departures from Ideal Behavior.

In the real world op-amps are not perfect. There is no “best” op-amp, thus one must trade off a range of parameters: input imperfections (offset voltage, drift, and noise; input current and noise; differential and common-mode range), output limitations (slew rate, output current, output impedance, output swing), amplifier characteristics (gain, phase shift, bandwidth, CMRR and PSRR), operating characteristics (supply voltage and current), and other considerations (package, cost, availability). See §4.4, Tables 4.1, 4.2a, and 4.2b, the more extensive tables in Chapters 5, and 8, and ¶L–N below.

#### ¶L. Input Limitations.

The *input offset voltage* ( $V_{os}$ ), ranging from about  $25\text{ }\mu\text{V}$  (“precision” op-amp) to  $5\text{ mV}$ , is the voltage unbalance at the input terminals. It’s an important parameter for precision circuits, and circuits with high closed-loop dc gain; the error seen at the output is  $G_{CL}V_{os}$ ). Some op-amps provide pins for external trimming of offset voltage (e.g., see Figure 4.43).

The *offset voltage drift*, or *tempco* ( $TCV_{os}$ , or  $\Delta V_{os}/\Delta T$ ), is the temperature coefficient of offset voltage; it ranges from about  $0.1\text{ }\mu\text{V}/^\circ\text{C}$  (“precision” op-amp) to  $10\text{ }\mu\text{V}/^\circ\text{C}$ . Even if you’re lucky and have an op-amp with low  $V_{os}$  (or you’ve trimmed it to zero),  $TCV_{os}$  represents the growth of offset with changing temperature.

The *input noise voltage density* ( $e_n$ ) represents a noisy voltage source in series with the input terminals. It ranges from about  $1\text{ nV}/\sqrt{\text{Hz}}$  (low-noise bipolar op-amp) to  $100\text{ nV}/\sqrt{\text{Hz}}$  or more (micropower op-amps). Noise voltage is important in audio and precision applications.

The *input bias current* ( $I_B$ ) is the (non-zero) dc current at the input terminals. It ranges from a low of about  $5\text{ fA}$  (CMOS low-bias op-amps, and “electrometer” op-amps) to

50 nA (typical<sup>47</sup> BJT op-amps) to a high of 10  $\mu$ A (wideband BJT-input op-amps). Bias current flowing through the circuit's dc source resistance causes a dc voltage offset; it also creates a current error in integrators and transresistance amplifiers.

The *input noise current* ( $i_n$ ) is the equivalent noise current added at the input. For most op-amps<sup>48</sup> it is simply the shot noise of the bias current ( $i_n = \sqrt{2qI_B}$ ); it ranges from about 0.1 fA/ $\sqrt{\text{Hz}}$  (CMOS low-bias op-amps, “electrometer” op-amps) to 1 pA/ $\sqrt{\text{Hz}}$  (wideband BJT op-amps). Input noise current flowing through the circuit's ac source impedance creates a noise voltage, which can dominate over  $e_n$ . The ratio  $r_n = e_n/i_n$  is the op-amp's *noise resistance*; for signal source impedances greater than  $r_n$  the current noise dominates.

Op-amps function properly when both inputs are within the *input common-mode voltage range* ( $V_{CM}$ ), which may extend to the negative rail (“single-supply” op-amps), or to both rails (“rail-to-rail” op-amps). Beware: many op-amps have a more restricted *input differential voltage range*, sometimes as little as just a few volts.

#### ¶M. Output Limitations.

The *slew rate* (SR) is the op-amp's  $dV_{out}/dt$  with an applied differential voltage at the input. It is set by internal drive currents charging the compensation capacitor, and ranges from about 0.1 V/ $\mu$ s (micropower op-amps) to 10 V/ $\mu$ s (general purpose op-amps) to 5000 V/ $\mu$ s (high-speed op-amps). Slew rate is important in high-speed applications generally, and in large-swing applications such as A/D and D/A converters, S/H and peak detectors, and active rectifiers. It limits the large-signal output frequency: a sinewave of amplitude  $A$  and frequency  $f$  requires a slew rate of  $SR=2\pi Af$ ; see Figure 4.54.

Op-amps are small devices, with *output current* deliberately limited to prevent overheating; see for example Figure 4.43, where  $R_5Q_9$  and  $R_6Q_{10}$  limit the output sourcing and sinking currents to  $I_{lim}=V_{BE}/R \approx 25$  mA, illustrated in Figure 4.45. If you need more output current, there are a few high-current op-amps available; you can also add an external unity-gain power buffer like the LT1010 ( $I_{out}$  to  $\pm 150$  mA), or a discrete push-pull follower.

The open-loop *output impedance* of an op-amp is generally in the neighborhood of 100  $\Omega$ , which is reduced by the loop gain to fractions of an ohm at low frequencies. Because an op-amp's open-loop gain  $G_{OL}$  falls as  $1/f$  over

most of its bandwidth (see ¶O below), however, the circuit's *closed-loop* output impedance rises approximately proportional to frequency; it looks inductive (Figure 4.53).

In general the *output swing* for an op-amp like Figure 4.43 extends only to within a volt or so from either rail. Many CMOS and other low-voltage op-amps, however, specify unloaded rail-to-rail output swings, see Figure 4.46.

Op-amps can be grouped into several *supply voltage* ranges: “low-voltage” op-amps have a maximum total supply voltage (i.e.,  $V_+ - V_-$ ) around 6 V, and generally operate down to 2 V; “high-voltage” op-amps allow total supply voltages to 36 V, and generally operate down to 5–10 V. In between there is a sparse class of what might be called “mid-voltage” op-amps, with total supply voltages in the neighborhood of 10–15 V. See Table 5.5. There are also op-amps that are truly high-voltage (to hundreds of volts), see Table 4.2b.

#### ¶N. Gain, Phase Shift, and Bandwidth.

Op-amps have large dc *open-loop gain*  $G_{OL(dc)}$ , typically in the range of  $10^5$ – $10^7$  (the latter being typical of “precision” op-amps, see Chapter 5). To ensure stability (see ¶O, below) the op-amp's open-loop gain falls as  $1/f$ , reaching unity at a frequency  $f_T$  (see Figure 4.47). This limits the closed-loop *bandwidth* to  $BW_{CL} \approx f_T/G_{CL}$ . Over most of the operating frequency range the op-amp's open-loop *phase shift* is  $-90^\circ$ , eliminated in the closed-loop response by feedback.

#### ¶O. Feedback Stability, “Frequency Compensation,” and Bode Plots

Finally, negative feedback can become *positive feedback*, promoting instability and oscillations, if the accumulated phase shift reaches  $180^\circ$  at a frequency at which the loop gain is  $\geq 1$ . This topic is foreshadowed in §4.6.2 in connection with capacitive loads, and it is discussed in detail in §4.9. The basic technique is *dominant-pole compensation*, in which a deliberate  $-6$  dB/octave (i.e.,  $\sim 1/f$ ) rolloff is introduced within the op-amp in order to bring the gain down to unity at a frequency lower than that at which additional unintended phase shifts rear their ugly heads (Figure 4.99). Most op-amps include such compensation internally, such that they are stable at all closed-loop gains (the unity-gain follower configuration is most prone to instability, because there is no attenuation in the feedback path). “Decompensated” op-amps are less aggressively compensated, and are stable for closed-loop gains greater than some minimum (often specified as  $G > 2, 5$ , or 10;

<sup>47</sup> The input current of “bias-compensated” BJT op-amps is typically around 50 pA.

<sup>48</sup> But not “bias-compensated” BJT op-amps, see §8.9.

Figure 4.95). Compensated op-amps exhibit an open-loop lagging phase shift of  $90^\circ$  over most of their frequency range (beginning as low as 10 Hz or less). Thus an external feedback network that adds another  $90^\circ$  of lagging phase shift at a frequency where the loop gain is unity will cause oscillation.

A favorite tool is the *Bode Plot*, a graph of gain (log) and phase (linear) versus frequency (log); see Figure 4.97. The *stability criterion* is that the difference of slopes between the open-loop gain curve and the ideal closed-loop gain curve, at their intersection, should ideally be 6 dB/octave, but in no case as much as 12 dB/octave.

# PRECISION CIRCUITS

## CHAPTER 5

In the preceding chapters we dealt with many aspects of analog circuit design, including the circuit properties of passive devices, transistors, FETs, and op-amps, the subject of feedback, and numerous applications of these devices and circuit methods. In all our discussions, however, we have not yet addressed the question of the best that can be done, for example, in minimizing amplifier errors (nonlinearities, drifts, etc.) and in amplifying weak signals with minimum degradation by amplifier “noise.” In many applications these are the most important issues, and they form an important part of the art of electronics. Therefore, in this chapter, we look at methods of precision circuit design (deferring the issue of noise in amplifiers to Chapter 8).

### Chapter overview

This is a *big* chapter – and an important one. It deals with a range of topics, which need not be read in order. As guidance, we offer this outline: we start with a careful examination of errors in circuits made with operational amplifiers, and explore the use of an error budget. We explore issues of unspecified parameters and “typical” versus “worst-case” component errors, and discuss ways to deal with them. Along the way we deal with some neglected topics such as diode leakage at the sub-picoamp level, “memory effect” in capacitors, distortion and gain nonlinearity, and an elegant way to remedy *phase* error in amplifiers. We discuss op-amp distortion in detail, with comparative graphs and test circuits.

Next we discuss the dark side of rail-to-rail op-amps: their open-loop output impedance, and input common-mode crossover errors. We provide detailed selection tables for precision, chopper and high-speed op-amps, and comparative graphs charting their noise, bias current, and distortion. We show how to interpret the multitude of op-amp parameters, and we discuss the tradeoffs you’ll have to make.

For those working in the low microvolt and nanovolt territory, we show the devastating effects of  $1/f$  noise, and how auto-zero (AZ) op-amps solve this problem; but there’s a tradeoff – the current noise of these devices that is often overlooked. As an interlude we look in some de-

tail at the cleverness of the front-end stage of an exemplary precision digital multimeter.

Then we advance to difference and instrumentation amplifiers – these go to the head of the class both in terms of digging out a difference signal in the presence of common-mode input, and in terms of gain accuracy and stability. We show their internal designs and how they’re used, with extensive tables and graphs comparing popular parts. Finally we take up fully differential amplifiers – these have differential inputs *and* outputs, and an output “common-mode control” input pin. Once again, we organize tables, internal circuit diagrams, and guidance for their use with high-performance ADCs.

For readers looking for the basics, this chapter can be skipped over in a first reading. Its material is not essential for an understanding of later chapters.

### 5.1 Precision op-amp design techniques

In the field of measurement and control there is often a need for circuits of high precision. Control circuits should be accurate, stable with time and temperature, and predictable. The usefulness of measuring instruments likewise depends on their accuracy and stability. In almost all electronic subspecialties we always have the desire to do things more accurately – you might call it the joy of perfection. Even if you don’t always actually *need* the highest precision, you can still delight in a full understanding of what’s going on.

#### 5.1.1 Precision versus dynamic range

It is easy to get confused between the concepts of *precision* and *dynamic range*, especially because some of the same techniques are used to achieve both. Perhaps the difference can best be clarified by some examples: a 5-digit multimeter has high precision; voltage measurements are accurate to 0.01% or better. Such a device also has wide dynamic range; it can measure millivolts and volts on the same scale. A precision decade amplifier (one with selectable gains of 1, 10, and 100, say) and a precision voltage reference may

have plenty of precision, but not necessarily much dynamic range. An example of a device with wide dynamic range but only moderate accuracy might be a 6-decade logarithmic amplifier (log-amp) built with carefully trimmed op-amps but with components of only 5% accuracy; even with accurate components a log-amp might have limited accuracy because of lack of log conformity (at the extremes of current) of the transistor junction used for the conversion, or because of temperature-induced drifts.

Another example of a wide-dynamic-range instrument (greater than 10,000:1 range of input currents) with only moderate accuracy requirement (1%) is the coulomb meter described in §9.26 of the previous edition of this book. It was originally designed to keep track of the total charge put through an electrochemical cell, a quantity that needs to be known only to approximately 5% but that may be the cumulative result of a current that varies over a wide range. It is a general characteristic of wide-dynamic-range design that input offsets must be carefully trimmed in order to maintain good proportionality for signal levels near zero; this is also necessary in precision design, but, in addition, precise components, stable references, and careful attention to every possible source of error must be used to keep the sum total of all errors within the so-called error budget.

### 5.1.2 Error budget

A few words on *error budgets*. There is a tendency for the beginner to fall into the trap of thinking that a few strategically placed precision components will result in a device with precision performance. On rare occasions this will be true. But even a circuit peppered with 0.01% resistors and expensive op-amps won't perform to expectations if somewhere in the circuit there is an input offset current multiplied by a source resistance that gives a voltage error of 10 mV, for example. With almost any circuit there will be errors arising all over the place, and it is essential to tally them up, if for no other reason than to locate problem areas where better devices or a circuit change might be needed. Such an error budget results in rational design, in many cases revealing where an inexpensive component will suffice, and eventually permitting a careful estimate of performance.

To add some spice to the subject, we note that there is some controversy in the engineering community surrounding this business of error budgets. One camp (which we might characterize as strict constructionists) insists that you allow for the true worst case, or you are guilty of violating good engineering practice. For example, if there are

18 gain-determining resistors of  $\pm 1\%$  tolerance in a circuit, then the guaranteed performance must be specified as  $\pm 18\%$ . The response of the other camp (which we might characterize as pragmatists) is “balderdash – it's overly constraining to allow an extremely unlikely possibility to limit the performance of a circuit design; and one can deal with such eventualities with component test procedures, finished circuit performance testing, and so on.” We'll revisit this controversy (and take sides in the debate) after running through an introductory example.

### 5.2 An example: the millivoltmeter, revisited

To motivate the discussion of precision circuits, let's revisit a circuit from the previous chapter. There we flirted, briefly, with issues of precision in §4.4.3, mostly to illustrate the effects of input offset voltage  $V_{OS}$  and input bias current  $I_B$  in a low-level dc application (a 0–10 mV millivoltmeter with  $10\text{ M}\Omega$  input resistance<sup>1</sup>). Back then, with wide-eyed naivety, we were astonished to see that our trusty LF411 op-amp was wholly inadequate to the task; it had way too much offset, and too much input current as well. We found a solution in the form of either a precision low-bias op-amp (an OPA336) or a chopper (also known as “auto-zeroing”) amplifier (an LTC1050).

As we'll soon see, our celebration of that “solution” was premature: we pronounced victory with an op-amp whose  $I_B$  alone caused the maximum allowable zero-input error of 1%. A careful design must take into account the cumulative effect of multiple sources of error.

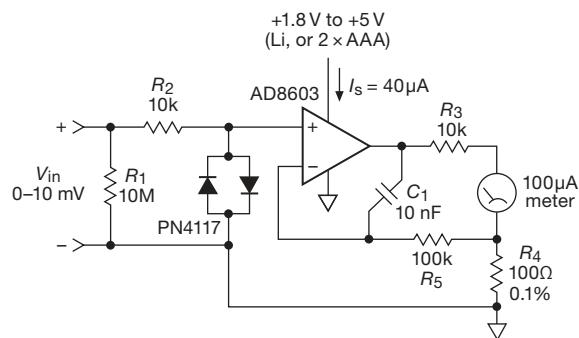
#### 5.2.1 The challenge: 10 mV, 1%, 10 MΩ, 1.8 V single supply

To make the problem more interesting, let's further constrain the specifications. This time we'll ask that the 0–10 mV meter operate from a single +3 V battery (either a lithium cell or a pair of alkaline AAA cells); that forces us to worry about “single-supply” operation, in which the op-amp must work down to zero volts at both input and output. Furthermore, it must work down to the end-of-life voltage of alkaline cells, which you see stated variously as 1.0 V/cell, or 0.9 V/cell; that means operation down to +1.8 V total supply voltage. And, as before, let's require an input resistance of  $10\text{ M}\Omega$  and insist that it indicates 0 mV ( $\pm 1\%$  of full scale) when the input is either shorted

<sup>1</sup> Note that it can be used as a sensitive *current* meter: with its  $10\text{ M}\Omega$  input resistance and 1% accuracy, it can measure currents down to  $10\text{ pA}$  ( $1\% \times 10\text{ mV}/10\text{ M}\Omega = 10\text{ pA}$ ).

or open. Note that this “zero-error” specification is different from a full-scale accuracy (“scale error”) specification: we might be happy with  $\pm 5\%$  full-scale accuracy, but we’d be most unhappy with a meter that reads 5% of full scale (here 0.5 mV) when there’s nothing connected to it.

Following the suggestion from last chapter’s design, let’s use current-sensing feedback, so the design is independent of the internal resistance of the analog meter. Figure 5.1 shows the circuit.



**Figure 5.1.** Accurate millivoltmeter powered with single lithium cell. The input protection clamp uses low-leakage diode-connected PN4117 JFETs.

### 5.2.2 The solution: precision RRIO current source

We use a precise current-sensing resistor  $R_4$ , in this case a 0.1%  $100\Omega$  resistor. Sounds exotic, but in fact these things are commonplace: the ever-helpful DigiKey website shows over 100,000 in stock, from five different suppliers, at prices down to \$0.20 (in quantities of 10). Note the routing of the input common (“–” terminal) connection directly to the low side of the sense resistor, a precaution that becomes increasingly important with small-value sense resistors, in which the wiring resistance of the ground return may add significant error.<sup>2</sup> Because the meter likely presents an inductive load (it’s a moving coil in a magnetic field, which is both inductive in its own right and reactive through its motor-like electromechanical properties), we took the precaution of dividing the feedback path in the usual way (through  $R_5$  at low frequencies,  $C_1$  at high frequencies; see for example Figure 4.76). The 10k output resistor  $R_3$  limits meter current for off-scale inputs.

The more challenging parts of this design are the input

protection network (which we blithely ignored in Chapter 4’s example), and, most critically, the choice of op-amp. First, the protection network: the requirements seem easy – clamp to a nondestructive op-amp input voltage (during input overvoltages), and draw less than  $\sim 10\text{ pA}$  leakage current at the full-scale input voltage (10 mV), in both forward and reverse directions. (That amount of diode current would reduce the input resistance by 1%.) As it turns out, the datasheets don’t ordinarily tell you how much current a diode draws at very low voltages. But if you go measure it, you’ll be surprised at what you find (Figure 5.2). Everyone’s favorite jellybean signal diode (1N914, 1N4148) is rather leaky, looking roughly like a  $10\text{ M}\Omega$  resistor at low voltages.<sup>3</sup> There are some specialized low-leakage diodes like the (somewhat hard to get) PAD-1 or PAD-5 that do much better; but you can do as well by using a diode-connected low-leakage JFET like the *n*-channel PN4117 (i.e., tie the source and drain together to form the cathode, and use the gate as the anode), or you can just use the diode terminal pairs from an ordinary *npn* transistor.<sup>4</sup> In this circuit the upstream 10k resistor  $R_2$  limits the clamp current, while having no effect on the circuit accuracy.

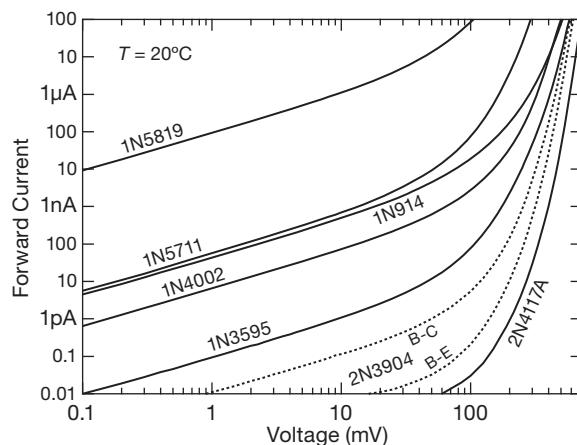
And now for the op-amp. This was the stumbling block in Chapter 4, and it has gotten only more difficult here, with the single low-voltage supply. We can separate the errors into a “zero” error and an overall scale factor error. The latter is the easy part: the circuit’s gain is accurately determined, so we merely require an accurate meter movement (if we don’t want any trim adjustments; or we could reduce the sense resistor and add a resistor to produce a trimmable gain greater than unity). It’s the “zero” requirement that is the tough part, because of the high sensitivity plus the high mandated input resistance. We require a worst-case combined effect of input offset and bias current of  $100\text{ }\mu\text{V}$  and  $10\text{ pA}$  individually. That is, each alone would cause a zero error of 0.1 mV ( $V_{err} = V_{OS} + I_B R_1$ ), so each must be smaller so that the worst-case combination meets specifications.

We looked at some promising contemporary op-amp offerings, which we’ve listed (along with the usual suspects) in Table 5.1 on page 296. The inexpensive jellybeans in the

<sup>3</sup> For all these diodes, the straight-line portion at low voltages represents a resistance in parallel with a non-conducting diode; so the low-leakage 1N3595 looks like  $10,000\text{ M}\Omega$  for  $V \lesssim 10\text{ mV}$ . These currents can be estimated from other diode parameters, namely reverse-bias leakage current at low reverse voltage or forward current at specified forward voltage. More on this in Chapter 1x.

<sup>4</sup> Or our friend John Larkin’s favorite, the collector-base junction of the BFT25, a low-cost 5 V *npn* microwave transistor. Its leakage is less than 10 fA when reverse biased, and <40 fA when forward biased up to 50 mV.

<sup>2</sup> Small-value sense resistors used for accurate measurement of high currents are available as 4-wire resistors. This arrangement is known as a *Kelvin connection*, and the sense resistor is sometimes called a *shunt*.



**Figure 5.2.** Diode datasheets are often skimpy with data like this, showing the low end of the current versus forward applied voltage. See Chapter 1x for lots more detail.

first three rows are hopeless for this job, with both offset voltage and bias current far worse than required (listed in the bottom row). And they all fail to operate at +1.8 V; and the first two fail the “single-supply” requirement. So much for el-cheapo op-amps.

The LPV521 is characteristic of the new breed of “low-voltage, low-bias, low-power” op-amps. It does those things just fine, but, in common with many CMOS op-amps, its  $V_{OS}$  specs are only so-so. It’s possible to do considerably better. For example, the CMOS precision OPA336 was our hero earlier; but look – its worse-case  $I_B$  consumes our entire error budget, and it would require  $V_{OS}$  trimming anyway, which is not so easily done in a single-supply setup (unlike our friend the LF411, this op-amp has no trim pins). We brushed off such niceties earlier; but this time let’s accept full responsibility for producing products that meet specs. That means adopting some serious worst-case discipline.

What choices are left? Most designers would reach for an auto-zero (chopper) amplifier for a precision application that does not require bandwidth and that is tolerant of noise. The best candidate we found is the ADA4051, with excellent  $V_{OS}$  specs, but five times too much  $I_B$ , if you believe the “maximum” specs. (Competing auto-zero amplifiers had greater bias current, greater minimum supply voltage, or both.) The last two op-amps are contemporary CMOS op-amps that qualify as “precision,” owing to careful design and production-line offset trimming. Both meet the error budget goals (but see below). We chose the AD8603 because it meets the 1.8 V supply spec, and as a bonus it operates at 35% the supply current of the LTC6078

(a dual op-amp, with no single version we could find).<sup>5</sup> As listed in Table 5.5, the AD8603 runs at 40  $\mu$ A quiescent current, and has 1 pA (max) input current and 50  $\mu$ V (max) offset voltage at 25°C.

Are we done? Not quite. The specs in the table are for operation at 25°C. It can get plenty warmer, with impressive increases in bias current for CMOS devices (for which the “bias” current is leakage current). Manufacturers will ordinarily provide worst-case (and sometimes typical) values at the high end of the operating temperature range (e.g., 85°C for “industrial” temperature range devices). For our chosen AD8603 they specify  $I_B(\text{max})=50 \text{ pA}$  at 85°C. No worst-case data for intermediate temperatures are given; but that value is consistent with a doubling every 10°C, so we can be confident that the circuit will meet specs<sup>6</sup> up to 50°C. It’s worth noting that manufacturers are sometimes rather lazy in setting worst-case leakage specs, as for example with the LPV521 in the table, whose ratio of “maximum” to “typical”  $I_B$  is 100:1. One industry guru attributes that to an unwillingness to test production parts to a tighter spec.

**Exercise 5.1.** We designed a  $\pm 10 \text{ mV}$  meter in Chapter 4 (Figure 4.56), but our design in Figure 5.1 is unipolar (i.e., 0 to  $+10 \text{ mV}$ ). Your boss has asked you to modify the design for  $\pm 10 \text{ mV}$  capability, using a  $\pm 100 \mu\text{A}$  zero-center meter and retaining the feature of operating with either a single lithium cell (or pair of AAA alkaline cells). Hint: you may wish to look at the AD8607 dual member of the AD8603 op-amp family. Extra credit: after finishing, you decide to impress your boss by enhancing the design so it will work with a 0–100  $\mu\text{A}$  meter.

### 5.3 The lessons: error budget, unspecified parameters

From this rather simple first example we’ve learned some important basic principles: (a) first, you’ve got to identify and quantify the sources of error within a circuit in order to create an error budget; and (b) strict worst-case design requires that all components (passive and active) be operated within their datasheet specifications, and that the effects of

<sup>5</sup> Five other parts we considered (see Tables 5.5, pages 320–321, and 5.6, page 335): the bipolar LT1077A, lower  $I_s$  and  $V_{OS}$  but too high  $V_s$  (min); the bipolar LT6003,  $I_s$  only 1  $\mu\text{A}$ , but too much offset voltage (a 5% error); the CMOS LMP2232A, poor 1.5% offset voltage; the MAX9617 auto-zero op-amp, with 0.1% (versus the AD8603’s 0.5%) offset voltage, but whoa, its input current through 10  $\text{M}\Omega$  amounts to a 1.4% offset; and the ISL28133 auto-zero, but its 3% offset (from input current) again reveals the input current weakness of auto-zero amplifiers.

<sup>6</sup> The datasheet does provide a plot of “typical”  $I_B$  versus temperature, confirming its exponential behavior of a doubling every 10°C.

**Table 5.1 Millivoltmeter Candidate Op-amps**

Part #	Input	$V_{os}$ typ max ( $\mu$ V)	$I_{bias}$ typ max (pA)	$V_{cm}$ neg lim (V)	$V_{out}$	$V_s$ (total) min max (V) (V)	$I_s$ typ ( $\mu$ A)	Price 100 pc (\$)	Notes
uA741	BJT	2000 6000	80k 500k	2	1.5V from rails	10 40	1500	0.27	old HV BJT
LF411	JFET	800 2000	50 200	3	1.5V from rails	10 40	1800	0.88	HV JFET
LM358A	BJT	2000 3000	45k 100k	0	0 to $V_+ - 1.2V$	3 32	500	0.21	old HV SS
LPV521	CMOS	100 1000	0.01 1	-0.1	R-R	1.8 5.5	0.5	1.05	low-bias LV RRIO
OPA336	CMOS	60 125	1 10	-0.2	R-R	2.3 5.5	20	1.70	chap 4 "solution"
ADA4051	CMOS	2 17	5 50	0	R-R	1.8 5	15	2.20	LV auto-zero RRIO
LTC6078	CMOS	7 25	0.2 1	0	R-R	2.7 5.5	110	1.75	LV low- $I_B$ low- $V_{os}$ RRIO dual
AD8603	CMOS	12 50	0.2 1	-0.3	R-R	1.8 5	40	1.40	LV low- $I_B$ low- $V_{os}$ RRIO
Ckt Limit	-	100	- 10	0	0 to anything	1.8 >3.6			must budget contributions

their guaranteed worst-case errors be added (as unsigned magnitudes) to determine the overall circuit's performance.

So far, so good; and in this example the op-amp choice allowed us to meet (and exceed) our target zero-error specifications (1% of full scale, with input open or shorted), even under guaranteed worst-case values of  $I_B$  and  $V_{os}$ .

#### A. Unspecified parameters: a pragmatic approach

But, looking a little closer, we've also seen in this design example that an unspecified parameter (forward diode current at low voltages, 0–10 mV) figured into the overall error.<sup>7</sup> What should one do under such circumstances?

The authors belong to the camp of “pragmatists” on this matter: first, you may have to be creative in reading the datasheet (as we did in the case of op-amp input current), particularly when the manufacturer's worst-case numbers represent a statement that “I don't want to test this parameter, so I'll put a conservative guess in the datasheet”; this is particularly relevant with parameters like leakage current, where the limits of automated test equipment (ATE) and testing time constraints encourage conservative worst-case datasheet specifications. Second, you may have to do some testing<sup>8</sup> of poorly specified (or unspecified) parameters (as we did with forward diode current). It may be sufficient to establish that the circuit effects of the unspecified parameter are completely insignificant (as here, when the current through the clamp diodes was less than 0.01 pA, or three orders of magnitude below budget); or, if it's a closer call,

you may have to set up a testing regime of incoming components to ensure that you meet specs. And third, you may have to deal with a situation in which there are many components contributing to the overall error budget by simply validating the performance of the overall circuit, subassembly, or complete instrument at final test.

This approach may appear cavalier. But the fact is that there are many situations in which you simply cannot meet challenging specifications while staying within the published worst-case specifications (or lack of specifications). Two examples help make this point: one of the authors designed and manufactured a line of battery-powered oceanographic instruments, intended for long-duration (from weeks to as long as a year) submerged observations and data logging. A typical instrument might have 200 or more 4000B-series CMOS ICs. The datasheet lists the 25°C quiescent current<sup>9</sup> as “0.04 $\mu$ A (typ), 10 $\mu$ A (max).” Great. So 200 of these puppies probably draw a total of 8 $\mu$ A, but they could draw (in a wildly improbable scenario) as much as 2 mA. A year's worth of operation would require 70 mAh (using typical values), but, under strict worst-case rules, we would have to allow for 17.5 Ah (amp-hours). Here's the rub: the substantial battery pack for these volume-constrained deep submergeable pressure housings provided only 5 Ah of capacity (with some safety margin of derating). And 80% of the battery capacity was budgeted for the sensors and recorders. So strict worst-case design would require quadrupling the battery pack (and

<sup>7</sup> As did the op-amp input current at modestly elevated temperatures.

<sup>8</sup> If you buy in large quantities the manufacturer may be willing to do these tests.

<sup>9</sup> Amusingly, the specs are the same for simple parts such as gates, or complex parts such as counters or arithmetic logic units.

expanding the pressure case), or, alternatively, removing a substantial volume of the instrumentation payload. The solution was (and still is) obvious: build the subcircuits, and test them for conforming quiescent current. They invariably worked just fine, and the testing served mostly to identify modules in which there was a defective component, usually caused by improper handling of the sensitive CMOS components.

A second example is a commercial instrument, namely a sensitive electrometer from Keithley. These things will measure currents down to *femtoamps* ( $10^{-15}\text{A}$ ), which requires a front-end stage of extraordinarily low bias current. They accomplish this with a JFET follower matched pair as input stage to a conventional precision op-amp, in a current-to-voltage configuration (the input is a summing junction, at zero volts). And to keep the gate current low, they operate the JFETs at a very low drain voltage of just  $+0.55\text{ V}$ , with the source terminal sitting just a fraction of a volt below the drain. Now, nowhere in the JFET datasheet will you find anything telling you what happens at such low voltages; and they won't tell you what the gate leakage is likely to be. You can throw up your hands and say that such an instrument cannot be made. Or you can do what Keithley did, which is to find a good source of JFETs and qualify them with in-house testing so that you can get on with the job.

Both examples illustrate that there are situations in which you simply cannot meet your design requirements while staying within the manufacturer's published worst-case specs. Having said that, we note that there are some engineers who will not deviate from strict worst-case specified component parameters in their circuit designs. They don't want to use special parts, and they won't touch such stuff with a 10-foot (3m) pole. We invite you to choose what you would do.

#### 5.4 Another example: precision amplifier with null offset

Having warmed up with the millivoltmeter, let's tackle a more complex design, one in which there are multiple error challenges. We describe the design choices and errors of this particular circuit within the framework of precision design in general, thus rendering painless what could otherwise become a tedious exercise.

We designed a precision amplifier (Figure 5.3) that lets you "freeze" the value of the input signal, amplifying any subsequent changes from that level by gains of exactly 1, 10, or 100. This might come in particularly handy in an experiment in which you wish to measure a small change in

some quantity (e.g., light transmission or RF absorption) as some condition of the experiment is varied. It is ordinarily difficult to get accurate measurements of small changes in a large dc signal, owing to drifts and instabilities in the amplifier. In such a situation a circuit of extreme precision and stability is required.

Here we show the example of a *strain gauge* sensor, which consists of a strain-sensitive resistive bridge whose elements change resistance (slightly!) in response to mechanical strain. A common resistance value is  $350\Omega$ ; and the sensitivity is such that, when biased with  $+5\text{V}$ , the differential output voltage across the bridge changes by  $\pm 10\text{mV}$  in response to the rated full-scale mechanical strain.<sup>10</sup> This small differential voltage sits on a dc level of  $+2.5\text{ V}$ , so you've got to begin with a good differential amplifier.

An important note at the outset: digital techniques offer an attractive alternative to the purely analog circuitry used here. A skilled designer would likely make use of precision analog/digital conversion techniques, perhaps in a hybrid implementation (in which a stable DAC is used to create the nulling signal within an analog circuit like ours), or perhaps in an all-digital scheme that relies on the intrinsic precision of a high resolution ADC alone.<sup>11</sup> Regardless, our all-analog example offers a smorgasbord of important lessons in precision design. But the reader can confidently look forward to exciting revelations in chapters to come.

#### 5.4.1 Circuit description

The front-end begins with an *instrumentation amplifier*  $U_1$ , a configuration of three op-amps that we'll talk about later (§5.15); these are differential-input amplifiers that excel in achieving high common-mode rejection, and allow gain selection with a single resistor (one or more are often provided internally). Here we've selected one with a good combination of low input current, offset drift, and noise, for reasons we'll explain later. Its gain of  $\times 100$  is followed with a noninverting  $\times 10$  gain stage ( $U_2$ ), for an overall gain of  $\times 1000$ ; that produces a full-scale output of  $\pm 10\text{ V}$  as input to the nulling circuitry ( $U_3-U_5$ ). If the input signal were single ended (e.g., from a thermocouple, photosensor, microwave absorption detector, or whatever), you would omit  $U_1$ , bringing in the signal at point "X," and adjusting the gain of  $U_2$  accordingly.

<sup>10</sup> The strain-gauge sensor sensitivity is "2mV-per-volt"; that's pretty low.

There are semiconductor strain sensors with higher sensitivity, but they may not be as stable.

<sup>11</sup> There's an example of the latter in §13.9.11C.

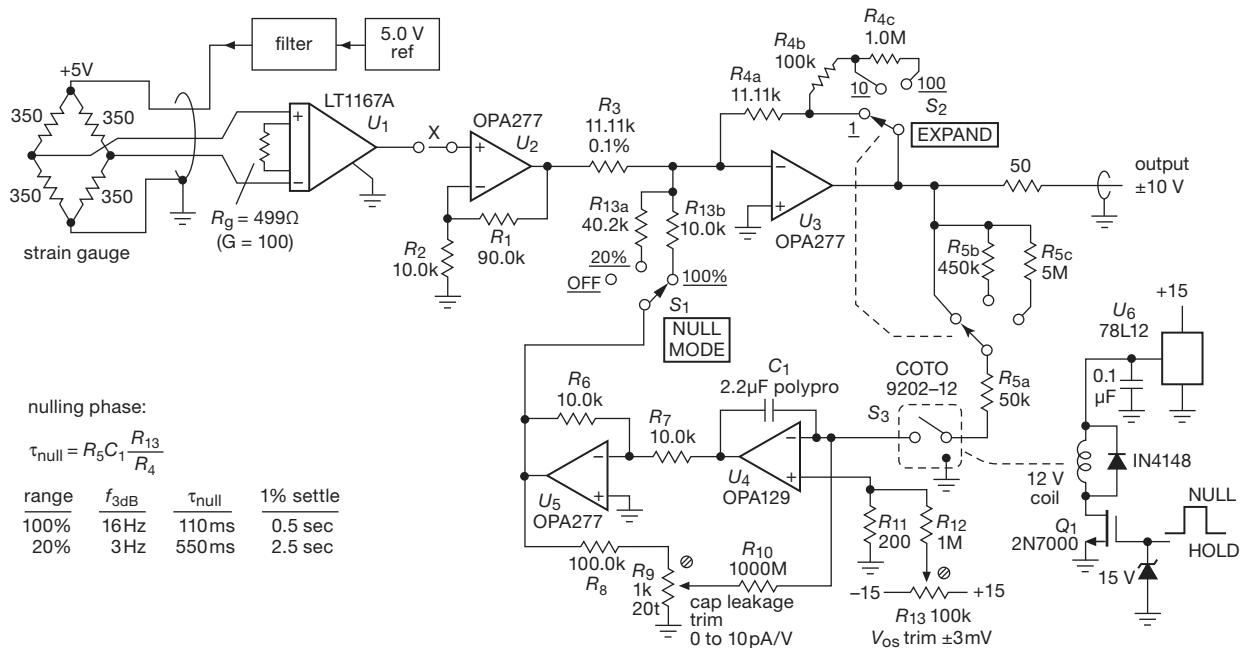


Figure 5.3. Autonulling dc laboratory amplifier. Gain-setting resistors are 0.1% tolerance.

The nailing circuitry works as follows: the amplifier stage  $U_3$  is configured in the inverting configuration, permitting a dc offset according to a current added to the summing junction. Nulling takes place when relay (switch)  $S_3$  is activated by a logic HIGH to coil driver  $Q_1$ 's gate. Then  $U_4$  charges the analog “memory” capacitor ( $C_1$ ) as necessary to maintain zero output. No attempt is made to follow rapidly changing signals, because in the sort of application for which this was designed the signals are essentially dc, and some averaging is a desirable feature. When the switch is opened, the voltage on the capacitor remains stable, resulting in an output signal from  $U_3$  proportional to the wanderings of the input thereafter. The gain about the input null level can be increased in decade steps (switch  $S_2$ ) to expand changes; the gain of the nailing integrator is switched accordingly to keep the feedback bandwidth constant. Switch  $S_1$  selects the full-scale nailing range (100%, 20%, or none).

There are a few additional features that we should describe before going on to explain in detail the principles of precision design as applied here: (a)  $U_5$ , in addition to providing the needed inversion of the nailing level, participates in a first-order leakage-current compensation scheme: the tendency of  $C_1$  to discharge slowly through its own leakage ( $\geq 100,000 \text{ M}\Omega$ , corresponding to a time constant of  $\geq 3$  days) is compensated by a small charging cur-

rent through  $R_{10}$  proportional to the voltage across  $C_1$ ; and (b) integrator  $U_4$  was selected for very low input current  $I_B$  (to minimize droop during “hold”), for which the tradeoff is relatively poor offset voltage  $V_{os}$ ; so we added an external offset trim ( $R_{11}-R_{13}$ ). This is not terribly critical, in any case, because an offset here merely causes a nonzero null of the same magnitude.

### 5.5 A precision-design error budget

For each category of circuit error and design strategy, we will devote a few paragraphs to a general discussion, followed by illustrations from the preceding circuit. Circuit errors can be divided into the categories of (a) errors in the external network components, (b) op-amp (or amplifier) errors associated with the input circuitry, and (c) op-amp errors associated with the output circuitry. Examples of the three are resistor tolerances, input offset voltage, and errors that are due to finite slew rate, respectively.

Let's start by setting out our error budget. It is based on a desire to keep input drift (from temperature and power-supply variations) down to the  $10 \mu\text{V}$  level, and nulled drift (primarily from capacitor “droop,” along with temperature and supply variations) below  $1 \mu\text{V}/\text{min}$  (referred to the input, or RTI). As with any budget, the individual items are arrived at by a process of tradeoffs, based on what can be

done with available technology. In a sense the budget represents the end result of the design, rather than the starting point. However, it will aid our discussion to have it now.

It's important to understand that the items in such a budget come from several sources: (a) parameters that are specified in the datasheet; (b) estimates of poorly specified (or unspecified) parameters; and (c) parameters that you may not even realize are important.<sup>12</sup> We might paraphrase these, respectively, as the *knowns*, the *known unknowns*, and the *unknown unknowns*.

### 5.5.1 Error budget

These are all in the form of worst-case voltage errors (at 25°C) referred to the instrument input.

#### 1. $\times 100$ Difference amplifier ( $U_1$ : LT1167A)

Offset voltage	40 $\mu$ V
Noise voltage (0.1–10 Hz)	0.28 $\mu$ Vpp (typ – no “max” spec)
Temperature	0.3 $\mu$ V/ $^{\circ}$ C
Power supply	28 nV/100 mV change
Input offset current $\times R_s$	0.11 $\mu$ V/350 $\Omega$ of $R_s$

#### 2. $\times 10$ Gain amplifier ( $U_2$ : OPA277)

Offset voltage	0.5 $\mu$ V
Temperature	10 nV/ $^{\circ}$ C
Time	2 nV/month (typ – no “max” spec)
Power supply	1 nV/100 mV change
Bias current	0.3 $\mu$ V
Load-current heating	5 nV at full scale (5 mW, 0.1 $^{\circ}$ C/mW)

#### 3. Output amplifier ( $U_3$ : OPA277)

Offset voltage	50 nV
Temperature	1 nV/ $^{\circ}$ C
Time	0.2 nV/month (typ – no “max” spec)
Power supply	0.1 nV/100 mV change
Bias current	30 nV
Load-current heating	5 nV at full scale (1 k $\Omega$ load)

#### 4. Hold amplifier ( $U_4$ : OPA129)

$U_4$ offset tempco	10 nV/ $^{\circ}$ C
Power supply	10 nV/100 mV change
Capacitor droop	0.4 $\mu$ V/min
(see current error budget)	

Charge transfer	1.1 nV
-----------------	--------

Current errors through  $C_1$  (needed for the preceding voltage error budget) are as follows:

Capacitor leakage	
Maximum (uncompensated)	(100 pA)
Typical (compensated)	10 pA
$U_4$ input current	0.25 pA
$U_4$ 's nulled $V_{OS}/R_{10}$	0.1 pA
Relay $S_3$ OFF leakage	10 pA (1 pA typ)
Printed-circuit-board leakage	5.0 pA

Not bad, though you *could* complain about the 40  $\mu$ V input offset – but we'd reply that a few tens of microvolts of *static* offset is of no concern in a nulling instrument, it's only the *drift* (with time and temperature) that matters. The various items in the budget will make sense as we discuss the choices faced in this particular design. We organize these by the categories of circuit errors listed earlier: network components, amplifier input errors, and amplifier output errors.

We'll address these errors quantitatively, in the context of Figure 5.3, beginning in §5.7.6, after looking at error sources a bit more generally in the next section.

### 5.6 Component errors

The degrees of precision of reference voltages, current sources, amplifier gains, etc., all depend on the accuracy and stability of the resistors used in the external networks. Even where precision is not involved directly, component accuracy can have significant effects, e.g., in the common-mode rejection of a differential amplifier made from an op-amp (see §§4.2.4 and 5.14), where the ratios of two pairs of resistors must be accurately matched. The accuracy and linearity of integrators and ramp generators depend on the properties of the capacitors used, as do the performances of filters, tuned circuits, etc. As we will see presently, there are places where component accuracy is crucial, and there are other places where the particular component value hardly matters at all.

Components are generally specified with an initial accuracy, as well as the changes in value with time (stability) and temperature. In addition, there are specifications of voltage coefficient (nonlinearity) and bizarre effects such as “memory” and dielectric absorption (for capacitors). Complete specifications also include the effects

<sup>12</sup> We discovered an example of the latter while measuring femtoamp leakage currents in a nice shielded test enclosure: after the box had been opened to change anything inside, it took quite a while for the measurements to settle down. Turns out that the process of moving things caused some rearrangement of surface charge on the Teflon-insulated wires, with a long relaxation time. Pease talks about this in his article “What's All This Teflon Stuff, Anyhow?” – see the footnote references in §5.10.7. We've experienced a similar bizarre manifestation with analog panel meters, where a swipe of the hand across the glass face can cause the needle to move way upscale... and stay there!

of temperature cycling and soldering, shock and vibration, short-term overloads, and moisture, with well-defined conditions of measurement. In general, components of greater initial accuracy will have their other specifications correspondingly better, in order to provide an overall stability comparable to the initial accuracy. However, the overall error that is due to all other effects combined can exceed the initial accuracy specification. Beware!

As an example, RN55C 1% tolerance metal-film resistors have the following specifications: temperature coefficient (tempco), 50 ppm/ $^{\circ}\text{C}$  over the range  $-55^{\circ}\text{C}$  to  $+175^{\circ}\text{C}$ ; soldering, temperature, and load cycling, 0.25%; shock and vibration, 0.1%; moisture, 0.5%. By way of comparison, the legacy 5% carbon-composition resistors (Allen-Bradley type CB) have these specifications: tempco, 3.3% over the range 25–85°C; soldering and load cycling, +4%, -6%; shock and vibration, ±2%; moisture, +6%. From these specs it should be obvious why you can't just select (using an accurate digital ohmmeter) carbon resistors that happen to be within 1% of their marked value for use in a precise circuit, but are obliged to use 1% resistors (or better) designed for long-term stability as well as initial accuracy. For the utmost in precision it is necessary to use ultra-precise resistors or resistor arrays, such as Susumu's RG-series of SMT (surface-mount technology) resistors (tolerance to 0.02%, tempco to 5 ppm/ $^{\circ}\text{C}$ ), Vishay's MPM-series metal-film networks (absolute tolerance to 0.05%, matched to 0.01%; absolute tempco to 25 ppm/ $^{\circ}\text{C}$ , matching tempco to 2 ppm/ $^{\circ}\text{C}$ ), or their even better "Bulk Metal Foil" types (absolute tolerance to 0.005%, matched to 0.001%; absolute tempco to 0.2 ppm/ $^{\circ}\text{C}$ , matching tempco to 0.1 ppm/ $^{\circ}\text{C}$ ).

### 5.6.1 Gain-setting resistors

In the preceding circuit (Figure 5.3), 0.1% resistors are used in the gain-setting network,  $R_1$ – $R_4$ , for accurately predictable gain. As we'll see shortly, the value of  $R_3$  is a compromise, with small values reducing offset current error in  $U_3$  but increasing heating and thermal offsets in  $U_2$ . Note that 1% resistors are used in the offset attenuator network,  $R_5$ – $R_{13}$ ; here absolute accuracy is irrelevant, and the stability of 1% metal-film resistors is altogether adequate.

### 5.6.2 The holding capacitor

#### A. Leakage

The largest error term in this circuit, as the error budget shows, is capacitor leakage in the holding capacitor,  $C_1$ . Capacitors intended for low-leakage applications give a leakage specification, sometimes as a leakage resistance,

sometimes as a time constant (megohm-microfarads). In this circuit  $C_1$  must have a value of at least a few microfarads in order to keep the charging rate from other current error terms small (see budget). In that range of capacitance, film capacitors (polystyrene, polypropylene, and polyester) have the lowest leakage. Polypropylene capacitors (from manufacturers such as Epcos, Kemet, Panasonic, Vishay, and Wima, generally with voltage ratings of 200–600 V) often have dc leakage specified in units of megohm-microfarads, with values in the range of 10,000–100,000  $\text{M}\Omega \mu\text{F}$ ; thus for a capacitance of 2.2  $\mu\text{F}$  this amounts to a parallel equivalent leakage resistance of at least 4.5–45 G $\Omega$ .

Even so, and adopting a plausible value of, say, 100 G $\Omega$ , that's equivalent to a leakage current of 100 pA at full output (10 V), corresponding to droop rates of roughly 3 mV/min at the output, the largest error term by far. For that reason we added the leakage-cancellation scheme described earlier. It is fair to assume that the effective leakage can be reduced to 10% of the capacitor's worst-case leakage specification (in practice, we can probably do much better). No great stability is required in the cancellation circuit, given the modest demands made of it. As we will see later when we discuss voltage offsets,  $R_{10}$  is kept intentionally large so that input voltage offsets in  $U_4$  aren't converted to a significant current error.

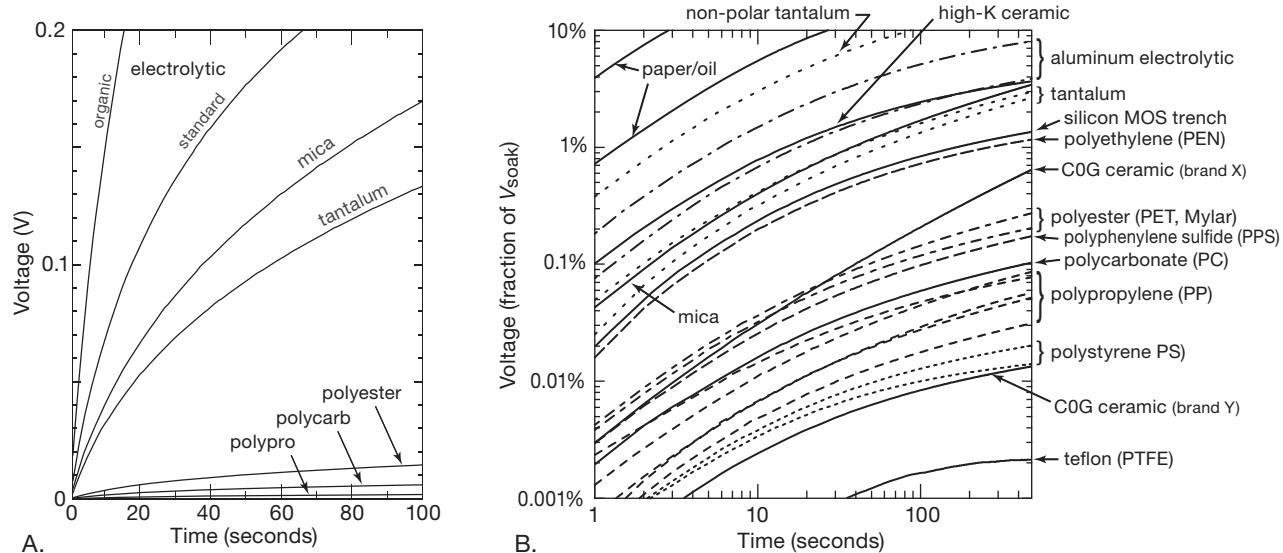
#### B. Dielectric absorption

We're not done with the capacitor, yet. An important effect, quite apart from resistive leakage, is capacitor "memory," officially known as *dielectric absorption*.<sup>13</sup> This is the tendency of capacitors to return, to some extent, to a previous state of charge, as shown in the measured data of Figure 5.4 (each capacitor was held at +10 V for a day or more, then discharged to 0 V for 10 s, then open-circuited and observed while it did its thing); see also the discussion in §§1x.3, 4.5.1, 4.5.6, and 13.8.4.

#### 5.6.3 Nulling switch

In the previous edition of this book, the analogous circuit (Figure 7.1, page 393) used MOSFETs (instead of relay  $S_3$ ) to activate the nulling circuit. That choice provided plenty of education, because we had to worry about (a) MOSFET channel leakage, devastatingly large at about  $\sim 1\text{ nA}$ , and (b) gate charge injection, of order 100 pC in that circuit. The solutions there were (a) the use of a series-connected

<sup>13</sup> It is not entirely clear that what is called "leakage" in high-quality capacitors is in fact distinct from dielectric absorption; see the footnote in §4.5.5.



**Figure 5.4.** Capacitors exhibit *memory effect* (dielectric absorption), a tendency to return to a previous state of charge. This is highly unhelpful in applications (such as analog sample-and-hold) in which a capacitor is used to retain an analog voltage. A. linear plot, showing the basic effect; B. log–log plot, revealing four decades of dirty laundry. Teflon is the uncontested winner; but it's hard to find, so the plastic film types (PS and PP) are generally your best choice. Ceramic COG can be excellent, but beware brand variations.

MOSFET pair, such that the downstream MOSFET had all four terminals (source, drain, gate, and substrate) ordinarily at zero volts, and (b) a sufficiently large holding capacitor such that the error was negligible, along with the observation that charge transfer was not of great concern because it resulted in a small offset of the auto-zero.

This time we have taken a more pragmatic (but less educational) approach, by using instead a small signal *relay*. The Coto 9202-12 is a small ( $4 \text{ mm} \times 6 \text{ mm} \times 18 \text{ mm}$ ) shielded relay, energized by 12 Vdc at 18 mA, with a specified OFF resistance of  $10^{12} \Omega$  minimum ( $10^{13} \Omega$  typical). The worst-case  $R_{\text{off}}$  value corresponds to a droop rate of 0.3 mV/min, but ten times less for “typical”  $R_{\text{off}}$ . Relays isolate better than transistor switches (higher  $R_{\text{off}}$  and lower  $C_{\text{off}}$ , here  $< 1 \text{ pF}$ ), and they have better ON performance as well (lower  $R_{\text{on}}$  than a low-capacitance analog switch, here  $< 0.15 \Omega$ ).

Of course, there *is* capacitance between the coil and the contacts, and thus an opportunity for the same sort of charge transfer as with a MOSFET switch (where the full-swing transitions at the gate couple capacitively to the drain and source). As we remarked in Chapter 3 (§3.4.2E), the total charge transferred is independent of the transition time and depends only on the total control-voltage swing and the coupling capacitance:  $\Delta Q = C_{\text{coup}} \Delta V_{\text{control}}$ . In this circuit, charge transfer results in a simple voltage error of the auto-zero, because the charge is converted to a volt-

age in the holding capacitor  $C_1$ . It's easy to estimate the error: the Coto relay specifies a coil-to-contact capacitance of 0.2 pF (for our grounded-shield configuration), and thus a corresponding charge transfer of  $\Delta Q = 2.4 \text{ pC}$  when the 12 V coil is energized.<sup>14</sup> That produces a voltage step of  $\Delta V_C = \Delta Q / C_1 = 1.1 \mu\text{V}$  across the  $2.2 \mu\text{F}$  capacitor  $C_1$ . This is completely within our error budget; in fact, we've likely overestimated the effect, because our calculation assumed that the entire coil undertook a 12 V step, whereas the average step is half that value.

For readers who harbor a deep-seated dislike of mechanical relays, we show in Figure 5.5 a switch implementation with series-connected JFETs. During HOLD the JFET gates are back-biased to  $-5 \text{ V}$ , by the somewhat tortured level-shifting circuit  $Q_1$ – $Q_4$ . The reader is invited to estimate the magnitude of droop (use the datasheet's  $I_{D(\text{off})}=0.1 \text{ pA}$ ) and of charge injection (use the datasheet's  $C_{\text{rss}}=0.3 \text{ pF}$ ) for this circuit.

## 5.7 Amplifier input errors

The deviations of op-amp input characteristics from the ideal that we discussed in Chapter 4 (finite values of input impedance and input current, voltage offset, common-mode rejection ratio, and power-supply rejection ratio, and

<sup>14</sup> Assuming care is taken in the wiring layout to maintain the low 0.2 pF capacitance of the HOLD signal.

Table 5.2 Representative Precision Op-amps

Part #	# per pkg <sup>a</sup>	Supply <sup>p</sup>		Input Current @25°C		Offset Voltage		CMRR min	$e_n$ @1kHz dB	GBW typ (nV/Hz)	Slew typ (MHz)	Swing to Supply IN	null pins	Cost DIP avail qty 25 (\$US)	Comments
		Range (V)	$I_Q$ (mA)	typ (pA)	max (pA)	$V_{os}$ typ ( $\mu$ V)	max ( $\mu$ V)								
<i>bipolar</i>															
LT1077A	1	2.2–44	0.05	7nA	9nA	9	40	0.4	97	27	0.23	0.08	- • - • • •	3.84	single-supply
LT1013	2,4	3.4–44	0.35	12nA	20nA	40	150	0.4	100	22	0.7	0.4	- • - • - •	3.13	single-supply
OPA277P	1,2,4	4–36	0.79	0.5nA	1nA	10	20	0.1	130	8	1	0.8	- - - - • •	3.17	improved OP-27
LT1012AC	1	8–40	0.37	25pA	0.1nA	8	25	0.2	114	14	0.5	0.2	- - - - • •	5.11	superbeta, comp
LT1677	1	3–44	2.8	2nA	20nA	20	60	0.4	109	3.2	7.2	2.5	• • • • • •	3.07	or AD8675, 0.5nA
LT1468	1	7–36	3.9	3nA	10nA	30	75	0.7	96	5	90	23	- - - - • •	4.26	0.7ppm distortion
<i>JFET</i>															
LF412A <sup>b</sup>	1,2	12–44	1.8	50	200	500	1000	-	80	25	4	15	- - - - - •	4.47	'411A $V_{os} < 0.5mV$
OPA827	1	8–40	4.8	15	50	75	150	1.5	104	3.8	22	28	- - - - - -	9.00	0.1–10Hz: 0.25 $\mu$ Vpp
<i>CMOS</i>															
LMC6482A <sup>b</sup>	2,4	3–16	0.5	0.02	4	110	750	1	70	37	1.5	1.3	• • • • - •	1.88	LMC7101 = SOT23
MAX4236A	1	2.4–6	0.35	1	500	5	20	0.6	84	14	1.7	0.3	- • • • - -	1.78	shdn, '37 decomp
OPA376	1,2,4	2.2–7	0.76	0.2	10	5	25	0.26	76	7.5	5.5	2	• • • • - -	1.32	etrim
<i>auto-zero</i>															
AD8628	1,2,4	2.7–6	0.85	30	100	1	5	0.002	120	22	2.5	1	• • • • - -	1.92	SOIC-8, SOT23-5

Notes: (a) boldface indicated number in a package for the part # listed. (b) not precision, listed for comparison. (p)  $I_Q$ , typical, per amplifier.

their drifts with time and temperature) generally constitute serious obstacles to precision circuit design, and they force trade-offs in circuit configuration, component selection, and the choice of a particular op-amp. The point is best made with examples, as we will do shortly. Note that these errors, or their analogs, exist for amplifiers of discrete design as well.

While reading through the following discussion, you may find it helpful to refer to Tables 5.2 and 5.3 where we list ten favorite precision op-amps (plus two inexpensive and not-precision comparees). Those listings add quantitative flesh to a somewhat abstract set of bony teachings.

### 5.7.1 Input impedance

Let's discuss briefly the error terms just listed. The effect of finite input impedance is to form voltage dividers in combination with the source impedance driving the amplifier, reducing the gain from the calculated value. Most often this isn't a problem, because the input impedance is bootstrapped by feedback, raising its value enormously. As an example, the OPA277P precision op-amp (with BJT-input, not FET-input stage) has a typical "differential-mode input impedance," of 100 M $\Omega$ . In a circuit with plenty of loop gain, feedback raises the input impedance to the datasheet's "common-mode input impedance" 250,000 M $\Omega$ . If that's not good enough, FET-input op-amps have astronomical values of  $R_{in}$ , for example 10<sup>13</sup>  $\Omega$  (differential) and 10<sup>15</sup>  $\Omega$  (common-mode) for the OPA129 that's used in this circuit.

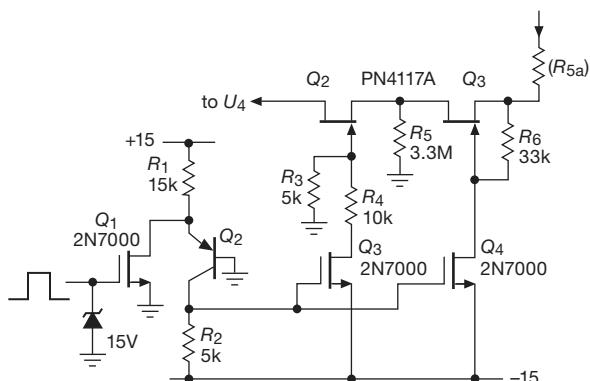


Figure 5.5. Electronic switch replacement of mechanical relay  $S_3$  in Figure 5.3. A lot of work, and no improvement in performance.

### 5.7.2 Input bias current

More serious is the input bias current. Here we're talking about currents measured in nanoamps, and this already produces voltage errors of microvolts for source impedances as small as 1 k $\Omega$ . Again, FET op-amps come to the rescue, but with generally increased voltage offsets as part of the bargain. Bipolar superbeta op-amps such as the LT1012 can also have surprisingly low input currents. As an example, compare the OPA277 precision bipolar op-amp with the LT1012 (bipolar, optimized for low bias current), the OPA124 (JFET, precision and low bias), the OPA129 (ultra-low-bias JFET), and the LMC6001 (CMOS, lowest-bias op-amp); these are some of the best you can get at the

**Table 5.3 Nine Low-input-current Op-amps**

Part #	Supply		Input current		$V_{os}$ max ( $\mu$ V)	$TCV_{os}$ typ ( $\mu$ V/ $^{\circ}$ C)	$TCV_{os}$ max ( $\mu$ V/ $^{\circ}$ C)
	$V_{total}$ (V)	$I_Q$ ( $\mu$ A)	typ (pA)	max (pA)			
<i>bipolar</i>							
OPA277P	10–36	790	500	1000	20	0.1	0.15
<i>superbeta</i>							
LT1012AC	8–40	370	25	100	25	0.2	0.6
AD706	4–36	750	50	200	100	0.2	1.5
<i>JFET</i>							
OPA124PB	10–36	2500	0.35	1	250	1	2
OPA129B	10–36	1200	0.03	0.1	2000	3	10
<i>MOSFET</i>							
MAX9945	4.8–40	400	0.05	-	5000	2	-
<i>CMOS, low-voltage</i>							
LMP7721	1.8–6	1300	0.003	0.02	150	1.5	4
LMC6001A	5–16	450	0.01	0.025	350	2.5	10
ADA4530-1	4.5–16	900	<0.001	0.02	50	0.13	0.5

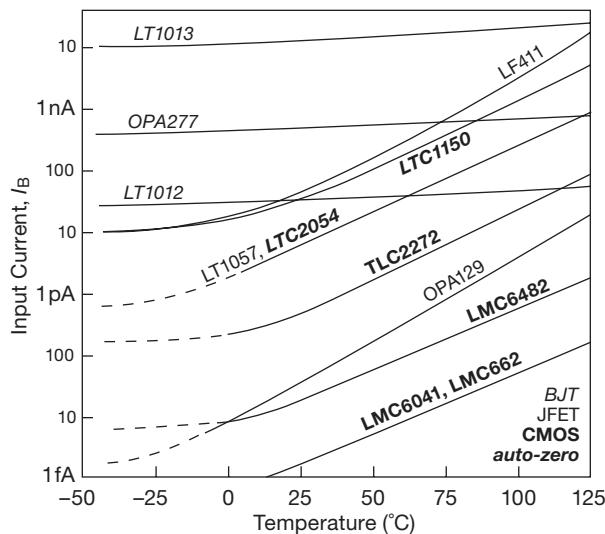
time of writing, and we've chosen the best grade of each one (Table 5.3; and see also Table 5.5 on pages 320–321 for greater detail, along with a wider selection of precision op-amps with low offset voltage; as well as relevant tables in Chapter 4x).

Well-designed FET amplifiers have extremely low bias currents, but with much larger offset voltages, compared with the precision OPA277. Because the offset voltage can always be trimmed, what matters more is the drift with temperature. In this case the FET amplifiers are 4 to 20 times worse. The op-amp with the lowest input current uses MOSFETs for the input stage. MOSFET op-amps are popular because of the proliferation of inexpensive units like TI's TLC270 series, as well as the ultra-low-bias-current devices like National's LMC6000-series parts. However, in contrast to JFETs or bipolar transistors, MOSFETs can have very large drifts of offset voltage with time, an effect that is discussed below. So the improvement in current errors you buy with a FET op-amp can be wiped out by the larger voltage error terms. With any circuit in which bias current can contribute significant error, it is often wise to ensure that both op-amp input terminals see the same dc source resistance (see, e.g., Figure 4.55); then the op-amp's *offset current* becomes the relevant specification. But be aware that a number of precision op-amps use a "bias-compensation" scheme to cancel (approximately) the input current in order to make that error term smaller (look back at Exercise 2.24 on page 125 to see how it's done). With op-amps of this type you generally don't gain anything by matching the dc resistances seen by the two inputs, as the residual bias current and the offset current are comparable, in a bias-compensated op-amp.

### A. Variation with temperature

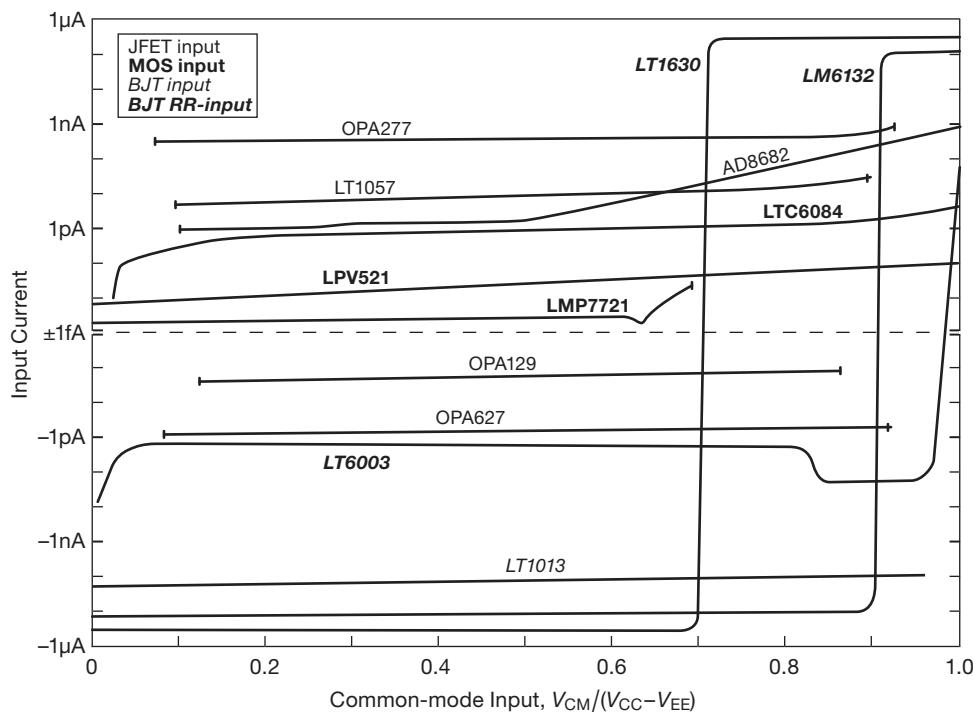
One additional point to keep in mind when using FET-input op-amps is that the input "bias" current is actually gate

leakage current, and it rises dramatically with increasing temperature: it roughly doubles for every  $10^{\circ}$ C increase in chip temperature, as seen in Figure 5.6. Because FET op-amps often run warm (our jellybean LF412, for example, dissipates 100 mW when run from  $\pm 15$  V supplies), the actual input current may be considerably higher than the  $25^{\circ}$ C figures you see on the datasheet.<sup>15</sup> By contrast, the input current of a BJT-input op-amp is actual base current, relatively constant with temperature. So a FET-input op-amp with impressive input-current specs on paper may not give such an improvement over a good superbeta bipolar unit. As the graph shows, for example, the LT1057 JFET-input op-amp with its  $\sim 3$  pA input current (at  $25^{\circ}$ C) will have an input current of about 100 pA at  $75^{\circ}$ C chip temperature, which is higher than the input current of the superbeta LT1012 at the same temperature. And our jellybean LF412 JFET op-amp has an input current that is comparable to that of the LT1012 at  $25^{\circ}$ C and many times higher at elevated temperatures.



**Figure 5.6.** Op-amp input current versus temperature, plotted from datasheet values. See also Figures 5.38 and 3.48. JFET-input op-amps are indicated in plain ("roman") type, BJT op-amps are *italic*, CMOS op-amps are **bold**, and auto-zero op-amps are **bold slant**.

<sup>15</sup> Making this quantitative, the LF412's maximum quiescent current is 6.5 mA, thus 195 mW dissipation when run from  $\pm 15$  V supplies. In a DIP-8 package that produces a  $22^{\circ}$ C temperature rise (the thermal resistance  $R_{\Theta JA}=115^{\circ}$ C/W), with a consequent quadrupling of the specified  $I_B=200$  pA (max). If the op-amp were driving a load, you'd have yet more dissipation. To put this in perspective, though, note that the driving impedance seen at the op-amp's input would have to be greater than 1 M $\Omega$  in order for this current-induced error to exceed the 1 mV (typ) input offset-voltage error.



**Figure 5.7.** Op-amp input current (at 25°C) versus common-mode input voltage, plotted over their operating range from datasheet values; BJT op-amps with rail-to-rail input stages (“BJT RR-input”) undergo an abrupt reversal of input-current polarity.

### B. Variation with common-mode input voltage

Finally, a very important caution: when comparing op-amp input currents, watch out for some op-amp designs whose  $I_B$  depends on the input voltage. This behavior is common in op-amps that are designed to operate over a rail-to-rail input (RRI) range, both FET-input and (especially) BJT-input types. The spec sheet usually lists  $I_B$  only at zero volts (or mid-supply), but a good datasheet will show curves as well. See Figure 5.7 for some typical  $I_B$  versus  $V_{in}$  behavior. The good performance of the OPA129 and OPA627 in this regard is due in part to their use of cascode input stages. The LMP7721 stands out not only for its 20 fA maximum input current, but for its prize-winning trajectory on this graph.

#### 5.7.3 Voltage offset

Voltage offsets at the amplifier input are obvious sources of error. Op-amps differ widely in this parameter, ranging from “precision” op-amps offering worst-case  $V_{OS}$  values generally in the 10s of microvolts to ordinary jellybean op-amps like the LF412 with  $V_{OS}$  values of 2–5 mV. At the time of writing,<sup>16</sup> the champion (by a slim margin) in the

(nonchopper, see below) world of low offsets is the bipolar OPA277P ( $\pm 20 \mu\text{V}$ , max), which, astonishingly, is equaled by the CMOS MAX4236A (though the latter’s drift is 12× worse, as one might expect).

Although many good single op-amps (but not duals or quads) have offset-adjustment terminals, it is still wise to choose an amplifier with inherently low initial offset  $V_{OS \text{ max}}$ , for several reasons. First, op-amps designed for low initial offset tend to have correspondingly low offset drift with temperature and with time. Second, a sufficiently precise op-amp eliminates the need for external trimming components (a trimmer takes up space, needs to be adjusted initially, and may change with time). Third, offset voltage drift and common-mode rejection are degraded by the unbalance caused by an offset-adjustment trimmer.

Figure 5.8 illustrates how a trimmed offset has larger drifts with temperature. It shows also how the offset adjustment is spread over the trimmer pot rotation, with best

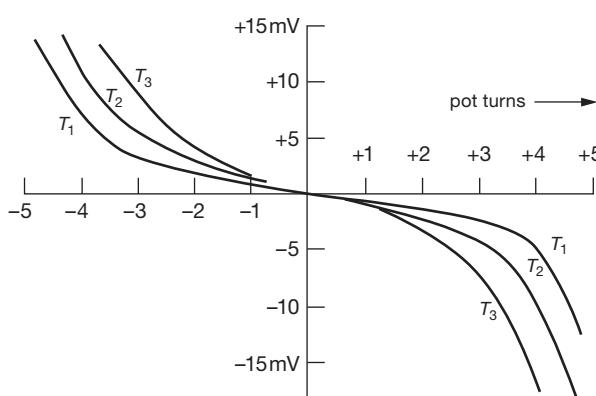
dence we added that “we expect to see further incremental improvements in this area.” That confidence was evidently misplaced: the Maxim website now says of the MAX400 “This product was manufactured for Maxim by an outside wafer foundry using a process that is no longer available. It is not recommended for new designs. The datasheet remains available for existing users.” *Sic transit...*

<sup>16</sup> In the previous edition of this book we awarded that honor to the MAX400M, with its specified worst-case  $V_{OS}$  of 10  $\mu\text{V}$ . With confi-

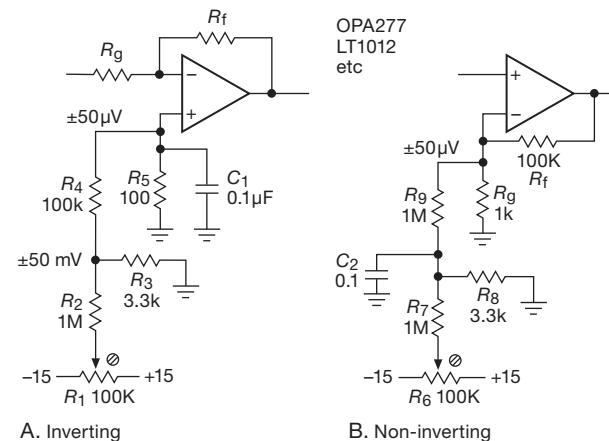
resolution near the center, especially for large values of trimmer resistance. Finally, you'll generally find that the recommended external trimming network provides far too much range, making it nearly impossible to trim  $V_{OS}$  down to a few microvolts; even if you succeed, the adjustment is so critical it won't stay trimmed for long. Another way to think about it is to realize that the manufacturer of a precision op-amp has *already* trimmed the offset voltage, in a custom test jig using "laser-zapping" techniques; you may be unable to do any better yourself. Our advice is (a) to use precision op-amps for precision circuits, and (b) if you must trim them further, arrange a narrow-range trim circuit similar that shown in Figure 5.3, with values adjusted to produce a full-scale range of  $\pm 50 \mu\text{V}$ , linear in trimmer rotation (e.g.,  $R_{11}=33 \Omega$  and  $R_{12}=10\text{M}$ ). Figure 5.9 shows how to arrange narrow-range external offset trims for both inverting and noninverting amplifier configurations.

Because voltage offsets can be trimmed to zero, what ultimately matters is the drift of offset voltage with time, temperature, and power-supply voltage. Designers of precision op-amps work hard to minimize these errors. You get the best performance from bipolar-input (as opposed to FET) op-amps in this regard, but input-current effects may then dominate the error budget. As shown in Table 5.2 on page 302 the best op-amps keep drifts below  $1 \mu\text{V}^\circ\text{C}$ ; the OPA277P typifies the best drift specification (for a non-chopper op-amp):  $\Delta V_{OS}=0.2 \mu\text{V}^\circ\text{C}$ , max.

Another factor to keep in mind is the drift caused by self-heating of the op-amp when it drives a low-impedance load. It is often necessary to keep the load impedance above 10k to prevent large errors from this effect. As usual, that may compromise the next stage's error budget from the effects of bias current! We'll see just such a problem in



**Figure 5.8.** Typical op-amp offset versus offset-adjustment potentiometer rotation for several temperatures.



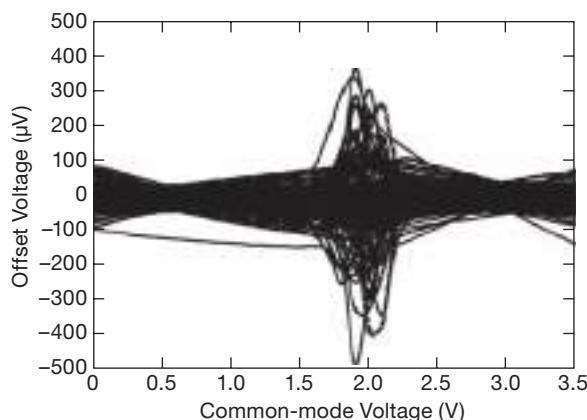
**Figure 5.9.** Narrow-range external trimming networks for precision op-amps.

this design example. For applications in which drifts of a few microvolts are important, the related effects of thermal gradients (from nearby heat-producing components) and thermal emf's (from voltages across junctions of dissimilar metals) become important. This will come up again when we discuss the ultraprecise *chopper-stabilized* amplifier in §5.11.

*An important caution:* when datasheets specify the particular measurement conditions for a parameter like  $V_{OS}$ , they mean it! A sobering example is shown in Figure 5.10, a plot of  $V_{OS}$  versus  $V_{CM}$  for the AD8615 op-amp, whose datasheet declares (on the front page) "Low offset voltage: 65  $\mu\text{V}$  max," but whose tabular data reveal that the measurement conditions are " $V_{CM} = 0.5 \text{ V}$  and  $3.0 \text{ V}$ ".

#### 5.7.4 Common-mode rejection

Insufficient common-mode rejection ratio (CMRR) degrades circuit precision by effectively introducing a voltage offset as a function of dc level at the input. This effect is usually negligible, because it is equivalent to a small gain change, and in any case it can be overcome by choice of configuration: an inverting amplifier is insensitive to op-amp CMRR, in contrast with a noninverting amplifier. However, in "instrumentation amplifier" applications you are looking at a small differential signal riding on a large dc offset, and a high CMRR is essential. In such cases you have to be careful about circuit configurations and, in addition, you must choose an op-amp with a high-CMRR specification. Once again, a superior op-amp like the OPA277 can solve your problems, with a CMRR (min)



**Figure 5.10.** This op-amp specifies a maximum offset voltage of  $\pm 60 \mu\text{V}$ . But it also specifies the following conditions:  $V_s = 3.5 \text{ V}$ , and  $V_{CM} = 0.5 \text{ V}$  or  $3.0 \text{ V}$ . Moral: don't ignore the footnotes!

at dc of 130 dB, compared with our jellybean LF411's meager specification of 70 dB. We discuss high-gain differential and instrumentation amplifiers later in the chapter, beginning at §5.13.

### 5.7.5 Power-supply rejection

Changes in power-supply voltage cause small op-amp errors. As with most op-amp specifications, the power-supply rejection ratio (PSRR) is referred to a signal at the *input*. For example, the OPA277 has a specified PSRR of 126 dB at dc, meaning that a 1 volt change in one of the power-supply voltages causes a change at the output equivalent to a change in the differential input signal of  $0.5 \mu\text{V}$ .

The PSRR drops with increasing frequency, approximately tracking the behavior of the open-loop gain, and a graph documenting this scurrilous behavior is often given on the datasheet. For example, the PSRR (relative to the negative rail) of our favorite OPA277 begins dropping at 1 Hz and is down to 95 dB (typ) at 60 Hz and 50 dB at 10 kHz. This rarely presents a problem, because power-supply noise is also decreasing at higher frequencies if you have used good bypassing. However, 120 Hz ripple could present a problem if an unregulated supply is used.

It is worth noting that the PSRR will not, in general, be the same for the positive and negative supplies. Thus the use of dual-tracking regulators doesn't necessarily bring any benefits. Note also that PSRR is often specified for  $G = 1$ , and can be considerably worse at higher gains; in fact, op-amps have been found that exhibit *gain* (!) from one rail to the output at moderate gain settings.

### 5.7.6 Nulling amplifier: input errors

Now we're ready to embark on a detailed discussion of the most serious error issues in the amplifier of Figure 5.3. The circuit begins with an optional precision instrumentation amplifier front-end  $U_1$  (more in §5.15), here chosen for its stable and accurate differential gain of  $100\times$ , low input current, and adequately low noise ( $9 \text{ nV}/\sqrt{\text{Hz}}$  typ at 10 Hz). Its worst-case offset voltage and tempco ( $\pm 40 \mu\text{V}$ ,  $0.3 \mu\text{V}/^\circ\text{C}$ ) specs are a factor of two worse than a precision op-amp like the OPA277 (in the best grade), but its 120 dB (min) CMRR as a difference amplifier, combined with 0.08% worst-case gain accuracy, 50 ppm/ $^\circ\text{C}$  (max) gain tempco, and low-voltage noise, make it a good front-end for a low-level bridge application like this. Although not important with the low source impedance in this example, its input current is satisfactorily low for a BJT-input amplifier, at just 0.35 nA max.<sup>17</sup>

For single-ended inputs  $U_1$  is omitted, and the signal is brought in at point “ $\times$ ” (add a  $470 \Omega$  series resistor, with a pair of low-leakage clamp diodes – see Figure 5.2 – to the rails, for overdrive protection). The OPA277's accuracy and stability rule here, though it is tempting to consider substituting a FET-input part; but the  $\gtrsim 10\times$  poorer  $V_{OS}$  tempco specification more than offsets the advantage of low input current, except with sources of very high impedance. The OPA277's 1 nA (max) bias current gives an error of  $1 \mu\text{V}/1 \text{k}\Omega$  source impedance, whereas the best-in-class JFET OPA627B (at \$35 apiece!), though providing negligible current error with its 5 pA (max) input current, would exhibit voltage offset drifts as large as  $3 \mu\text{V}/4^\circ\text{C}$  ( $4^\circ\text{C}$  is considered a typical laboratory ambient temperature range of variation). In this circuit one could happily add an offset trim to  $U_2$ , preferably in the manner of Figure 5.9. As mentioned earlier, feedback bootstraps the input impedance to  $250 \text{ G}\Omega$  and eliminates any gain errors from finite source impedance, up to  $25 \text{ M}\Omega$  (for a gain error less than 0.01%).

$U_2$  drives an inverting amplifier ( $U_3$ ), with  $R_3$  chosen as a compromise between heat-produced thermal offsets in  $U_2$  and bias-current offset errors in  $U_3$ . The value chosen keeps heating down to 5 mW (at 7.5 V output, the worst case), which works out to a temperature rise of  $0.8^\circ\text{C}$  (the op-amp has a thermal resistance  $R_{\Theta JA}$  of about  $0.15^\circ\text{C}/\text{mW}$ , see §9.4), with a consequent maximum voltage offset of  $\Delta V_{OS} = T C V_{OS} \Delta T = 0.12 \mu\text{V}$ . The  $11 \text{ k}\Omega$

<sup>17</sup> In fact, if noise is of primary concern you could substitute the  $\times 4$  quieter INA103 instrumentation amplifier at the front-end, paying the price in input offset current: a whopping  $1 \mu\text{A}$  (thus  $\pm 350 \mu\text{V}$ ) of static offset voltage created by the  $350 \Omega$  differential source resistance here.

source impedance seen by  $U_3$  results in an error due to bias-current offset, but, with  $U_3$  inside a feedback loop with  $U_4$  and  $U_5$  trimming the overall offset to zero, all that matters is the drift in the current error term. The OPA277 provides a graph of typical bias-current change with temperature (not often specified by manufacturers), from which the error result of  $0.2 \mu\text{V}/4^\circ\text{C}$  in the error budget is calculated. Reducing the value of  $R_3$  would improve this term, at the expense of the heating term in  $U_2$ .

The dc input impedance of  $U_3$  comes closer to presenting a problem. To estimate the error, we compare  $U_3$ 's differential input impedance of  $100 \text{ M}\Omega$  with the worst-case (i.e., with gain set for  $\times 100$ ) impedance seen driving its input. The latter is just the feedback resistance ( $1\text{M}$ ) divided by the loop gain  $G_{\text{OL}}/G_{\text{CL}}$ , thus  $10\Omega$ . So the worst-case loading effect is 1 part in  $10^7$ , three orders of magnitude less than 0.01% error. This is one of the toughest examples we could think of, and even so the op-amp input impedance presents no problem, thus demonstrating that, in general, you can ignore the effects of op-amp input impedances.

Drifts in offset voltage in both  $U_2$  and  $U_3$  over time, temperature, and power-supply variations affect the final error equally and are tabulated in the budget. It is worth pointing out that they are all automatically cancelled at each “zeroing” cycle, and only short-term drifts matter anyway. These errors are all in the microvolt range, thanks to a good choice of op-amp.  $U_4$  has larger drifts, but it must be a FET type to keep capacitor current small, as already explained. Note that errors in  $U_4$ 's output are amplified by the gain setting of  $U_3$ ; thus they are specified as *input* errors in the budget.

Note the general philosophy of design that emerges from this example: you work at the problem areas, choosing configurations and components as necessary to reduce errors to acceptable values. Tradeoffs and compromises are involved, with some choices depending on external factors (e.g., the use of a FET-input op-amp for  $U_2$  would be preferable for source impedances greater than about  $10 \text{ k}\Omega$ ).

## 5.8 Amplifier output errors

As we discussed in Chapter 4, op-amps have some serious limitations associated with the output stage. Limited slew rate, output crossover distortion (§2.4.1A), and finite open-loop output impedance can all cause trouble, and they can cause precision circuits to display astoundingly large errors if not taken into account.

### 5.8.1 Slew rate: general considerations

As we mentioned in §4.4.1K, an op-amp can swing its output voltage only at some maximum rate. This effect originates in the frequency-compensation circuitry of the op-amp, as we will soon explain in a bit more detail. One consequence of a finite slew rate is to limit the output swing at high frequencies to a maximum of  $V_{\text{pp}} = S/\pi f$ , as shown in §4.4.2 and plotted here in Figure 5.11.

A second consequence is best explained with the help of a graph of slew rate versus differential-input signal (Figure 5.12). The point to be made here is that a circuit that demands a substantial slew rate must operate with a substantial voltage error across the op-amp's input terminals. This can be disastrous for a circuit that pretends to be highly precise: the feedback loop is in error, more so as the output slews more rapidly, thus producing a distorted output waveform. (Look ahead at the measured distortion plots in Figure 5.19 on page 311 to see this effect, for example in the LT1013 for which  $S=0.8 \text{ V}/\mu\text{s}$ .)

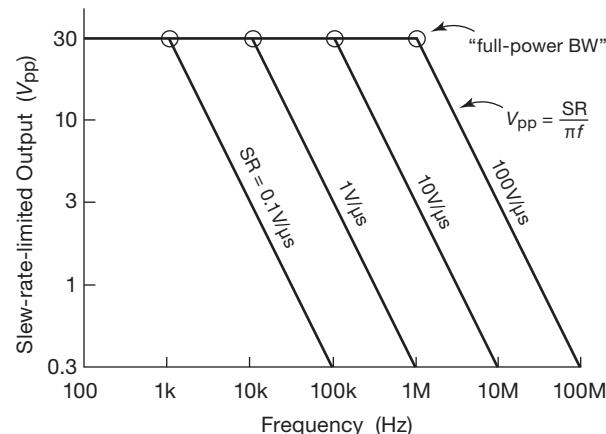
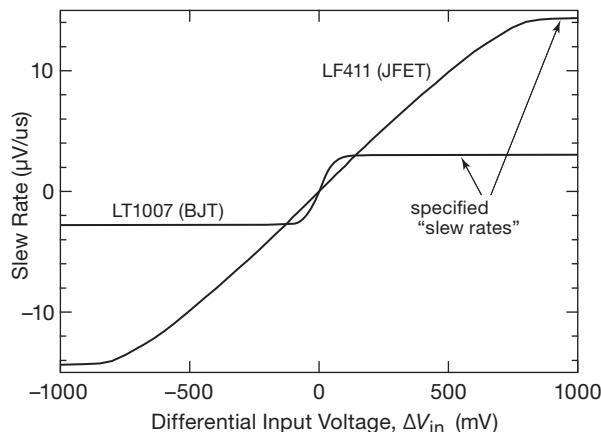


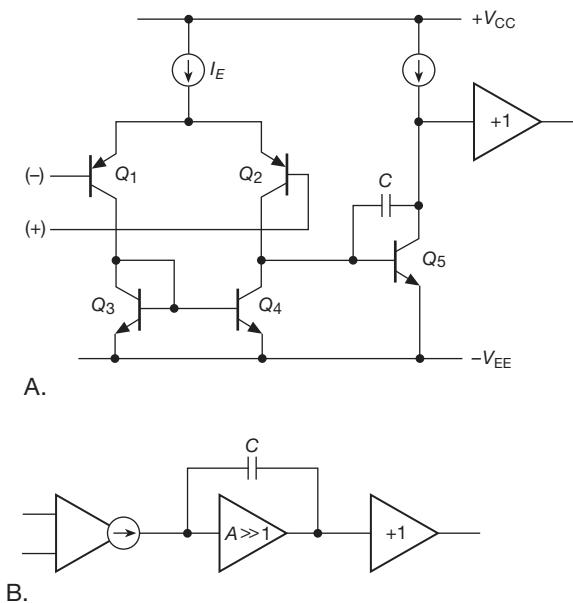
Figure 5.11. Maximum output swing versus frequency.

Let's look at the innards of an op-amp in order to get some understanding of the origin of slew rate (see §4x.9 for a more extensive discussion). The vast majority of op-amps can be summarized with the notional “Widlar circuit” shown in Figure 5.13. A differential input stage,<sup>18</sup> loaded with a current mirror, drives a stage of large voltage gain with a compensation capacitor from output to input. The output stage is a unity-gain push-pull follower. The compensation capacitor  $C$  is chosen to bring the open-loop gain of the amplifier down to unity before the phase shifts

<sup>18</sup> We've simplified it slightly: the input stage of Widlar's original LM101 used a *pnp* differential pair, but it was configured as a common-base amplifier driven by an *npn* follower pair.



**Figure 5.12.** A substantial differential-input voltage is required to produce the full op-amp slew rate, as shown in these measured data. For BJT-input op-amps it takes  $\sim 60$  mV to reach full slew rate; for JFETs and MOSFETs it's more like a volt.



**Figure 5.13.** Typical op-amp internal compensation scheme.

caused by the other amplifier stages have become significant. That is,  $C$  is chosen to put  $f_T$ , the unity-gain bandwidth, near the frequency of the next amplifier rolloff pole, as described in §4.9. The input stage has very high output impedance, and it looks like a current source to the next stage.

The op-amp is slew-rate limited when the input signal drives one of the differential-stage transistors nearly to cutoff, driving the second stage with the total emitter current

$I_E$  of the differential pair. For a BJT input stage this occurs with a differential input voltage of about 60 mV, at which point the ratio of currents in the differential stage is 10:1. At this point  $Q_5$  is slewing its collector as rapidly as possible, with all of  $I_E$  going into charging  $C$ . The transistor  $Q_5$  and  $C$  thus form an integrator, with a slew-rate-limited ramp as output. It's not hard to derive an expression for the slew rate, knowing how bipolar transistors work – see the discussion in §4x.9. The bottom line is that the classic BJT-input op-amp circuit of Figure 5.13 has a slew rate  $S$  given by  $S \approx 0.3f_T$ .

To get a higher slew rate, then, you can choose an op-amp with greater bandwidth  $f_T$ ; if you are operating at closed-loop gains greater than unity, you can use a compensated op-amp (with its higher  $f_T$  value). But there are ways (as explained in §4x.9) to beat the limit  $S \approx 0.3f_T$  (which assumed a unity-gain-compensated op-amp with a BJT differential input configured for maximum gain, i.e., with  $R_E=0$ ). Namely: (a) use an op-amp with reduced input-stage transconductance (either a FET-input op-amp or a BJT-input op-amp with emitter degeneration); (b) use an op-amp with a different input-stage circuit, specifically designed for enhanced slew rate – examples are the “cross-coupled transconductance reduction” technique (used in the TLE2142 family; see the cross-coupled input stage circuit shown in §4x.9), and the Butler “wide dynamic range transconductance stage” (used for example in the OP275 and OP285; see the Butler input stage circuit in §4x.9); (c) use a current-feedback (CFB) op-amp, or a CFB variant (with a buffered inverting input) that mimics an ordinary voltage-feedback (VFB) op-amp.

These tricks work. If we define an enhancement factor  $m$  (i.e.,  $S=0.3mf_T$ ), the LF411 (with JFET input) plotted in Figure 5.12 has  $m = 12$ , compared with the bipolar LT1007 ( $m=1.0$ ); the TLE2141 (with cross-coupled BJT input stage) has  $m = 25$ , and the OP275/285 (with Butler input stages) have  $m = 8$ ; the LT1210 (a CFB op-amp) has  $m = 55$  with the recommended feedback resistor; and the LT1351 (a CFB in VFB’s clothing) has  $m = 220$ .

For a deeper look at slew rate, turn to the extended discussion in Chapter 4x (§4x.9).

## 5.8.2 Bandwidth and settling time

Slew rate measures how rapidly the output voltage can change. The op-amp slew-rate specification usually assumes a large differential input voltage (60 mV or more), which (in spite of its potential for creating output distortion) is not unreasonable, given that an op-amp whose output isn't where it's supposed to be will have its input driven

hard by feedback, assuming a reasonable amount of loop gain. Of perhaps equal importance in high-speed precision applications is the time required for the output to get where it's going following an input change. This *settling-time* specification (the time required to get within the specified accuracy of the final value and stay there; see Figure 5.14) is always given for devices such as digital-to-analog converters, where precision is the name of the game, but it is not normally specified for op-amps.

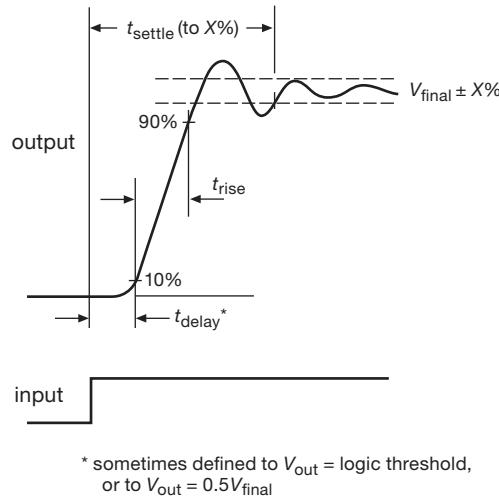


Figure 5.14. Settling time defined.

We can estimate op-amp settling time by considering first a different problem, namely, what would happen to a perfect voltage step somewhere in a circuit if it were followed by a simple  $RC$  lowpass filter (Figure 5.15). It is a simple exercise to show that the filtered waveform has the settling times shown. This is a useful result, because you often limit bandwidth with a filter to reduce noise (more on that later in the chapter). To extend this simple result to an op-amp, just remember that a compensated op-amp has a 6 dB/octave rolloff over most of its frequency range, just like a lowpass filter. When connected for closed-loop gain  $G_{CL}$ , its “bandwidth” (the frequency at which the loop gain drops to unity) is approximately given by

$$f_{3dB} = f_T/G_{CL}.$$

As a general result, a system of bandwidth  $B$  has response time  $\tau \approx 1/(2\pi B)$ ; thus the equivalent “time constant” of the op-amp is

$$\tau \approx G_{CL}/2\pi f_T.$$

The settling time is then roughly  $5-10\tau$ .

Let's try our prediction on a real case. The TLE2141

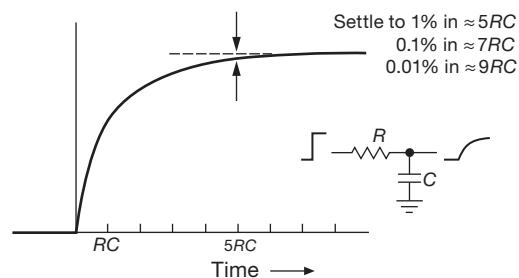


Figure 5.15. Settling time of an  $RC$  lowpass filter.

from TI is a precision fast-settling op-amp, with an  $f_T$  of 5.9 MHz. Our simple formula then estimates the inverting-configuration (i.e.,  $G=2$ ) response time to be 54 ns, thus a settling time of 378 ns ( $7\tau$ ) to 0.1%. This is in good agreement with the datasheet's value of 340 ns.

There are several points worth making: (a) our simple model gives us only a lower bound for the actual settling time in a real circuit; you should always check the slew-rate-limited rise time, which may dominate. (b) Even if slew rate is not a problem, the settling time may be much longer than our idealized “single-pole” model, depending on the op-amp's compensation and phase margin. (c) The op-amp will settle more quickly if the frequency compensation scheme used gives a plot of open-loop phase shift versus frequency that is a nice straight line on a log-log graph (as in Figure 5.17); op-amps with wiggles in the phase-shift graph are more likely to exhibit overshoot and ringing, as in the upper waveform shown in Figure 5.14. (d) A fast settling time to 1%, say, doesn't necessarily guarantee a fast settling time to 0.01%, since there may be a long tail (Figure 5.16). (e) There's no substitute for an actual settling-time specification from the manufacturer.

Table 5.4 lists a selection of high-speed op-amps suitable for applications that demand high  $f_T$ , high slew rate, fast settling time, and reasonably low offset voltage.

### 5.8.3 Crossover distortion and output impedance

Some op-amps (for example the classic single-supply LM324/358) use a simple push-pull follower output stage, without biasing the bases two diode drops apart, as we discussed in §2.4.1. This leads to “class-B” distortion near zero output, because the driver stage has to slew the bases through  $2V_{BE}$  as the output current passes through zero (Figure 5.18). This crossover distortion can be substantial, particularly at higher frequencies where the loop gain is reduced; see the measured data in Figure 5.19. It is greatly reduced in op-amp designs that bias the output

**Table 5.4 Representative High-speed Op-amps<sup>x</sup>**

Part #	# per pkg <sup>a</sup>	Supply <sup>b</sup>		$I_{in}$ @25°C		Offset Voltage			$e_n$ typ <sup>r</sup>	GBW (MHz)	Slew typ	$I_{out}$ typ	$C_{in}$ (pF)	dist. graph	Swing to Supply		null pins	DIP avail	cost qty 25 (\$US)	Comments	
		range (V)	$I_Q$ (mA)	$V_{os}$ typ (mV)	$\Delta V_{os}$ typ ( $\mu\text{V}/^\circ\text{C}$ )	$V_{os}$ max (mV)	$\Delta V_{os}$ typ ( $\mu\text{V}/^\circ\text{C}$ )	$e_n$ typ <sup>r</sup> ( $\text{nV}/\sqrt{\text{Hz}}$ )							$IN$	$OUT$					
<i>bipolar</i>																					
LT1468	1	7–36	3.9	3nA	0.03	0.08	0.7	5	90	23	22	4	-	-	-	-	•	•	4.26	0.7 ppm dist	
LT1360	1,2,4	5–36	4	0.3 $\mu\text{A}$	0.3	1	9	9	50	800	34	3	-	-	-	-	•	•	2.75	C-Load™	
LM6171	1,2	5–36	2.5	1 $\mu\text{A}$	1.5	3	6	12	100	3600	90	-	•	-	-	-	•	•	2.57	VFB+CFB	
AD844	1	9–36	6.5	0.2 $\mu\text{A}$	0.05	0.3	1	9	330 <sup>g</sup>	2000	60	2	-	-	-	-	•	•	5.23	CFB, comp pin	
AD8021 <sup>b</sup>	1	4.5–26	7	7.5 $\mu\text{A}$	0.4	1	0.5	2.1	925	420	60	1	•	-	-	-	-	-	2.42	comp pin, 16-bits	
<i>JFET</i>																					
OPA604A	1,2	9–50	5.3	50	1	5	8	10	20	25	36	10	-	-	-	-	•	•	2.93	3ppm, dual '2604	
OPA827A	1	8–40	4.8	15	0.08	0.15	1.5	3.8	22	28	30	9	-	-	-	-	-	-	9.00	quiet, accurate	
ADA4637	1	9–36	7.0	1	0.12	0.3	1	6.1	80	170	45	8	-	-	-	-	d	10.12	decomp, G>7		
<i>low-voltage bipolar</i>																					
LT6220	1,2,4	2.2–13	0.9	15nA	0.07	0.35	1.5	10	60	20	35	2	-	•	•	•	-	-	1.75	SOT-23	
LMH6723	1,2,4	4.5–13	1	2 $\mu\text{A}$	1	3	-	4.3	370	600	110	1.5	-	-	-	-	-	-	2.03	CFB, SOT23-5	
ADA4851	1,2,4	3–12.6	2.5	2.2 $\mu\text{A}$	0.6	3.4	4	10	125	200	85	1.2	-	-	•	•	-	-	1.40	SOT23-5, shdn pin	
LT1818	1,2	4–12.6	9	2 $\mu\text{A}$	0.2	1.5	10	6	400	2500	70	2	-	-	-	-	-	-	1.35	VFB+CFB, fast	
LT6200	1,2	3–12.6	16.5	10 $\mu\text{A}$	0.2	1.2	8	0.95	165	50	70	4	-	•	•	•	-	-	2.99	1% dist at 50MHz	
LT6200-10	1	3–12.6	16.5	10 $\mu\text{A}$	0.2	1.2	8	0.95	1600	450	70	4	-	•	•	•	-	-	2.99	fastest RRIO	
OPA698 <sup>e</sup>	1	5–13	16	3 $\mu\text{A}$	2	5	15 <sup>m</sup>	5.6	450	1100	55	1	n	-	-	•	-	-	4.14	clipping	
<i>low-voltage JFET</i>																					
OPA656	1	9–13	14	2	0.25	1.8	2	7	230	290	50	2.8	•	-	-	-	-	-	5.59	low $e_n \cdot C_{in}$ noise	
OPA657	1	9–13	14	2	0.25	1.8	2	7	1600	700	50	4.5	-	-	-	-	-	-	10.01	decomp, G>7	
ADA4817	1,2	5–10.6	19	2	0.4	2	7	4	1050	870	70	1.5	-	•	-	-	-	-	4.93	lowest $e_n \cdot C_{in}$	
<i>CMOS</i>																					
AD8616	2	2.7–6	1.7	0.2	0.02	0.06	1.5	7	24	12	150	7	-	•	•	•	-	-	1.52	'8615 SOT23-5	
LMP7717	1,2	1.8–6	1.15	0.05	0.01	0.15	1	6.2	88	28	15	15 <sup>c</sup>	-	•	•	•	-	-	2.18	decomp, G>10	
OPA350	1,2,4	2.5–7	5.2	0.5	0.15	0.5	4	7	38	22	40	6.5	-	•	•	•	-	•	1.67	6ppm	

Notes: (a) boldface indicates number in a package for the part number listed. (b) for G<20 use ext Cc chosen to set  $f_{3dB}=200\text{MHz}$ . (c) 15mA sinking, 47mA sourcing. (d) for DIP-8 see OPA637. (e) OPA699 decomp. (g) at G=10. (m) max. (n) distortion plot for OPA699. (p)  $I_Q$ , typical, per amplifier. (r) at 1kHz. (x) see also the fast op-amp table in Chapter 4x.

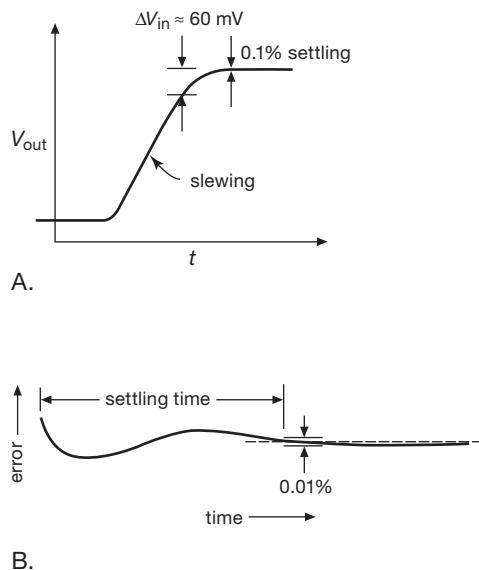
push-pull pair into slight conduction (“class-AB”), for example the LT1013, which is an improved version of the LM324. The right choice of op-amp can have enormous impact on the performance of low-distortion audio amplifiers. Perhaps this problem has contributed to what the audiophiles refer to as “transistor sound.” Some modern op-amps, particularly those intended for audio applications, are designed to produce extremely low crossover distortion. Examples are the LT1028, the AD797, and the excellent “LME49000” series from NSC, e.g., the LME49710. The latter, for example, has less than 0.0001% distortion over the full audio band of 20–20 kHz. (That’s the claim, anyway; we may be overly gullible!) These amplifiers all have very low noise voltage, as well; the LT1028, for example, vies for the title of world noise-voltage champion, with  $e_n = 1.7 \text{ nV}/\sqrt{\text{Hz}}$  (max) at 10 Hz. See the expanded plots of op-amp distortion in Figures 5.43 and 5.44, where various op-amps compete for the title of king of low distortion. High-voltage op-amps have an advantage below

10 kHz, whereas the low-voltage types have an advantage above 200 kHz.

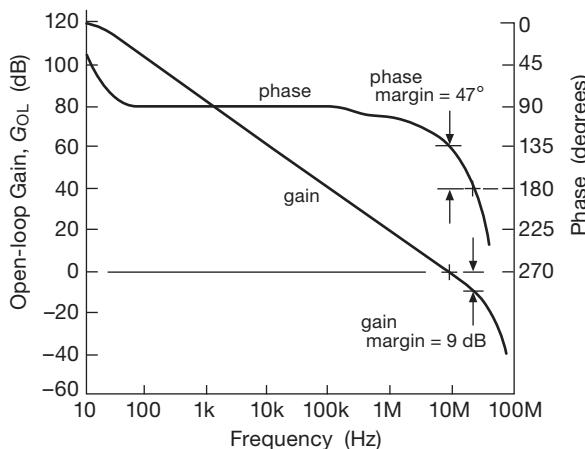
The open-loop output impedance of a typical split-supply op-amp is highest when the output is near ground, because the output transistors are operating at their lowest current into the (ground-retumed) load. The output impedance also rises at high frequency as the transistor gain drops off, and it may rise slightly at very low frequencies because of thermal feedback on the chip.

It is easy to neglect the effects of finite open-loop output impedance, thinking that feedback will cure everything. But when you consider that some op-amps have open-loop output impedances of a few hundred ohms, it becomes clear that the effects may not be negligible, especially at low to moderate loop gains. Figures 5.20 and 5.21 shows some typical graphs of op-amp output impedance, both with and without feedback.

The finite output impedance contributes also to instability when driving capacitive loads, as we discussed in

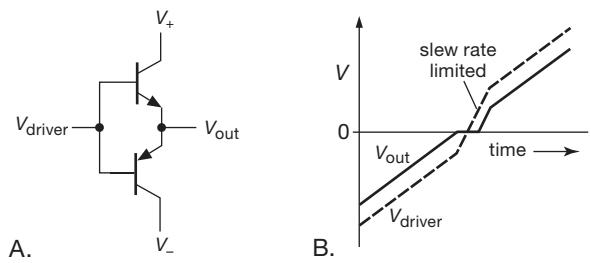


**Figure 5.16.** A. The slewing decreases when input error approaches 60 mV. B. Settling to high precision can be surprisingly lengthy.

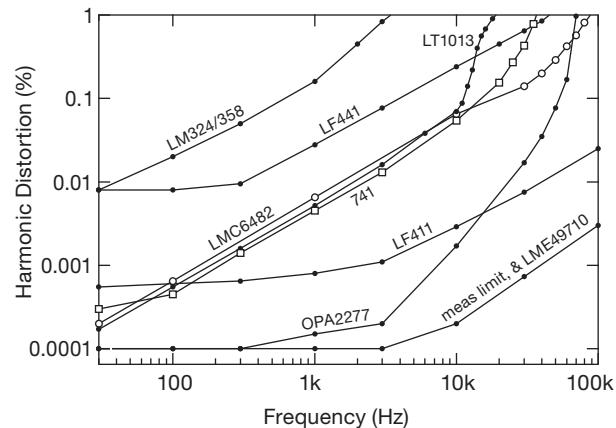


**Figure 5.17.** OP-42 gain and phase versus frequency.

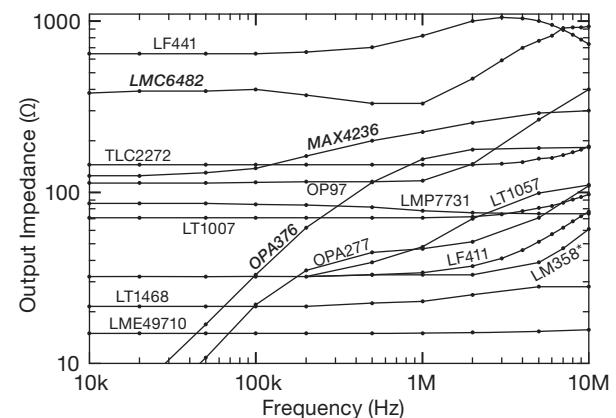
§4.6.2, owing to the additional lagging phase shift within the feedback loop that is created by  $R_{out}$  in combination with  $C_{load}$ . Several common solutions were shown in Figure 4.78, including a split feedback path or the inclusion of a unity-gain buffer within the loop. The latter is worth a mention here.



**Figure 5.18.** Crossover distortion in class-B push-pull output stage.



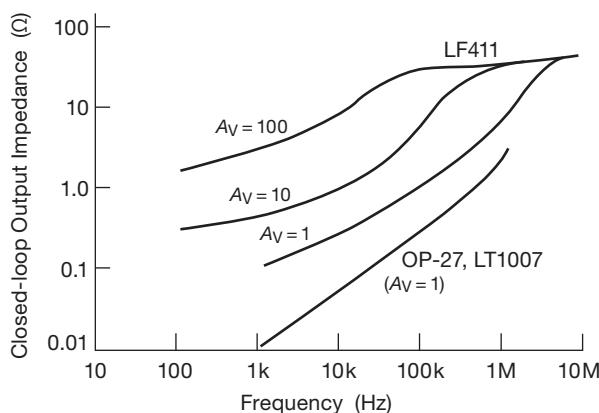
**Figure 5.19.** Measured harmonic distortion versus frequency for several popular op-amps (1 Vrms output, unloaded). See also Figures 5.43 and 5.44.



**Figure 5.20.** Measured open-loop output impedance versus frequency for a selection of op-amps. Parts shown in **bold** have CMOS output circuits. \* With output pulldown resistor.

#### 5.8.4 Unity-gain power buffers

If the technique of split feedback paths is unacceptable, one solution is to add a unity-gain high-current buffer inside



**Figure 5.21.** Closed-loop output impedance versus frequency for the LF411 and LT1007 op-amps, from manufacturers' datasheets.

the loop, as, for example, in the general-purpose laboratory amplifier of Figure 4.87. The LT1010 in that circuit has adequate bandwidth ( $>10\text{ MHz}$ ) in which it adds little phase shift; thus it can be within the feedback loop with a small amount of external compensation.

These “power boosters” can, of course, be used for loads that require high current (for example, driving a terminated coax cable), regardless of whether or not there are problems with capacitance. And unity-gain buffers are useful even with loads of only moderate current, in the context of precision circuit design, because they prevent thermal drifts by keeping the heat out of the low-offset amplifier. You can see a couple of examples of power boosters in Figures 5.47 and 13.119, as well as in the discussion in Chapter 4x.

### 5.8.5 Gain error

There's one more error that arises from finite open-loop gain, namely, an error in closed-loop gain owing to finite loop gain.

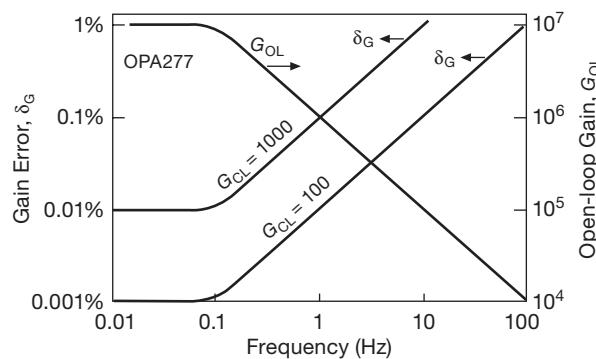
We calculated in Chapter 2 (§2.5.2) the expression for closed-loop gain in a feedback amplifier,  $G = A/(1+AB)$ , where  $A$  is the open-loop gain and  $B$  is the “gain” of the feedback network. You might think that the  $A \geq 100\text{ dB}$  or so of op-amp open-loop gain is plenty, but when you try to construct extremely precise circuits you are in for a surprise. From the preceding gain equation it is easy to show that the “gain error,” defined as

$$\delta_G = \text{gain error} \equiv \frac{G_{\text{ideal}} - G_{\text{actual}}}{G_{\text{ideal}}},$$

is just equal to  $1/(1+AB)$  and ranges from 0 for  $A = \infty$  to 1 (100%) for  $A = 0$ .

**Exercise 5.2.** Derive the foregoing expression for gain error.

The resulting frequency-dependent gain error is far from negligible. For instance, an LF411 with its 106 dB of low-frequency open-loop gain will have a gain error of 0.5% at low frequencies when configured for a closed-loop gain of 1000. Worse yet, the open-loop gain drops 6 dB/octave above 20 Hz, so our amplifier would have a gain error of a whopping 10% at 500 Hz! Figure 5.22 plots calculated gain error versus frequency for the OPA277, with its extraordinary 140 dB of low-frequency open-loop gain, when configured for closed-loop gains of 100 and 1000. It should be obvious that you need plenty of gain and a high  $f_T$  to maintain accuracy at even moderate frequencies.



**Figure 5.22.** OPA277 gain error.

We plotted these curves from the graph of open-loop gain versus frequency given in the datasheet. Even if your op-amp datasheet provides a curve, it's best to work backward from the specified  $f_T$  (i.e., the datasheet's GBW; see Figure 5.42 and associated discussion) and dc open-loop gain, figuring the open-loop gain at the frequency of interest, and thus the gain error (as above) as a function of frequency. This procedure yields

$$\delta_G = \frac{1}{1 - jBf_T/f} \approx \frac{f}{Bf_T},$$

where  $B$  is, as usual, the gain of the feedback network, and the approximation is valid for the useful case  $Bf_T/f \gg 1$ . Of course, in some applications, such as filters,  $B$  may also depend on frequency.

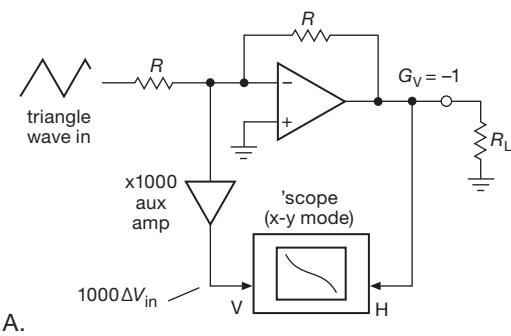
**Exercise 5.3.** Derive the foregoing result for  $\delta_G(f)$ .

### 5.8.6 Gain nonlinearity

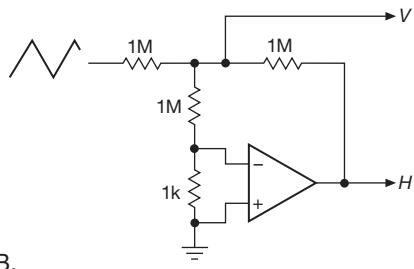
Op-amps have lots of open-loop gain at low frequencies, and the excess ( $G_{OL}/G_{CL}$ ) is the loop-gain feedback mechanism that contributes to accuracy and the reduction of the op-amp's intrinsic nonlinearities, as discussed first

in §2.5.3. Ideally, then, we want lots of open-loop gain in a precision circuit. And that's why auto-zero amplifiers (§5.11) and precision op-amps are built with high open-loop gains, for example  $\sim 160$  dB for the auto-zero LMP2021, and  $\sim 150$  dB for the precision LT1007.

For *accuracy*, then, we want lots of loop gain. For purposes of *linearity*, however, it's OK to have less loop gain – what matters more is the intrinsic linearity of the op-amp, combined with an open-loop gain characteristic that changes linearly (if at all) with output swing. The intrinsic linearity is strongly influenced by the output-stage design, particularly when the amplifier is driving a load: crossover distortion is always bad, as is an output stage that is asymmetric in its source/sink capabilities (like the LM358, with a Darlington *npn* pullup and single *pnp* pull-down). And a poor layout within the chip can create nonlinearities from the thermal offsets produced by local heating when driving a load.



A.

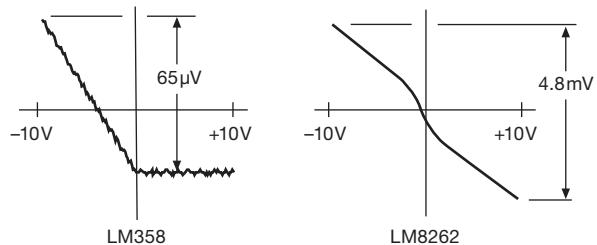


B.

**Figure 5.23.** Low-frequency gain nonlinearity test circuit. A. Notional: an auxiliary amplifier makes visible the  $\mu$ V-scale differential input voltage versus output swing. B. Circuit used by Pease for measurements in AN-1485 (see footnote on the current page).

In a nice set of measurements, Bob Pease<sup>19</sup> explored the low-frequency gain nonlinearity of a selection of op-amp

types (sadly, none from other manufacturers), operating as unity-gain inverters with a full-swing output; he made measurements both when the op-amps were unloaded, and when they were driving a  $1\text{k}\Omega$  load. The basic scheme is shown in Figure 5.23A, where a 'scope looks at the amplifier input error versus output swing. For Pease's actual measurements he used the subtle variant in Figure 5.23B, in which the op-amp amplifies its error by  $\times 1000$ , delivering the bad news directly.



**Figure 5.24.** Gain nonlinearity traces for two op-amps with output-stage deficiencies. In these  $x$ - $y$  displays the vertical axis shows the (small) differential input signal required to produce the (full-swing) output signal indicated on the horizontal axis. To estimate gain error, divide the vertical deviation from a best-fit straight line by the full-swing output.

The sorts of things you see (with a loaded op-amp) are shown in Figure 5.24, where we've sketched Pease's traces for the aforementioned LM358 (afflicted with asymmetric source/sink) and the LM8262 (fast, but afflicted with some crossover distortion). An exemplary op-amp like the very low distortion LM4562 presents an ideal nearly-horizontal straight line. The LF411 (single) and LF412 (dual), our JFET jellybeans, show an interesting contrast: according to Pease, the LF411 chip layout is sub-optimal (in terms of gain and thermal effects), with great effort rewarded by better results in the dual LF412.

Here are some of his summary results for op-amps driving a somewhat lighter load ( $4\text{k}\Omega$ ). In general, the measured gain nonlinearity, when *unloaded*, was far smaller than these listed values. Keep in mind that these measurements were made at very low frequencies (generally just a few hertz), where the loop gain is maximum.

<sup>19</sup> National Semiconductor App Note AN-1485: *The Effect of Heavy Loads on the Accuracy and Linearity of Operational Amplifier Circuits*

(or, "What's All this Output Impedance Stuff, Anyhow?"). Gain nonlinearity data can be found in some datasheets, for example the AD620 instrumentation amplifier.

HV BJT ( $V_{sig} = \pm 10$ V)		
LM8262	12 ppm	xover dist.
LM358	1 ppm	asym. output stage
LF411	1.4 ppm	poor layout – thermal
LF412	0.3 ppm	better layout
LM4562	0.025 ppm	pro-audio, $G_{OL} = 10^7$

CMOS RRO ( $V_{sig} = \pm 4$ V)		
LMC6482	1.1 ppm	jellybean
LMC6062	0.2 ppm	precision

CMOS auto-zero ( $V_{sig} = \pm 2$ V)		
LMP2012	0.2 ppm	precision

### 5.8.7 Phase error and “active compensation”

We've talked mostly about the *gain* error caused by limited op-amp bandwidth (and therefore falling loop gain with increasing frequency). But limited loop gain also produces *phase* error, which can be important in applications such as video, interferometry, and so on. And the effect is not at all negligible – recall (§1.7.9) that a single *RC*-like rolloff creates a phase shift of  $\sim 6^\circ$  at a frequency of  $f_C/10$ , and  $\sim 0.6^\circ$  at  $f_C/100$ ; the latter is two full decades below the  $-3$  dB breakpoint. If we model an op-amp's rolloff of open-loop gain similarly (a “single-pole” rolloff), we can expect comparable phase shifts.

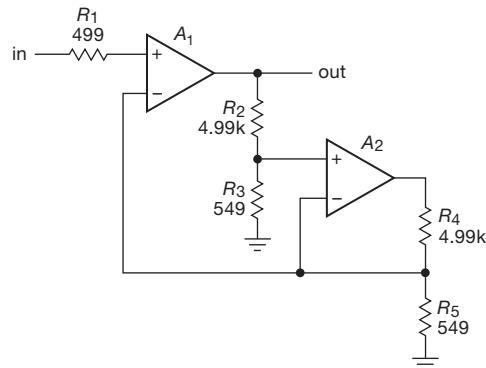
In this approximation, the resultant phase shift for an op-amp voltage amplifier is thus given by

$$\phi = \tan^{-1} \left( \frac{f}{f_C} \right) \approx \frac{f}{f_C} \text{ (radians)},$$

where the  $-3$  dB breakpoint  $f_C$  is the frequency at which the loop gain has fallen to unity:  $f_C = f_T/G_{CL}$ . Here  $G_{CL}$  is the closed-loop gain (as set by the feedback network), and  $f_T$  is the gain-bandwidth product (GBW) of the op-amp (for a single-pole rolloff that's the same as the frequency at which the open-loop gain is unity; but for typical op-amps, with more complicated rolloffs, you want to use the GBW figure). Multiply by  $57.3$  ( $180/\pi$ ) to get the answer in degrees. The approximate result (the last expression) is reasonably accurate for small-to-moderate phase shifts, up to 0.5 radian, say.

There are several ways to address this problem. The simplest is to use an amplifier of greater bandwidth. If you don't want to (or cannot) do that, another possibility is to introduce an *RC* network in the feedback path to cancel the phase error (in *s*-plane language, you are introducing a zero to cancel a pole). This can be effective, but requires

“tuning” of the compensation network to match the frequency response of the particular op-amp specimen itself; and because the op-amp's characteristics change with temperature, the network must do likewise. A third possibility is to cascade two stages, each configured for lower gain (and therefore smaller phase error).



**Figure 5.25.** Phase-error reduction by means of “active compensation,” exploiting the closely matched frequency responses of dual op-amp pair  $A_1$  and  $A_2$ .

But an elegant solution is *active compensation*, a clever technique that uses a second matching op-amp to create a replica of the error, which can then be subtracted from the main amplifier. Figure 5.25 shows how this can be done.<sup>20</sup> The bandwidth of the main amplifier is unchanged, but its phase error is reduced dramatically, as shown in the SPICE simulation and measured data of Figure 5.26. There is some peaking in the amplitude response – about +3 dB at the frequency at which the phase error is  $45^\circ$  – but generally insignificant within the frequency range over which the phase error is small (e.g., +0.1 dB at  $f=0.1f_T/G_{CL}$ ). A circuit configured for low closed-loop gain will generally exhibit greater peaking.<sup>21</sup> Under the assumption that the

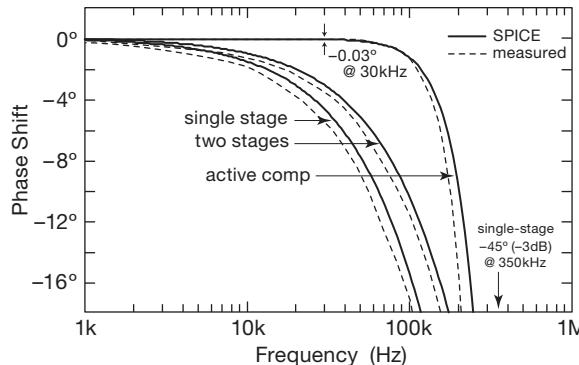
<sup>20</sup> See “Active Feedback Improves Amplifier Phase Accuracy,” by J. Wong, *EDN Magazine*, 17 Sept 1987; reprinted as Analog Devices AN-107. Wong credits the idea to Soliman in a 1979 paper, and Soliman credits the idea to Brackett and Sedra in a 1976 paper. But Wong's paper is the most useful reference for understanding the configuration of Figure 5.25.

<sup>21</sup> In SPICE simulations we found that the peaking increased to  $\sim 7$  dB for the LF412 model configured for  $G=2$ ; this can be tamed by adding a compensation capacitor  $C_c$  across feedback resistor  $R_2$ . Choosing  $C_c$  to match the op-amp's  $f_T$  (i.e.,  $C_c = 1/2\pi f_T R_2$ ) reduced the peaking to 4 dB, at the expense of tripling the (pretty small) phase error. In his article, James Wong warns that the technique may result in an unstable amplifier for low gains, below  $G=5$  for example. He shows also how the technique can be further improved if  $A_2$  is made from two amplifiers.

amplifiers are matched, it can be shown that this technique produces a phase shift given approximately by

$$\phi \approx \left( \frac{f}{f_C} \right)^3 \text{ (radians)},$$

again accurate for small-to-moderate phase shifts ( $\lesssim 30^\circ$ , say, i.e., the small-angle approximation).

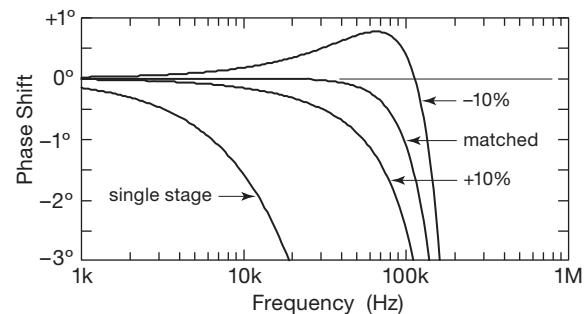


**Figure 5.26.** SPICE simulation and measured data of the phase shift versus frequency for the circuit of Figure 5.25, implemented with an LF412 dual JFET op-amp. For comparison the analogous data are plotted for both a single  $G=10$  stage and for a cascade of two stages, each with  $G=\sqrt{10}$ . The measured part's  $f_C$  was 295 kHz, somewhat lower than the SPICE model's 350 kHz.

Real op-amps are not perfectly matched. To see how a mismatch in  $f_T$  affects the phase compensation, we ran a SPICE simulation with an  $f_T$  mismatch of  $\pm 10\%$  (Figure 5.27). Evidently our test-bench part, chosen quite at random (dip fingers into parts bin, grasp first part touched, extract, measure), has a considerably better  $f_T$  matching, as suggested by Wong: ‘‘Monolithically matched dual or quad op-amps can provide the frequency-matching characteristics (to within 1% to 2%) necessary for the success of the active-feedback approach.’’<sup>22</sup>

It’s interesting to compare predicted phase shifts for several scenarios mentioned at the outset: (a) a single amplifier of given bandwidth (call it  $f_{T0}$ , 3 MHz for the LF412), configured for a closed-loop gain  $G = 10$ ; (b) two cascaded stages, each with  $G = \sqrt{10}$ ; (c) the active compensation method of Figure 5.25; and (d) a single amplifier of greater bandwidth ( $10f_{T0}$ , say). Here are the calculated results:

<sup>22</sup> We went back to the bench and measured a handful of LF412 dual op-amps. Among different specimens the  $f_T$  values ranged over  $\pm 20\%$ , but within any single part the  $f_T$ ’s of its two op-amps matched typically to 0.1%, with one outlier showing a 1.5% mismatch.



**Figure 5.27.** Active compensation of phase error requires matched op-amp bandwidths, as seen in this SPICE simulation for which the  $f_T$  of the compensation op-amp  $A_2$  has been varied  $\pm 10\%$  relative to that of the signal path op-amp  $A_1$ .

	Single, $G=10$	2-stage, each $G=\sqrt{10}$	Active, comp	Single, $f_T=10f_{T0}$
$0.001f_{T0}$	$-0.57^\circ$	$-0.36^\circ$	$-0.00006^\circ$	$-0.006^\circ$
$0.003f_{T0}$	$-1.7^\circ$	$-1.1^\circ$	$-0.0015^\circ$	$-0.17^\circ$
$0.01f_{T0}$	$-5.7^\circ$	$-3.6^\circ$	$-0.06^\circ$	$-0.57^\circ$
$0.03f_{T0}$	$-17.2^\circ$	$-10.9^\circ$	$-1.5^\circ$	$-1.7^\circ$
$0.1f_{T0}$	$-45^\circ$	$-36^\circ$	$-45^\circ$	$-5.7^\circ$

It’s clear that the remarkable (and underutilized) technique of active compensation represents an efficient use of resources. The noninverting  $G = 2$  case looks especially useful, e.g., to drive backterminated 75  $\Omega$  video cables.<sup>23</sup>

## 5.9 RRIO op-amps: the good, the bad, and the ugly

In Chapter 4 (§§4.4.1, 4.4.2, and 4.6.3) we introduced rail-to-rail op-amps, including (a) op-amps that operate properly with common-mode inputs over the full supply voltage range (RRI), (b) op-amps that can swing their outputs over the full supply range (RRO), and (c) op-amps that can do both (RRIO). With lower supply voltages increasingly in vogue, you see many new op-amps with these desirable capabilities.

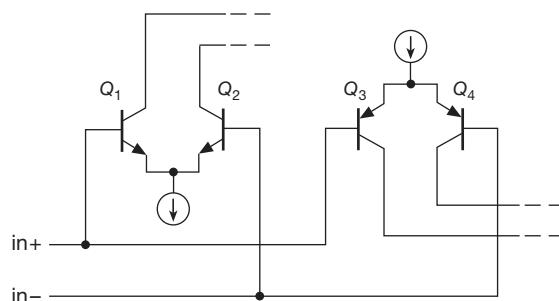
Desirable, but to be used with caution. These benefits come at a cost, which we’ll discuss here in the context of precision design (with further discussion in Chapter 4x). In circuits that strive for accuracy there are some hidden compromises in the designs of these op-amps about which the datasheet may be, uh, understated (or completely silent). Here are the important ones.

<sup>23</sup> Or sometimes called ‘‘double-terminated’’, as in Figure 12.110. We suggest trying your amplifier of choice, taking care to terminate the second op-amp with 150  $\Omega$ .

### 5.9.1 Input issues

#### A. Input-current crossover

Most RRIO op-amps use a complementary pair of differential input stages, with their inputs driven in parallel, to handle the full supply voltage range (Figure 5.28). This causes a shift in input current because the signal path changes from one pair to the other, as seen clearly in Figure 5.7 (particularly the BJT-input RRIO op-amps: LT1630, LM6132). An abrupt change in input current causes input errors from finite driving impedance. Some RRIO op-amps avoid this problem by using an on-chip charge pump to generate a supply voltage beyond the rail, so a single input amplifier allows rail-to-rail inputs. Examples are the OPA360-series,<sup>24</sup> the AD8505 and ADA4505, the MAX4162-series, and the MAX4126-series. Except for the BJT-input MAX4126, these all use MOS inputs.



**Figure 5.28.** A typical rail-to-rail input circuit consists of a pair of complementary differential amplifiers, with downstream circuitry to select the active pair's output.

In situations where you need RRO but don't need full rail-to-rail *input* (a voltage amplifier with  $G>2$ , say), be sure to consider an RRIO op-amp with input extending to the negative rail only (sometimes called "ground sensing"). Note also that, by using an op-amp in an inverting circuit configuration, you avoid this problem completely (but you probably would not choose an RRIO op-amp for such a configuration anyway).

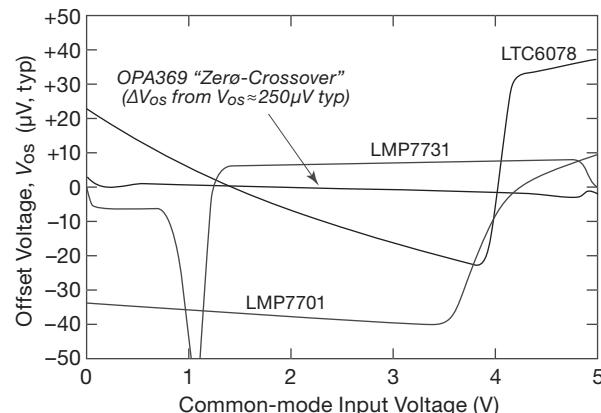
#### B. Input offset-voltage crossover

The dual input stages of RRIO op-amps cause similar mischief in terms of their input offset voltage  $V_{OS}$ , as seen in Figure 5.29. The abrupt change can occur close to either end of the supply range, as seen in the LMP7701 and LMP7731 op-amps from the same manufacturer. These curves were adapted from their respective datasheets,

which typically display a figure showing a tangle of overlapping curves measured on multiple op-amp samples (if they're willing to show any data at all about this seamy topic). Here you can see by comparison the uncomplicated (and downright *boring*) behavior of an RRIO op-amp with an on-chip charge pump powering a single input amplifier. This variation of  $V_{OS}$  with  $V_{CM}$  is not only undesirable, it is also unpredictable, as you can see in Figure 5.30.

This problem is nicely circumvented by the use of an inverting configuration, which holds constant the common-mode input voltage. More generally, always consider using an inverting configuration to prevent *any* circuit misbehavior caused by op-amp dependence on  $V_{CM}$ .<sup>25</sup>

The OPA350 datasheet shows a nice example (Figure 5.31) of input crossover effects in RRIO op-amps, namely a 17 dB increase in audio distortion in a  $G=1$  follower when the 3 Vpp sinewave input is shifted upward to enter the crossover region.<sup>26</sup> The same graph illustrates nicely how increased closed-loop gain causes increased distortion owing to decreased loop gain.



**Figure 5.29.** Op-amps with rail-to-rail inputs usually exhibit a shift of  $V_{OS}$  as the input voltage passes control from one input pair to the other. The OPA369 circumvents this by using a single input pair, powered beyond the rail by an on-chip charge pump.

### 5.9.2 Output issues

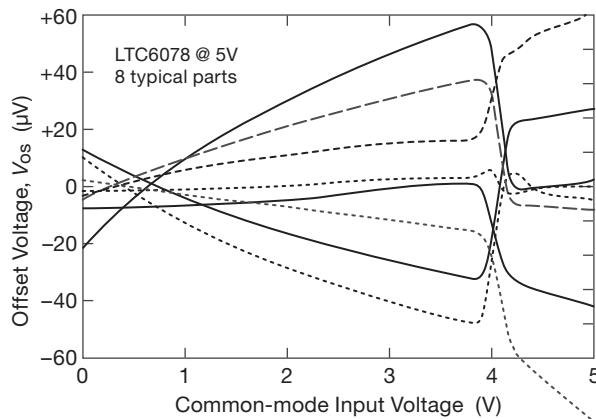
#### A. Output impedance

The output stage of a conventional (not RRO) op-amp is ordinarily a complementary push-pull follower (or some

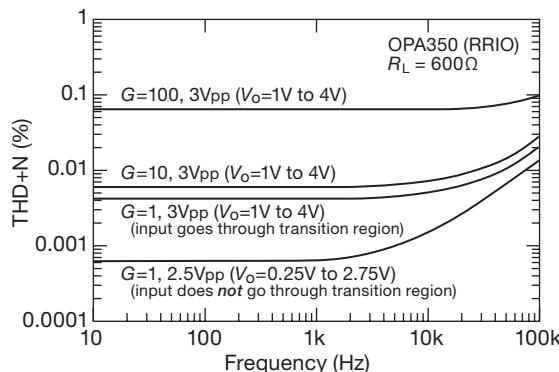
<sup>25</sup> As Jim Williams liked to say, "Use an inverting configuration, unless you can't."

<sup>26</sup> See also Bonnie Baker's article (in the *Baker's Best* series) "Where did all that racket come from?" in *EDN Magazine*, 23 April 2009, available at edn.com.

<sup>24</sup> Playfully named "Zerø-Crossover" amplifiers, or ZCOs.



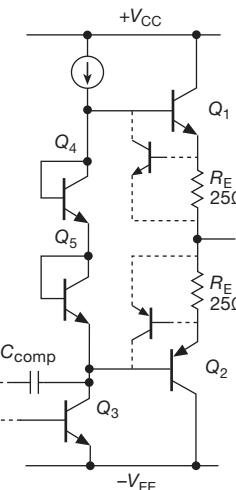
**Figure 5.30.** The shift of offset voltage in an RRI op-amp can be unpredictable (even as to the sign of the effect!), as seen in these data, adapted from the unusually forthcoming manufacturer's datasheet.



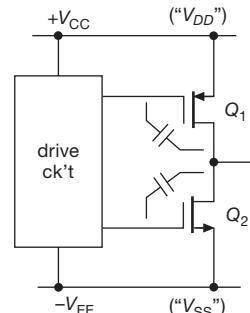
**Figure 5.31.** Distortion versus frequency for the OPA350 RRIO op-amp. The two lowest curves show the dramatic increase in distortion when the input signal enters the input crossover region. Increasing closed-loop gain causes further distortion because of reduced loop gain.

variation thereupon), biased with some conduction overlap to prevent crossover distortion at mid-supply (see §5.8.3). By contrast, the output complementary pair in an RRO op-amp is configured as a push–pull common-source *amplifier*; see Figure 5.32. That's necessary for the output to reach the rails (absent a second set of beyond-the-rails supply voltages). But it creates problems, owing to its inherently high output impedance.

The high  $Z_{out}$  means that the output-stage gain (and therefore the loop gain) depends on the value of load resistance; and a capacitive load creates large phase shifts, compromising the loop stability (see, for example, Figure 4.79). These problems are addressed in part by use of



A. Follower (not RRO)



B. Amplifier (RRO)

**Figure 5.32.** The classic (not rail-to-rail) op-amp output stage is a push–pull unity-gain follower with inherently low output impedance, biased (via  $Q_4Q_5$ ) to suppress crossover distortion; it has straightforward biasing and current limiting. By contrast, a rail-to-rail output stage (usually implemented in CMOS) is a push–pull common-source amplifier ( $G > 1$ ) with inherently high output impedance; it requires considerable trickery in its biasing and current limiting.

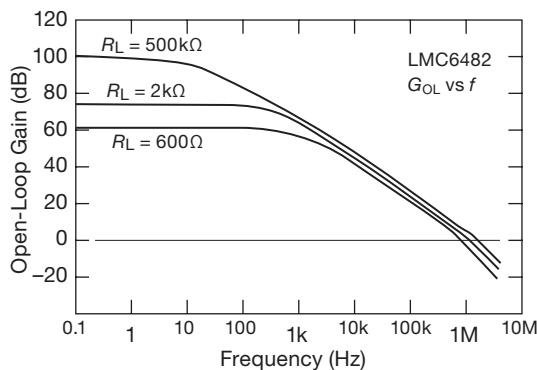
internal feedback around the output stage (the capacitors in Figure 5.32B), so that the gain and output impedances are reasonably well controlled except at low frequencies – see for example Figures 5.33 and 5.34.<sup>27</sup>

## B. Saturation at the rails

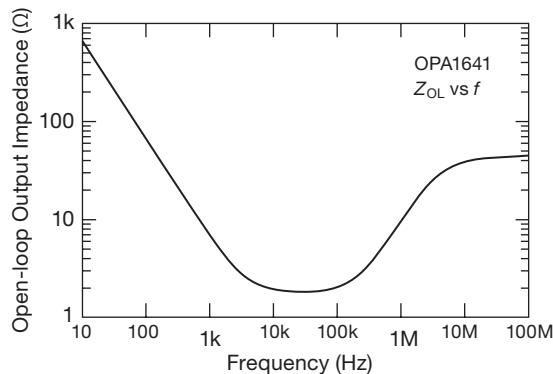
Some “rail-to-rail output” op-amps (in particular, those with a BJT output stage) don’t quite make it the last few millivolts; that’s because the output transistor’s saturation voltage is not zero. (This is not usually a problem with MOSFET outputs, which look like an  $R_{on}$  to one rail or the other when driven full range.) Usually this doesn’t matter, because what you care about most is getting full use of a limited supply voltage (when operating with low-voltage supplies). But it does matter, for example, if you’ve got a single-supply setup in which the op-amp is driving an ADC whose conversion range goes clear down to ground.

In such a case be sure to check the specifications. Some

<sup>27</sup> It’s unusual to see plots (or even tabulated values) of *open-loop* output impedance on datasheets; and in cases where a graph is shown, it rarely extends to very low frequencies. It is likely that other op-amps, including some with conventional (follower) output stages, also exhibit a rise in open-loop output impedance at very low frequencies. This is rarely of concern, though, owing to the very high loop gain down there.



**Figure 5.33.** The low-frequency gain of rail-to-rail output op-amps may depend strongly on load resistance, as seen here for the LMC6482.



**Figure 5.34.** For some RRO op-amps the open-loop output impedance rises markedly at low frequencies, owing to internal capacitive negative feedback around the output stage that becomes ineffective at low frequencies. But, not to worry, there's lots of loop gain at low frequencies in typical op-amp applications.

RRO op-amps will warn you that the output will not reach the negative rail (e.g., 10 mV for the bipolar LT6003); others will instruct you to add an external pull-down resistor or current sink (e.g., the bipolar LT1077, which saturates to 3 mV with no pull-down, and 0.1 mV with a 5kΩ pull-down). Op-amps with clean MOSFET saturation will tell you not to worry – the unloaded output will go all the way to ground (e.g.,  $\lesssim 0.1$  mV for the CMOS AD8616 or AD8691).

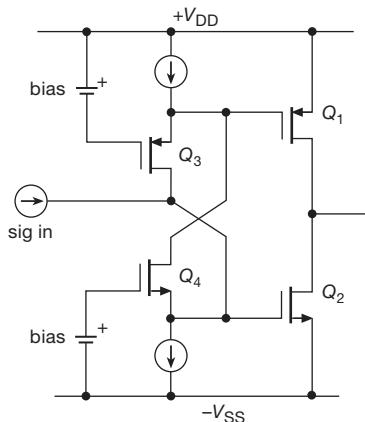
### C. Distortion

The rail-to-rail output stage (Figure 5.32B) presents real challenges to the chip designer when it comes to quiescent biasing and reduction of crossover distortion. Despite heroic efforts, these amplifiers generally perform some 20–

40 dB worse than their conventional (non-RRO) counterparts in terms of distortion, as seen in the pair of plots in Figures 5.43 (non-RRO) and 5.44 (mostly RRO);<sup>28</sup> see also the SPICE plots in §4x.11.

### D. Monticelli's output circuit

An elegant RRO circuit solution was devised by Monticelli,<sup>29</sup> and it is shown here in simplified form in Figure 5.35 (there's a full discussion in §4x.11). It has the effect of biasing the push-pull pair  $Q_1 Q_2$  in such a way that there is current overlap at crossover, and, better still, there is continuing current through both transistors *throughout the output swing*. We might call this “push–pull class-A” mode (though it seems to have been named already: “class-AA”). It is used, for example, in the CMOS OPA365 and in the BJT OPA1641. And it works – these parts have  $-114$  dB and  $-126$  dB harmonic distortion, respectively.



**Figure 5.35.** The Monticelli rail-to-rail output circuit.

Here's a capsule description of the Monticelli circuit's operation: first, think of  $Q_3$  and  $Q_4$  each as unity-gain current amplifiers whose source terminal is the “summing

<sup>28</sup> In fairness, we note that some of the poorer “distortion” results (which are actually THD+N – distortion plus noise) may be due to the lower supply voltages of the RRO op-amps, necessitating lower signal levels, thus causing noise to loom larger.

<sup>29</sup> See his patent US4570128, and his IEEE JSSC paper (SC-21, #6, 1986), in which he says “The output stage (Figure 8) must solve a level shifting problem that has plagued rail-to-rail designs for some time. Elaborate solutions have been proposed that combine multiple embedded feedback loops that are in effect op amps within op amps. To succeed as a general-purpose quad, a simpler solution had to be found.” Although originally developed at NSC, this circuit (or close variations) is popular with op-amp designers at Analog Devices and at TI (even before it swallowed NSC).

junction" (because the gate is held at fixed voltage). Now imagine an increasing input signal current, which reduces the net current sunk at  $Q_4$ 's source. This reduces its  $V_{GS}$ , which increases  $Q_2$ 's  $V_{GS}$ , thereby increasing the output pull-down current. Meanwhile, the reduced drain current in  $Q_4$  causes less of  $Q_3$ 's source current to be diverted, thus increasing the  $V_{GS}$  of  $Q_3$ ; that causes a reduction in  $Q_1$ 's  $V_{GS}$ , and therefore a lower output pullup current. The overall quiescent current is set by the dc bias applied to  $Q_3$  and  $Q_4$ . So it's a nicely balanced circuit, with a single-ended current input and a push-pull current output.

This is one cool circuit! In §4x.11 there's a more complete description, including SPICE simulations of a BJT implementation, and comparison with a conventional (not rail-to-rail) class-AB push-pull emitter follower. This inherently symmetrical circuit also works well with differential current drives to the drains of both  $Q_3$  and  $Q_4$ , a configuration you'll often see.

## 5.10 Choosing a precision op-amp

If there's no such thing as a perfect operational amplifier, then that's especially true for precision op-amps. Although sufficient perfection may be achieved in a few parameters, the design tradeoffs required for achieving this invariably degrade other parameters. For example, if we need a very quiet medium-frequency op-amp, a world-class quiet IC, we won't be able to enjoy world-class low-input-bias currents.<sup>30</sup> That's because the amplifier will use bipolar input transistors, which will have to be operated at fairly high collector currents, and you know what that means for the base currents (e.g., look at the LT1028). Another example: if we want micropower operating current, we won't be able to enjoy world-class fast settling time, because we won't be able to have a high  $f_T$  and fast slew rates; that takes power, and lots of it.

In this section we take an in-depth look at the process of choosing a precision op-amp that is right for the job at hand, linked closely to a broad selection of exemplary parts in Tables 5.5 (pages 320–321) and 5.6 (on page 335). If you've got a circuit design that you've been struggling with, this section should hit the spot. The nitty-gritty level of detail that follows is essential to the careful design that distinguishes an excellent circuit from a compromise-ridden also-ran. For the casual reader, on the other hand,

the level of detail in the following treatment may be, well, "not superficial enough."<sup>31</sup>

As we begin our tour of precision op-amp parameters and their significance, we invite you to bury yourself in the data. With your circuit design goals in mind, start with an important op-amp parameter and look for the best choices. After zeroing in on a winning value, you can examine other parameters for that op-amp: do some of the other parameters for our winning op-amp now look like poor choices? Maybe your op-amp isn't a winner after all. Or perhaps you've got to return to your design goals and adjust them in accordance with reality, and repeat the process. Remember always that "engineering is the art of compromise."

### 5.10.1 "Seven precision op-amps"

Seven is a nice number, and, in preparation for the extended discussion of the very practical issue of choosing a precision op-amp, we provide in Table 5.5 (pages 320–321) a comparison of the important specifications for an updated listing of seven of our favorite precision op-amps. The problem is, we just couldn't restrict ourselves to a mere seven – it's closer to seven dozen! Spend some time with it (and check off your own seven faves!) – it will give you a good feeling for the trade-offs you face in high-performance design with op-amps. Note particularly the trade-offs of offset voltage (and drift) versus input current for the best bipolar and JFET op-amps. You also get the lowest noise voltage from bipolar op-amps, trending downward with increasing bias current; we'll see why that happens later in Chapter 8 when we discuss noise. The awards for low-input current, however, always go to the FET op-amps, again for reasons that will become clear later. In general, choose FET op-amps for low-input current and current noise; choose bipolar op-amps for low-input voltage offset, drift, and voltage noise.

Among FET-input op-amps, those using JFETs dominate the scene, particularly where precision combined with low noise is needed (but not *all* JFET op-amps: note that our jellybean favorites, the LF411/412, are not precise enough to qualify for membership in the table). That dominance is being challenged, though, by some low-voltage CMOS parts like the factory-trimmed MAX4236A and OPA376, and by parts like the TLC4501A that use tricks such as auto-zeroing at power-up.<sup>32</sup>

<sup>30</sup> In §8.6.3 we show a discrete op-amp circuit where both of these goals are achieved.

<sup>31</sup> A phrase lifted from a student's reply in an end-of-course questionnaire: "This course was not superficial enough for me."

<sup>32</sup> There was traditionally a problem peculiar to MOSFETs, which has been largely solved through process improvements. MOS transistors

Table 5.5 "Seven" precision op-amps (page 1: high voltage)

Part #	#	Supply <sup>p</sup> per pkg	Input Current		Offset Voltage		Noise <sup>t</sup>		F <sub>R</sub>		Settle <sup>s</sup>		Swing to		Cost				
			Range (V)	I <sub>o</sub> (mA)	typ max	V <sub>os</sub> ( $\mu$ V)	typ max	$\Delta V_{os}$ ( $\mu$ V)	V <sub>npp</sub> min dc <sup>b</sup> ( $\mu$ V) dB	V <sub>npp</sub> max dc <sup>b</sup> ( $\mu$ V) dB	$i_n$ 1kHz (fA/ $\sqrt{\text{Hz}}$ )	GBW typ (MHz)	Slew typ (V/ $\mu$ s)	C <sub>in</sub> dist. pf	C <sub>in</sub> IN OUT pf	DIP avail qty 25 shdn pN + -	DIP price (\$US)	Comments	
<i>HV bipolar</i>																			
□ LT1077A	1	2.2-44	0.05	7nA	9	40	0.4	1.6	97	0.5	27	65	0.23	0.08	-	-	3.84		
□ LT1490A	2,4	2.5-44	0.04	1nA	8nA	110	500	2	4	84	1	50	0.2	0.07	-	-	3.25		
□ AD8622A	2,4	4-36	0.22	45nA	200nA	10	125	0.5	1.2	125	0.2	11	150	0.56	0.48	-	-	5.33	
□ LT1013 <sup>f</sup>	2,4	3.4-44	0.35	12nA	20nA	40	150	0.4	2	100	0.55	22	70	0.7	0.4	-	-	3.13	
□ OPA277	1,2,4	4-36	0.79	0.5nA	1nA	10	20	0.1	0.15	130	0.22	8	200	1	0.8	16	0.1	3.17 improved OP-27	
□ TLE2141A	1,2,4	4-44	3.5	0.7 $\mu$ A	1.5 $\mu$ A	175	500	1.7	-	85	0.5	10.5	1900	5.9	45	0.4	-	1.15 cross-coupled slew	
□ LT1677	1	3-44	2.8	2nA <sup>e</sup>	20nA <sup>e</sup>	20	60	0.4	2	109	0.09	3.2	1200	7.2	5	4.2	-	an "RIO LT1007"	
□ AD8675	1,2	9-36	2.5	0.5nA	2nA	10	75	0.2	0.6	114	0.1	2.8	300	10	2.5	-	-	2.22 0.6 ppm dist, dual=76	
□ OPA2209	1,2,4	4.5-40	2.2	1nA	4.5nA	35	150	1	3	120	0.13	2.2	500	18	6.4	2.1	-	4.04 0.25 ppm dist, SOT23	
□ LT1007 <sup>g</sup>	1	4-44	2.7	10nA	35nA	10	25	0.2	0.6	117	0.06	2.5	400	8	2.5	-	-	2.48 '37 decomp 60MHz	
□ ADA4004	1,2,4	9-36	2.2	40nA	90nA	40	125	0.7	1	110	0.15	1.8	3500	12	2.7	-	-	4.20 family, SOT23	
□ LT1468	1	7-36	3.9	3nA	10nA	30	75	0.7	2	96	0.3	5	600	90	23	0.8	4	4.26 0.7ppm dist	
□ AD8597	1,2	9-36	4.8	40nA	210nA	10	120	0.8	2.2	120	0.08	1.1	4300	10	16	2.0	12	3.71 1ppm dist	
□ LT1028A <sup>h</sup>	1	8-44	7.4	25nA	90nA	10	40	0.2	0.8	108	0.04	0.85	4700	75	15	-	5	6.48 lower $I_B$ than AD697	
<i>HV superbeta</i>																			
□ LT61010An	1,2,4	2.7-40	0.14	20	110	10	35	0.2	0.8	107	0.4	14	100	0.35	0.11	45	4	2.22 replace LT1012, w/RRO	
□ LT1012AC <sup>m</sup>	1	2.4-40	0.37	25	100	8	25	0.2	0.6	114	0.5	14	(20)	0.5	0.2	-	-	5.11 has overcomp pin	
□ OP97E	1,2,4	4.5-40	0.40	30	100	10	25	0.2	0.6	114	0.5	14	(20)	0.9	0.2	-	-	5.52 dual='297, quad='497	
□ AD706	2,4	4-36	0.8	50	200	30	100	0.2	1.5	106	0.5	15	50	0.8	1.5	-	2	4.05 AD704 quad	
□ LT1884A	2,4	3.5-40	0.85	150	400	25	50	0.3	0.8	114	0.4	9.5	50	2	0.9	10	-	5.23 LT1882 slower, lower $I_B$	
<i>HV JFET</i>																			
□ AD795	1	8-36	1.3	1	2	100	500	3	10	90	1	9	0.6	1.6	1	11	2.2	-	
□ OPA124PB	1	10-36	2.5	0.35	1	100	250	1	2	100	1.6	8	0.5	1.5	1.6	10	3	-	
□ OPA140	1,2,4	4.5-40	1.8	0.5	10	30	120	0.35	1	126	0.25	5.1	0.8	11	20	0.9	-	6.40 substrate pin	
□ AD711C	1,2	9-36	2.5	15	25	100	250	2	5	86	2	16	10	4	20	1.0	5.5	3.75 0.5ppm dist	
□ LT1055Cx	1	20-40	2.8	10	50	120	700	3	12	85	2	15	1.8	4.5	12	1.8	4	2.19 AD712 low-cost dual	
□ ADA4000	1,2,4	5-36	2	5	40	200	1700	2	-	80	1	16	10	12	2.7	-	5.5	2.52 1057 dual, 1058 quad	
□ OPA192	1,2,4	4.5-40	1	5	20	5	25	0.2	0.5	120	1.3	5.5	10	1.5	10	20	0.9	1.46 jellybean, AD711 subst	
□ OPA134	1,2,4	5-36	4	5	100	500	200	2	-	86	-	8	3	8	20	1	-	3.87 CMOS, e-Trim™	
□ OPA1641	1,2,4	4.5-40	1.8	2	20	1000	3500	-	-	120	-	5.1	0.8	20	11	-	-	1.60 0.8 ppm dist, jellybean	
□ AD8320A	1,2	10-27	2.5	2	10	85	250	0.5	1	90	1.8	6	5	25	60	0.6	15	2.76 0.5 ppm dist, family	
□ OPA827	1	8-40	4.8	15	50	75	150	1.5	-	104	0.25	3.8	2.2	22	28	0.55	9	11.86 caution, $\pm 12\text{V}$ max	
□ OPA27B	1	10-36	7	1	5	40	100	0.4	0.8	106	0.6	4.5	1.6	16	55	0.55	7	9.00 cheaper than 627	
□ OPA337B	1	10-36	7	1	5	40	100	0.4	0.8	106	0.6	4.5	1.6	80	135	0.45	7	30.00 ADA4627B 2nd-source	
□ AD449KH	1	10-36	0.6	75fA <sup>z</sup>	0.1	150	250	2	5	90	4	35	0.5	1	3	5	-	28.00 L=60fA, TO-99 pkg	
□ OPA29B	1	10-36	1.2	30fA <sup>z</sup>	0.1	500	2000	3	10	80	4	17	0.1	1	2.5	-	2	12.00 lowest $I_B$ for HV part	
<i>HV chopper</i>																			
□ LTC1150	1	4.8-32	0.8 <sup>s</sup>	10	100	0.5	10	0.01	0.05	110	1.8	high (f <sub>18</sub> )	2.5	3	-	-	6.00 only HV AZ w/int caps		
□ ADA4638	1	4.5-33	0.85	45	90	0.5	4.5	-	0.08	130	1.2	66	100	1.5	4	9	-	5kHz noise spike	
□ OPA2188	1,2,4	4-40	0.42	160	850	6	25	0.03	0.09	120	0.25	8.8	750	2	0.8	27	9.5	-	new

Table 5.5 "Seven" precision op-amps (page 2: low voltage)

Part #	#	Supply <sup>a</sup> per pkg (V)	Input Current Range mA	Offset Voltage @25°C		Noise <sup>t</sup>		GBW MHz		Slew typ typ		Swing to Supply Pin		DIP avail qty 25 (\$US)	Comments				
				typ	max	V <sub>os</sub> typ (μV)	max (μV)	min dB	C <sub>m</sub> nV/ V	i <sub>n</sub> fA/ V	GBW typ (MHz)	Slew typ typ	C <sub>in</sub> pF	IN OUT					
<i>LV bipolar</i>																			
□ L76003	1,2,4	1.5-18	0.001	40	140	175	500	2	5	73	3	325	12	0.003	0.001	slow	-		
□ EL8176	1	2.4-6	0.05	0.5nA	2nA	25	100	0.7	-	90	1.5	28	160	0.4	0.13	-	-		
□ □ LMP7731	1,2	1.8-6	2.2	1.5nA	30nA	6	500	0.5	101	0.08	2.9	1100	22	2.4	-	-	2.30 SOT-23, charge pump		
□ □ LT6220	1,2,4	2.2-13	0.9	15nA	150nA	70	350	1.5	5	85	10	800	60	20	0.3	-	1.98 SOT-23, 3nV/√Hz		
□ □ LT6230	1,2,4	3-12.6	3.3	5μA	10μA	100	500	0.5	3	96	0.18	1.1	2400	215	70	0.05	-	1.75 SOT-23	
□ □ L76230	1,2,4	3-12.6	3.3	5μA	10μA	100	500	0.5	3	96	0.18	1.1	2400	215	70	0.05	-	2.72 -10=decomp, 1.3GHz	
<i>LV JFET</i>																			
□ OPA656	1	9-13	14	2	20	250	1800	2	12	80	-	7	1.3	230	290	0.02	2.8	-	-
□ □ L76003	1,2,4	1.8-6	0.01	0.02	1	10	150	0.3	0.75	81	2.3	60	0.1c	0.13	0.06	slow	-	-	
□ □ LMP2322A	1,2,4	1.8-6	0.01	1	10	-	800	2	-	61	-	105	-	0.12	0.04	-	-	3.29 '31 single, '34 quad	
□ □ TS611A	1,2	1.8-6	0.04	0.2	1	12	50	1	4.5	85	2.3	25	50	0.4	0.1	23	-	0.82 '612 dual	
□ □ AD8603	1,2,4	1.8-6	0.04	0.2	1	7	25	0.2	0.7	95	1	18	0.25c	0.75	0.05	24	-	1.36 SOT-23, 1pA	
□ □ LTC6078	2,4	2.7-6	0.05	0.2	1	-	70	0.2	0.8	93	1.3	13	0.5	3.6	1	6	-	3.54 V <sub>os</sub> degrades near V <sub>+</sub>	
□ □ LTC6081	2,4	2.7-6	0.33	0.2	1	-	500	5	20	0.6	2	84	0.2	14	0.6	1.7	0.3	4.75 degrade V <sub>cm</sub> >V <sub>c</sub> -1.5V	
□ □ MAX4236A	1	2.4-6	0.35	1	-	-	110 <sup>v</sup>	1	-	70	-	37	30	1.5	-	3	-	1.78 not AZ, '4237 decomp	
□ □ LMC6482A	2,4	3-16	0.5	0.02	4	110 <sup>v</sup>	750	1	-	-	-	-	-	-	-	-	-	1.88 '7101=SOT-23	
□ □ OPA376	1,2,4	2.2-7	0.76	0.2	10	5	25	0.26	1	76	0.8	7.5	0.25c	5.5	2	2	-	1.32 CMOS Etrim™	
□ □ OPA364	1,2	1.8-5.5	0.85	1	10	-	500	3	-	74	10	17	0.6	7	5	1.5	-	2.18 chg-pump, 20ppm dist	
□ □ TLC4501A	1,2	4-7	1	1	60	10	40	1	-	90	1.5	12	0.6	4.7	2.5	2.2	●	3.06 0.3s self-cal at pwr-on	
□ □ OPA743	1,2,4	3.5-13	1.1	1	10	1500	7000	8	-	66	11	30	2.5	7	10	15	●	1.58 jellybean	
□ □ LMP7715	1,2	1.8-6	1.15	0.05	1	10	150	1	4	85	-	5.8	10	17	9.5	-	15	-	
□ □ OPA376	1,2,4	1.8-6	1.15	0.05	1	10	150	1	4	85	-	6.2	10	88	28	-	15	-	
□ □ AD8692	1,2,4	2.7-6	0.85	0.2	1	400	2000	0.3	6	68	1.6	8	50	10	5	1	-	1.36 0.6ppm dist	
□ □ AD8616	1,2,4	2.7-6	1.7	0.2	1	23	60 <sup>f</sup>	1.5	7	80	2.4	7	50	24	12	0.5	7	-	
□ □ OPA350	1,2,4	2.5-7	5.2	0.5	10	150	500	4	-	74	-	7	4	38	22	0.5	6.5	-	
□ □ OPA380	1,2	2.7-7	7.5	3	50	4	25	0.03	1	100	3	5.8	10	90	2	3	-	5.39 transimp, w/auto-zero	
□ □ LMC6001A	1	5-16	0.45	10fA	25fA	-	350	2.5	10	83	-	22	0.13	1.3	1.5	-	-	12.19	
□ □ LMP7721	1	1.8-6	1.3	3fA <sup>y</sup>	20fA	26	150	4	83	-	7	10	17	10	-	-	-	11.89 very low en for 20fA!	
□ □ MAX9617	1	1.6-6	0.06	10	140	0.8	10	0.01	0.12	116	0.42	100	1.5	0.7	-	-	●	-	
□ □ AD8638	1,2	5-16	1	1.5	40	3	9	0.01	0.06	118	1.2	60	noisy	1.35	2.5	3	4	-	
□ □ LTC2050H	1	2.7-11	0.8	7	50	0.5	3	-	0.03	120	1.5	-	-	3	2	-	-	3.31 auto-zero, SOT23	
□ □ AD8551	1,2,4	2.7-6	0.7	10	50	1	5	0.01	0.04	120	1	42	(2)	1.5	0.4	50	-	2.19 replaces LTC1050	
□ □ OPA735	1,2	2.7-13	0.6	100	200	1	5	0.01	0.05	115	2.5	135	40	1.8	1.5	-	10	-	
□ □ L76003	1,2,4	1.5-18	0.001	40	140	0.8	10	0.01	0.12	116	0.42	100	1.5	0.7	-	-	●	1.60 charge pump	
□ □ EL8176	1	2.4-6	0.05	0.5nA	2nA	25	100	0.7	-	90	1.5	28	160	0.4	0.13	-	-	-	
□ □ LMP7731	1,2	1.8-6	2.2	1.5nA	30nA	6	500	0.5	101	0.08	2.9	1100	22	2.4	-	-	-	-	
□ □ LT6220	1,2,4	2.2-13	0.9	15nA	150nA	70	350	1.5	5	85	10	800	60	20	0.3	-	-	-	
□ □ LT6230	1,2,4	3-12.6	3.3	5μA	10μA	100	500	0.5	3	96	0.18	1.1	2400	215	70	0.05	-	-	-
□ □ L76230	1,2,4	3-12.6	3.3	5μA	10μA	100	500	0.5	3	96	0.18	1.1	2400	215	70	0.05	-	-	-
□ □ L76230	1,2,4	3-12.6	3.3	5μA	10μA	100	500	0.5	3	96	0.18	1.1	2400	215	70	0.05	-	-	-
□ □ L76230	1,2,4	3-12.6	3.3	5μA	10μA	100	500	0.5	3	96	0.18	1.1	2400	215	70	0.05	-	-	-
□ □ L76230	1,2,4	3-12.6	3.3	5μA	10μA	100	500	0.5	3	96	0.18	1.1	2400	215	70	0.05	-	-	-
□ □ L76230	1,2,4	3-12.6	3.3	5μA	10μA	100	500	0.5	3	96	0.18	1.1	2400	215	70	0.05	-	-	-
□ □ L76230	1,2,4	3-12.6	3.3	5μA	10μA	100	500	0.5	3	96	0.18	1.1	2400	215	70	0.05	-	-	-
□ □ L76230	1,2,4	3-12.6	3.3	5μA	10μA	100	500	0.5	3	96	0.18	1.1	2400	215	70	0.05	-	-	-
□ □ L76230	1,2,4	3-12.6	3.3	5μA	10μA	100	500	0.5	3	96	0.18	1.1	2400	215	70	0.05	-	-	-
□ □ L76230	1,2,4	3-12.6	3.3	5μA	10μA	100	500	0.5	3	96	0.18	1.1	2400	215	70	0.05	-	-	-
□ □ L76230	1,2,4	3-12.6	3.3	5μA	10μA	100	500	0.5	3	96	0.18	1.1	2400	215	70	0.05	-	-	-
□ □ L76230	1,2,4	3-12.6	3.3	5μA	10μA	100	500	0.5	3	96	0.18	1.1	2400	215	70	0.05	-	-	-
□ □ L76230	1,2,4	3-12.6	3.3	5μA	10μA	100	500	0.5	3	96	0.18	1.1	2400	215	70	0.05	-	-	-
□ □ L76230	1,2,4	3-12.6	3.3	5μA	10μA	100	500	0.5	3	96	0.18	1.1	2400	215	70	0.05	-	-	-
□ □ L76230	1,2,4	3-12.6	3.3	5μA	10μA	100	500	0.5	3	96	0.18	1.1	2400	215	70	0.05	-	-	-
□ □ L76230	1,2,4	3-12.6	3.3	5μA	10μA	100	500	0.5	3	96	0.18	1.1	2400	215	70	0.05	-	-	-
□ □ L76230	1,2,4	3-12.6	3.3	5μA	10μA	100	500	0.5	3	96	0.18	1.1	2400	215	70	0.05	-	-	-
□ □ L76230	1,2,4	3-12.6	3.3	5μA	10μA	100	500	0.5	3	96	0.18	1.1	2400	215	70	0.05	-	-	-
□ □ L76230	1,2,4	3-12.6	3.3	5μA	10μA	100	500	0.5	3	96	0.18	1.1	2400	215	70	0.05	-	-	-
□ □ L76230	1,2,4	3-12.6	3.3	5μA	10μA	100	500	0.5	3	96	0.18	1.1	2400	215	70	0.05	-	-	-
□ □ L76230	1,2,4	3-12.6	3.3	5μA	10μA	100	500	0.5	3	96	0.18	1.1	2400	215	70	0.05	-	-	-
□ □ L76230	1,2,4	3-12.6	3.3	5μA	10μA	100	500	0.5	3	96	0.18	1.1	2400	215	70	0.05	-	-	-
□ □ L76230	1,2,4	3-12.6	3.3	5μA	10μA	100	500	0.5	3	96	0.18	1.1	2400	215	70	0.05	-	-	-
□ □ L76230	1,2,4	3-12.6	3.3	5μA	10μA	100	500	0.5	3	96	0.18	1.1	2400	215	70	0.05	-	-	-
□ □ L76230	1,2,4	3-12.6	3.3	5μA	10μA	100	500	0.5	3	96	0.18	1.1	2400	215	70	0.05	-	-	-
□ □ L76230	1,2,4	3-12.6	3.3	5μA	10μA	100	500	0.5	3	96	0.18	1.1	2400	215	70	0.05	-	-	-
□ □ L76230	1,2,4	3-12.6	3.3	5μA	10μA	100	500	0.5	3	96	0.18	1.1	2400	215	70	0.05	-	-	-
□ □ L76230	1,2,4	3-12.6	3.3	5μA	10μA	100	500	0.5	3	96	0.18	1.1	2400	215	70	0.05	-	-	-
□ □ L76230	1,2,4	3-12.6	3.3	5μA	10μA	100	500	0.5	3	96	0.18	1.1	2400	215	70	0.05	-	-	-
□ □ L76230	1,2,4	3-12.6	3.3	5μA	10μA	100	500	0.5	3	96	0.18	1.1	2400	215	70	0.			

Finally, the so-called *chopper-stabilized* (here and in Table 5.6 on page 335) amplifiers form the most important exception to the generalization that FET op-amps, particularly MOSFET types, suffer from larger initial offsets and much larger drifts of  $V_{OS}$  with temperature and time than do bipolar-transistor op-amps. In fact, these devices (known also as *auto-zero* or *zero-drift* amplifiers) are the amplifiers with the *smallest* offset voltage and drift, typically in the  $\pm 1 \mu\text{V}$  and  $\pm 0.05 \mu\text{V}/^\circ\text{C}$  range. They use MOSFET analog switches and amplifiers to sense, and correct, the residual offset error of an ordinary op-amp (which itself is often built with MOSFETs, on the same chip). This is not without compromise, however: chopper-stabilized amplifiers have some unpleasant characteristics that make them unsuited for many applications, as we'll see in §5.11.

### 5.10.2 Number per package

The first column in Table 5.5 gives the choices available for the number of devices per package (the number in **boldface** shows which choice matches the part number). We generally list single op-amp parts, even though in practice the dual op-amps are more useful and popular (in some cases distributors don't even stock the single types). Special features such as pins for external offset-nulling, compensation, and shutdown are available only for the single op-amp package types and are indicated in the right-hand columns. Generally the specs are identical for the different dies and appear on the same datasheet, but not always.

### 5.10.3 Supply voltage, signal range

It's likely your first consideration will be supply-voltage range and signal levels. High-voltage parts (able to operate from  $\pm 15 \text{ V}$ , i.e.,  $30 \text{ V}$  total) are listed first in the table,

are susceptible to a unique debilitating effect that neither FETs nor bipolar transistors have. It turns out that sodium-ion impurity migration and/or phosphorus polarization effects in the gate insulating layer can cause offset voltage drifts under closed-loop conditions, in extreme cases as much as  $0.5 \text{ mV}$  over a period of years. The effect is increased for elevated temperatures and for a large applied differential-input signal, with some datasheets showing a typical  $5 \text{ mV}$  change of  $V_{OS}$  over 3000 hours of operation at  $125^\circ\text{C}$  with  $2 \text{ V}$  across the input. This sodium-ion disease can be alleviated by introducing phosphorus into the gate region. Texas Instruments, for example, uses a phosphorus-doped polysilicon gate in its "LinCMOS" series of op-amps (TLC270-series) and comparators (TLC339 and TLC370-series). These popular inexpensive parts come in a variety of packages and speed/power selections and maintain respectable offset voltages with time ( $50 \mu\text{V}$  eventual offset drift per volt of differential input).

with parts ordered more or less by  $I_Q$ , the quiescent supply current, in each category. Battery-powered applications benefit from low supply currents, but some low-drift applications do as well, because op-amp self-heating temperature effects will be less. Some parts offer a version with a power-shutdown (SHDN) pin. For example, the LT6010 current drops from  $135 \mu\text{A}$  to  $12 \mu\text{A}$  in shutdown (but a *gotcha*: the shutdown pin itself takes another  $15 \mu\text{A}$ ) and it takes  $25 \mu\text{s}$  to turn ON or OFF. Other parts do better, e.g., the OPA364 draws  $0.9 \mu\text{A}$  when off.

Circuits operating with high-voltage supplies benefit from using high signal levels, such as  $\pm 10 \text{ V}$  full scale. An offset voltage, say  $V_{OS}=40 \mu\text{V}$ , is a smaller fraction of  $20 \text{ V}_{pp}$  than it is of  $0$  to  $4 \text{ V}$ . With the exception of chopper op-amps, you don't get any offset-voltage improvement for low-voltage parts.

Low-voltage parts finish the table (most are  $5.5 \text{ V}$  maximum total supply, but some permit  $11 \text{ V}$  or higher, suitable for  $\pm 5 \text{ V}$  operation), but it's important to realize that many "high-voltage" parts are designed and specified to work well at low voltages, even under  $3 \text{ V}$ . Some work well with  $\pm 3 \text{ V}$  to  $\pm 5 \text{ V}$  supplies and should not be rejected simply because they can also work at higher voltages. But be warned, you need to examine the common-mode input range and output-swing range. For example, although a  $44 \text{ V}$  op-amp like the LT1490 works with  $3 \text{ V}$  supplies and allows rail-rail inputs and outputs, another fine  $44 \text{ V}$  LTC op-amp, the low-noise LT1007 (which works down to  $4 \text{ V}$ ) is limited to inputs and outputs no closer than  $2 \text{ V}$  from the rails – nearly useless when running from  $\pm 2 \text{ V}$ . Clearly it's not meant to be a low-voltage part. Your quick guide to these issues is the "swing to supply" columns; the LT1490 has all four checks whereas the LT1007 has no checks.

### 5.10.4 Single-supply operation

If you're running with low supply voltages, you may want to use a single-supply power arrangement. Op-amps capable of single-supply use have at a minimum the ability to operate their inputs and outputs to the negative rail (i.e., ground). Many permit operation with outputs also to the positive rail, and claim rail-to-rail outputs on the front page of the datasheet. But be warned, there's usually a performance degradation when outputs are near the supply rails. Some op-amps offer zero-volt or below-the-rail operation if you add a pulldown resistor.<sup>33</sup>

Seven high-voltage op-amps on our list offer single-

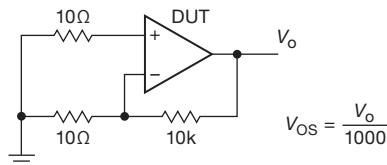
<sup>33</sup> Many op-amps can do this, but without saying so. That's because their pullup transistor and drivers work all the way down to the negative

supply operation; some, like the LT1013, excel at it. Two offer full rail-to-rail input-output, or RRIO, operation. The fast-slewing TLE2141 is especially interesting (fast settling, but high bias current), as is the low-noise LT1677 (lower bias current, but slow slewing and settling). All but two of the low-voltage parts offer single-supply operation.

There are some precision op-amps with low supply currents, down to 10–60  $\mu\text{A}$ , although this will severely limit your choices for other parameters. There's even a respectable 0.85  $\mu\text{A}$  (and 1.8 V) op-amp, the LT6003. Some op-amp types, such as the JFETs, don't offer any low-power parts. Nonetheless you might choose one for its low noise and low bias current.

### 5.10.5 Offset voltage

Perhaps the single parameter most often associated with precision amplifiers is input-voltage error. To measure small offset voltages, use the op-amp's gain to magnify the effect, as shown in Figure 5.36. Offset voltage was our required parameter to gain entry into the table; few op-amps with maximum offset voltages above 250  $\mu\text{V}$  made the grade. There are plenty of parts with <10  $\mu\text{V}$  *typical* offset voltages, but “typical” isn't a reliable spec when you're in the business of manufacturing precision instruments. Bipolar input stages have an advantage over JFET and CMOS in the table, but they suffer from higher input bias currents. The “superbeta” parts are a pleasant exception, especially at high temperatures (see Figure 5.38), but none of these parts has inputs that operate to either supply rail.



**Figure 5.36.** Offset-voltage test circuit. The  $\times 1000$  voltage gain makes submillivolt offsets in the device under test (DUT) easily measurable. Input *current* effects are negligible, owing to the small ( $10\Omega$ ) resistances seen at the op-amp's inputs. Add a  $200\Omega$  resistor at the output if you want to drive a cable.

Many low-offset parts have disappeared since the second edition of our book, especially in the JFET category. They've become too expensive for the market and have lost the competition with low-voltage chopper and auto-zero op-amps. The latter have a few token representatives in this

rail, without sourcing current to the output. Some require a minimum pull-down current, e.g., 0.5 mA for the OPA364.

table, but they have their own selection table (Table 5.6 on page 335) that you should examine if they look good for your design. They have their own problem issues, such as current noise, which we discuss in §5.11.

Offset voltage variation with common-mode input voltage is a serious issue for some parts, especially for RRIO op-amps (see Figure 5.29), and is not addressed on the table. It's always important to follow up an initial choice from the table with a careful examination of the datasheet. For example, the OPA364 and MAX9617 use internal charge pumps to power their input stages, eliminating this problem completely.

Offset voltage drift with temperature is an important parameter when stability of measurements matters. This parameter is not production tested. The maximum drift spec may not be very reliable, and some manufacturers have stopped providing such a spec.

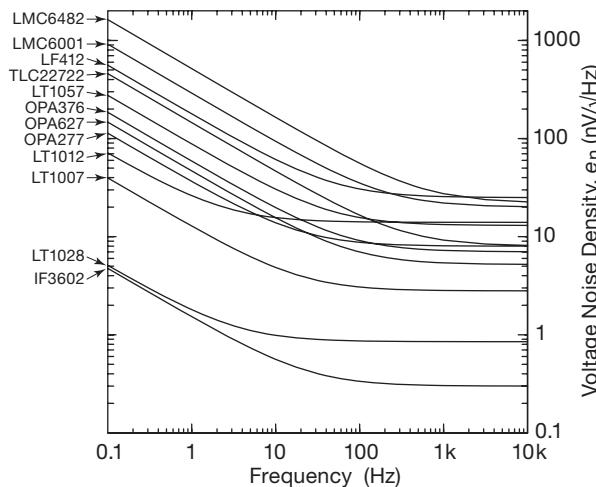
Offset voltage drift with time is a parameter that used to appear in precision op-amp datasheets, with values of order 300–400 nV/month; a few high-performance parts like the LT1007 claimed 200 nV/month, and chopper-stabilized op-amps generally claim drifts of 50 nV/month.<sup>34</sup> This is somewhat uncharted territory, and some people claim that drift slows with time, or perhaps it is more akin to a random walk, in either case suggesting that a drift specification should perhaps have units of nV/ $\sqrt{\text{month}}$ .

### 5.10.6 Voltage noise

Voltage noise is the in-band variation of op-amp input-offset voltage that's indistinguishable from signal. It's useful to view it as a “noise spectral density” function  $e_n(f)$ , which tells you the rms noise voltage in a 1 Hz bandwidth (see §8.2.1) centered at frequency  $f$ . Figure 5.37 shows an idealized plot of input voltage-noise density as a function of frequency for some of the op-amps in the table. For most op-amps  $e_n(f)$  is essentially flat for frequencies above its “ $1/f$  corner frequency,” with  $e_n$  increasing below the corner frequency, approximately as  $1/\sqrt{f}$ . (Auto-zero, or “chopper-stabilized,” op-amps are not shown. They behave differently, because low-frequency “noise” is removed by the auto-zero process, so their  $e_n$  is flat at low frequencies. We'll discuss them shortly; see §5.11.)

The voltage-noise column in Table 5.5 (pages 320–321) shows  $e_n$  at its commonly specified frequency of 1 kHz,

<sup>34</sup> Although it's possible to do considerably better, for example the measured 6 nV/month reported by Bob Pease for the LMP2011. See Table 8.3 and Figures 8.60, 8.61, 8.110, and 8.110. The IF3602 is a large-geometry dual JFET available from InterFet, shown for comparison.



**Figure 5.37.** Voltage noise density  $e_n$  for a selection of representative op-amps, showing the increase of noise power below the  $1/f$  corner frequency. Some op-amps that have good specs at 1 kHz don't look so good at 0.1 Hz. Figure 5.54 shows the resulting voltage noise  $v_n$  when such a noise density is integrated over frequency.

comfortably in the flat region above the  $1/f$  corner of most op-amps;  $e_n$  varies from  $0.85 \text{ nV}/\sqrt{\text{Hz}}$  for the high-current LT1028 to  $325 \text{ nV}/\sqrt{\text{Hz}}$  for the current-starved LT6003.

We discuss noise in greater detail below, and in Chapter 8, but let's start with the simple relationship relating voltage-noise *density*  $e_n$  (given in units of  $\text{nV}/\sqrt{\text{Hz}}$ ) to the value of *integrated* total voltage noise  $V_n$  (in units of nV, or  $\mu\text{V}$ ; and either rms or peak-to-peak) over some frequency passband. In the flat (i.e., white noise) region of  $e_n$  at frequencies above the  $1/f$  corner, the integrated noise voltage is simply  $V_n = e_n \sqrt{\text{BW}}$ .

As we'll see, for circuits with bandwidths of 1–10 kHz or more, the noise voltage  $V_n$  is dominated by the noise density at high frequencies. Datasheets for most op-amps give an  $e_n$  value at 1 kHz, but some also specify it at 10 kHz, 100 kHz, or even 1 MHz; and they usually provide graphical plots of  $e_n$  versus frequency. Because the integrated noise voltage is dominated by  $e_n$  at the high-frequency end of the operating range, be sure to use that value (or a frequency-weighted eyeball average) in the simple formula above. Using a high-frequency  $e_n$  value is particularly important for transimpedance amplifiers, which suffer from " $e_n \omega C$ " current noise ( $i_n = e_n 2\pi f C_{in}$ ) at high frequencies.

### A. “ $1/f$ ” noise

We discuss  $1/f$  noise in greater detail in Chapter 8 (where we show how to determine the  $1/f$  noise corner, etc., in §8.13.4, but here we address the practical question “What effect do the scary-looking rising noise-density curves in Figure 5.37 have on my circuit noise?” The noise density  $e_n$  is indeed higher at low frequencies, but that density gets multiplied by a smaller frequency span. Put another way, the total noise *voltage* (as contrasted with the noise *density*  $e_n$ ) contributed by an op-amp depends both on its  $e_n$  and on the circuit's bandwidth. More precisely, the mean square noise voltage is the integral of  $e_n^2$  over the bandpass:

$$v_n^2 = \int_{f_a}^{f_b} e_n^2(f) df,$$

where  $e_n(f)$  is the noise spectral density (often plotted in datasheets), and the bandpass (or observation band) extends from  $f_a$  to  $f_b$ . We then get the rms noise voltage by taking the square root of  $v_n^2$ .

We perform the integrations and show the devastating effects of  $1/f$  noise in the integrated-noise plots later, in Figure 5.54 (in the context of auto-zero op-amps). More bandwidth means more noise, and all of the curves rise as  $\sqrt{f}$  at the high end. The op-amps are ranked in order by their high-frequency  $e_n$  values; it's interesting to compare their positions in Figure 5.37 with the rankings in Figure 5.54. At the low frequency end, the noise voltage of conventional op-amps levels off, because their rising  $1/f$  noise density makes up for the reduced bandwidth,<sup>35</sup> whereas the noise voltage of the auto-zero amplifiers continues its downward trend.

Let's use Figure 5.54 to explore a revealing example. The LT1012 has an  $e_n$  (at 1 kHz) of  $14 \text{ nV}/\sqrt{\text{Hz}}$ , and a 2.5 Hz corner frequency.<sup>36</sup> If it were used in a precision amplifier with a high-frequency cutoff beyond 1 Hz, for example, it would be less noisy than an OPA277, even though the latter has a lower noise density at 1 kHz of  $8 \text{ nV}/\sqrt{\text{Hz}}$ , because the latter has a 20 Hz noise corner (indicated by black dots). But the OPA277 wins back some respect when we see that it's dramatically quieter than a competing  $9 \text{ nV}/\sqrt{\text{Hz}}$  part, the TLC2272, which suffers from a much higher 330 Hz noise corner.<sup>37</sup>

<sup>35</sup> Well, not exactly: if the noise power density were truly to continue rising as  $1/f$ , the integral would diverge at zero frequency (dc). For Figure 5.54 we set the low frequency limit to be 0.01 Hz.

<sup>36</sup> How to determine the corner frequency? See the discussion in §8.13.4.

<sup>37</sup> The LT1012 and OPA277 are BJT parts, and the TLC2272 is a CMOS op-amp. The '2272 boasts a minuscule 60 pA max bias current, far better than the '277's 1 nA, but not much better than the '1012's amazing 100 pA.

The integrated-noise plots are revealing, but it's helpful to have a single number in the table to evaluate our op-amps. The "V<sub>npp</sub>" parameter in the table is the peak-to-peak voltage noise over a 0.1-10 Hz band. This shows an op-amp's "dc noise" as seen in the flat part of the curves in Figure 5.54. The values range up to 11  $\mu$ V<sub>pp</sub> (but parts like the LMC6482 avoid the competition by not listing any spec at all). The LT1028 is the winner with 35 nV<sub>pp</sub>, but the LMP7731 is a notable part at 80 nV<sub>pp</sub>, with its RRIO capabilities and its SOT-23 package. The ADA4075, with its 1 mV offset, didn't make the precision table, but its 60 nV<sub>pp</sub> noise level is attractive.

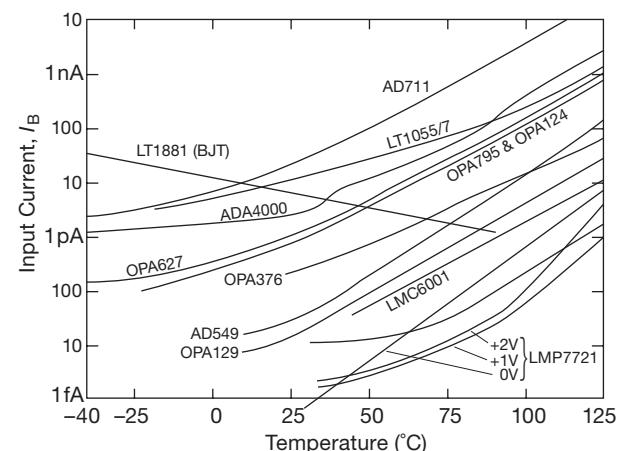
Precision op-amps that suffer from  $1/f$  noise (i.e., all except auto-zero types) have a V<sub>npp</sub> spec whose lower frequency limit is usually 0.1 Hz. If you need a lower starting frequency (such as 0.01 Hz, used in the plots), multiply the low-frequency noise-voltage value listed by the square root of the number of additional low-end decades you want (that's an interesting  $1/f$  noise factoid). As long as the  $1/f$  corner is multiple decades higher, you can ignore the white-noise contribution to the spectrum.

The V<sub>npp</sub> parameter is your primary clue about an op-amp's long-term drift performance.

### 5.10.7 Bias current

The available input bias currents range from femtoamps to microamps (nine orders of magnitude!). In some applications this is the parameter that rules out entire classes of op-amp choices. Parts featuring very low typical input currents often have unimpressive *maximum* specs; this is due to the difficulty and expense of automated testing at currents below 10 pA or so. For example, the \$1.88 low-cost LMC6482A CMOS op-amp has a 20 fA "typical" spec, but the datasheet shows a 4 pA maximum value – that's 200× worse.<sup>38</sup> However if you're willing to pay over \$10, you can get an LMP7721 with a 20 fA maximum spec.

As we've said earlier (and will say again), the "bias" current of JFET and CMOS op-amps is a *leakage* current, and it increases exponentially with temperature; see Figure 5.38. That's the bad news. The *good* news (as we've also said before) is that there are some op-amps (like the LT1012 and AD706) that have low JFET-like input currents, but which have BJT inputs and therefore enjoy better



**Figure 5.38.** Input current versus temperature for a representative set of op-amps from Table 5.5 on pages 320–321, taken from manufacturers' datasheets. See also Figures 5.6 and 3.48.

high-temperature performance and improved offset voltages and drift, see Figure 5.6.

Low-input-current op-amps generally have higher offset voltage and offset voltage drift, and they're usually more noisy. The OPA627 and ADA4627 JFET op-amps are exceptions, and they've served us well, but they're expensive. Happily the new JFET OPA827 has lower noise and offers some price relief. The AD743, not in the table because of its 1 mV offset voltage, sports a 2.9 nV/ $\sqrt{\text{Hz}}$  spec. Looking at CMOS parts (which are low voltage), we find that the LMP7715 is the best low-noise contender, at 5.8 nV/ $\sqrt{\text{Hz}}$ , but a low-cost part like the AD8616, with 1 pA bias and 7 nV/ $\sqrt{\text{Hz}}$  noise, may offer a good compromise.

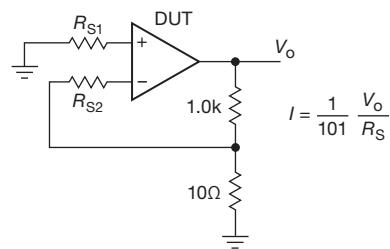
High-speed op-amps often have high input bias currents, typically 200 nA to as much as 20  $\mu$ A. They also tend to have high offset voltages, above 0.5 mV, so most parts in this category didn't even make it into the precision table, but appear in the high-speed op-amp category in Table 5.4 on page 310. TI's fast OPA656 and '657 low-voltage JFET op-amps, with 2 pA typical bias current, 290 V/ $\mu$ s slew rate, and 20 ns settling time, demanded and were granted residence in both tables. The OPA380 is a 90 MHz op-amp meant for fast transimpedance applications, featuring 50 pA and 25  $\mu$ V maximum offset. This part uses an auto-zero circuit to achieve 25  $\mu$ V offset, but avoids excess current noise with an isolating filter (Figure 5.41).

#### A. Measuring bias current

To measure input currents (or offset currents) down to the nanoamp level or so, you can use the simple circuit in Figure 5.39. For *really* small currents, though, you've got

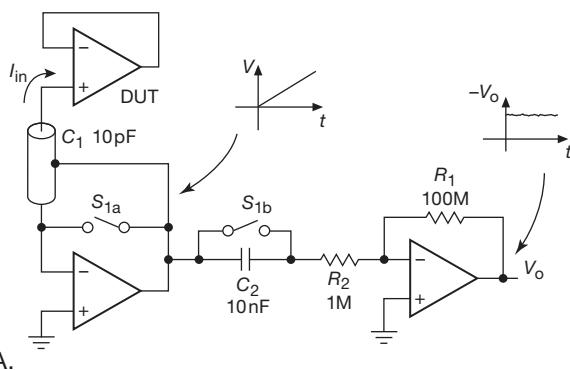
<sup>38</sup> Note 13 on the LMC6482 datasheet says "Guaranteed limits are dictated by tester limitations and not device performance. Actual performance is reflected in the typical value." That's illuminating, but not entirely helpful for the designer of a mass-production instrument.

to play some clever tricks: a current down in the femtoamps ( $10^{-15}$  A) develops only microvolts across a gigaohm! (And you'll never see that, because offset voltages are much larger.) Instead, accumulate the tiny input current in an integrator (itself constructed with an ultra-low-input-current op-amp), as shown in Figure 5.40A.<sup>39</sup> Here the short length of shielded cable (with Teflon dielectric) serves as the integrator's feedback capacitance  $C_1$  (standard 50Ω coax has a capacitance of almost exactly 1 pF/cm, see Appendix H). You can watch the ramp directly, or, if you want to get fancy, add a differentiator as shown.



**Figure 5.39.** Input current (set one of the  $R_s$ 's to zero) and input offset current ( $R_{s1} = R_{s2}$ ) test circuit. Use  $R_s$  values large enough so that the voltage developed across them is at least some 10s of millivolts, so that errors due to offset voltage can be ignored. Add a 200Ω resistor at the output if you want to drive a cable.

A somewhat simpler method, which has worked well for us, is to wire the op-amp as a follower, with a small capacitor from the (+) input to ground (Figure 5.40B); the op-amp's input current then generates an input ramp, faithfully reproduced at the output. At first we struggled with memory effects in mica and film capacitors, but we finally settled on an air (variable) capacitor, of the sort that were used to tune AM radios in the good ol' days. With the capacitor set to 365 pF we got an output ramp of 0.20 mV/s from an LMC6482, and thus an input current of 73 fA. You can "reset" this circuit by collapsing the supply rails. Be sure to put the whole business in a metal box: these open inputs are really sensitive beasts!



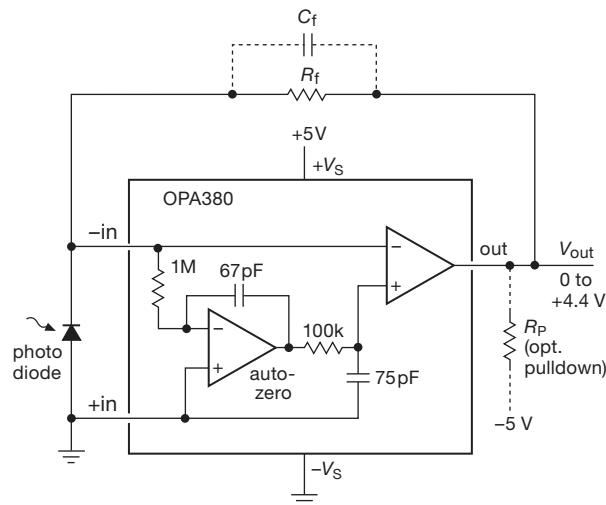
**Figure 5.40.** Integrate the input current to measure in the picoamp range (and below). A. With a separate MOS-input op-amp integrator whose input current is in the low femtoamps (e.g., an LMP7721,  $I_B=3\text{fA}$  typ,  $20\text{fA}$  max). Use electromechanical relays (not MOS switches) for  $S_1$ , for example the COTO 9202 series shown in Figure 5.3. B. More simply, let the device's input current charge a small capacitor, and observe the ramp at the  $G=1$  output.

### 5.10.8 Current noise

The op-amp's input current noise density  $i_n$  flows through the source impedance seen at the amplifier's input terminals, contributing an equivalent noise voltage density  $i_n Z_s$ ; this is often negligible compared with the amplifier's  $e_n$ . We can define a "noise impedance" for the op-amp,  $Z_n \equiv e_n / i_n$ , so that we can safely ignore current noise when the source impedance  $Z_s \ll Z_n$ .

Typical  $i_n$  values range from  $0.1\text{fA}/\sqrt{\text{Hz}}$  to  $50\text{fA}/\sqrt{\text{Hz}}$  for CMOS and JFET op-amps, and up to  $5\text{pA}/\sqrt{\text{Hz}}$  for low-noise BJT-input op-amps that operate at relatively high input currents. The superbeta BJT LT1012 (with its low input current) does considerably better, at  $20\text{fA}/\sqrt{\text{Hz}}$ . But note that the high- $i_n$  LT1028 is the *voltage*-noise winner, for which we pay the penalty of high current noise. Its noise impedance  $Z_n = 850\Omega$ , which means that unusually low circuit resistances must be used, say  $300\Omega$  or less, to obtain the full benefit from its low voltage noise. By contrast, the BJT LT1013 single-supply op-amp has  $Z_n = 315\text{k}\Omega$ , a comfortably high value.

<sup>39</sup> Based on nice techniques worked out by Paul Grohe and Bob Pease at National Semiconductor. See Paul Rako's article "Measuring Nanoamperes" (EDN, 26 April 2007), and two riffs by Pease from his series in Electronic Design: "What's All This Teflon Stuff, Anyhow?" (14 Feb 1991) and "What's All This Femtoampere Stuff, Anyhow" (2 Sept 1993). Nicely readable versions currently available at <http://electronicdesign.com/test-amp-measurement/whats-all-teflon-stuff-anyhow> and .../whats-all-femtoampere-stuff-anyhow.



**Figure 5.41.** The OPA380 achieves  $4\mu\text{V}$  typical voltage offset using an auto-zero feature, yet it has only  $10\text{fA}/\sqrt{\text{Hz}}$  current noise at  $10\text{kHz}$ . It is ideal for transimpedance applications such as the photodiode preamp shown here.

The manufacturer's spec gives the current-noise density at a high frequency like  $1\text{kHz}$  or  $10\text{kHz}$ , chosen to be well above the current-noise  $1/f$  corner frequency. The current-noise  $1/f$  corner typically comes at much higher frequencies than the voltage-noise  $1/f$  corner. For example, the OPA277 has a  $20\text{Hz}$  voltage-noise corner, but a  $200\text{Hz}$  current-noise corner; for an LT1007 it's  $2\text{Hz}$  and  $120\text{Hz}$ . The differing  $1/f$  corner frequencies mean that we'll get different  $Z_n$  values at low frequencies. Returning to the LT1028, for example, its relatively higher current noise at low frequencies lowers  $Z_n$  to  $212\Omega$  at  $10\text{Hz}$ . This requires that we push down further our circuit resistances for optimum performance, to  $100\Omega$  or less.

At high frequencies the current noise may consist largely of the fundamental (and unavoidable) *shot noise*, the statistical fluctuations of electron flow (eq'n 8.6). For an input bias (or leakage) current  $I_B$  that lower bound is  $i_n = \sqrt{2qI_B}$ ; for a bias current of  $10\text{pA}$  that evaluates to  $i_n=1.8\text{ fA}/\sqrt{\text{Hz}}$  (from which you can conveniently scale up or down by the square root of  $I_B$ ). The typical and maximum bias-current specs vary widely, as we saw; evidently many manufacturers simply list the calculated shot-noise value corresponding to the typical bias-current spec. For example, the LT1013 has  $I_B=12\text{nA}$  typ, from which we can calculate a current-noise density of  $62\text{ fA}/\sqrt{\text{Hz}}$ ; the manufacturer's spec is  $70\text{ fA}/\sqrt{\text{Hz}}$ .

An important exception comes in the case of BJT-input op-amps with bias-current cancellation circuits: that greatly reduces the dc input current, but not the cur-

rent noise. For example, the LT1007, with its quiet  $e_n=2.5\text{nV}/\sqrt{\text{Hz}}$ , has a  $10\text{nA}$  bias-current spec, from which we calculate a shot-noise current of  $56\text{ fA}/\sqrt{\text{Hz}}$ , but the manufacturer's spec is  $i_n=400\text{ fA}/\sqrt{\text{Hz}}$ . That's seven times too high! What's going on? To achieve low-noise voltage they run the input transistors at high collector currents. That creates a high base current, which would be the op-amp's input current if they hadn't used the old base-current-cancellation trick (§4x.10). So the dc bias current is small, but the current noise is large. The ultralow- $e_n$  LT1028 also uses bias-current cancellation, keeping its bias current down to  $25\text{nA}$ , but with a current noise 10 times higher than the calculated shot-noise value. And for the LT6010, whose input current is brought down to just  $20\text{ picoamps}$  (the kind of low currents you see in FET-input op-amps), the current noise is 40 times larger than calculated shot noise.

In other words, it's important to realize that the input noise current of a bias-cancelled op-amp will be considerably larger than you would expect if you were to calculate the shot noise arising from the net (i.e., cancelled) input bias current. Rather, you need to calculate the shot noise from the uncancelled base currents (and then apply a factor of  $\sqrt{2}$  to account for the additional noise in the cancellation current). For example, using the LT6010's value of  $I_B = \pm 20\text{ pA}$  (typ), you'd incorrectly estimate a shot noise current of  $i_n \approx 2.5\text{ fA}/\sqrt{\text{Hz}}$ , whereas the datasheet lists a typical value (at  $1\text{kHz}$ , well above the  $1/f$  corner) of  $100\text{ fA}/\sqrt{\text{Hz}}$ ; similarly for the LT1028 (which specifies  $1000\text{ fA}/\sqrt{\text{Hz}}$ , versus the  $90\text{ fA}/\sqrt{\text{Hz}}$  you would incorrectly estimate from the net  $I_B$ ). Table 5.5 does not tell you if a BJT amplifier employs bias-current cancellation, but there's a convenient bias-cancel column in the low-noise BJT op-amp Table 8.3a on page 522. Op-amps with bias-current cancellation generally do not have rail-to-rail input stages.<sup>40</sup>

A *caution*: some datasheets list greatly optimistic values for  $i_n$ , evidently making exactly this error. For example, the bias-cancelled LT1012's datasheet shows a typical  $i_n$  leveling off to  $6\text{ fA}/\sqrt{\text{Hz}}$  (beyond the  $1/f$  corner), which is what you'd calculate from the specified net (i.e., cancelled) input current of  $\pm 100\text{ pA}$  max, whereas you would expect a value about 10 times larger (assuming the uncancelled

<sup>40</sup> The LT1677 listed in Table 8.3a is an exception. Its datasheet has a graph labelled "Input Bias Current Over the Common Mode Range," showing that the bottom  $1.4\text{V}$  and top  $0.7\text{V}$  of the common-mode range suffer from high bias currents. The op-amp's offset voltage is also degraded in these regions. But, hey, we warned you about that in §5.9.1!

base current is about  $100\times$  larger). We were skeptical of the datasheet's claimed  $i_n$ , so we measured it (along with others that appeared to be similarly in error), and found<sup>41</sup>  $i_n \approx 55\text{ fA}/\sqrt{\text{Hz}}$ . This error shares some of the characteristics of an *epidemic*, having infected also the datasheets of auto-zero op-amps. For example, the exemplary AD8628A (listed in Table 5.6 on page 335) specifies an input noise-current density of  $5\text{ fA}/\sqrt{\text{Hz}}$ ; imagine our surprise when we measured a value 30 times larger. Not to be outdone, the MCP6V06 auto-zero op-amp's specification of  $0.6\text{ fA}/\sqrt{\text{Hz}}$  is rather at odds with its measured  $170\text{ fA}/\sqrt{\text{Hz}}$ . See the discussion in §§5.11.

It's important to note that, in the case of chopper and auto-zero op-amps, the input current noise spec is usually given at a low 10 Hz frequency, because that's below the region of very high switch charge-injection current noise (see §3.4.2E). If you take the trouble to measure an auto-zero's input noise, you'll see something like the plots shown in Figure 5.52. This is an unfortunate situation, with insufficient guidance (and perhaps deliberate obfuscation) from manufacturers. We discuss this further in §5.11.

### 5.10.9 CMRR and PSRR

The common-mode rejection ratio, CMRR, tells how much the input offset voltage  $V_{OS}$  varies with common-mode input voltage. The problem, of course, is that such a change in  $V_{OS}$  masquerades as a change in the input signal voltage.

The CMRR values vary from 70 dB (min) for our favorite LMC6482 (an inexpensive CMOS dual op-amp), up to 130 dB for the precision OPA277. Degradation of CMRR at high frequencies often matters, and there's usually a plot in the op-amp's datasheet (check out a few for yourself, for example these two; we show CMRR plots for other types of op-amp in Figures 5.73 and 5.82). For example, the LMC6482 typical CMRR starts falling upward of 1 kHz, and it's down to 80 dB at 10 kHz. It's interesting that both the OPA277 and the AD8622 (another expensive high performer at dc) degrade to about 80 dB by 10 kHz, joining our jellybean CMOS friend. Other parts do better, such as the LT1007 (114 dB typ at 10 kHz). And to repeat a warning we made elsewhere: the CMRR spec often applies to only a limited common-mode range; read the datasheet carefully.

*Note this universal cure:* a time-honored way to avoid CMRR troubles is to use an inverting configuration.

The power-supply rejection ratio, PSRR (not shown in Table 5.5 (pages 320–321) tells how much  $V_{OS}$  varies with power-supply voltage. Typical dc values are 60–80 dB for the LMC6482, up to 130 dB for the OPA277 (but only 100 dB for the AD8622). Study your datasheets!

Frequently one rail is much worse than the other, especially for ac PSRR, because of the op-amp's compensation capacitor (see Figure 4.43, where  $Q_5$  and  $Q_6$  are referenced to the negative rail). For example, the OPA277 suffers by an extra 25 dB on its negative rail. The ac PSRR is significant in two regions, 100–120 Hz (and harmonics) for power-supply ripple, and at high frequencies for cross-talk from other circuitry.

A common defense against PSRR problems, in sensitive applications such as low-level input stages, is to add an  $RC$  filter in the supply rails.

### 5.10.10 GBW, $f_T$ , slew rate and “m,” and settling time

It's tempting to think one can never have too much GBW (gain-bandwidth product, or  $f_T$ , its original name that we favor, see Figure 5.42). After all, a higher GBW means greater loop gain, and higher loop gain means lower error (gain, phase, distortion). What's more, with higher  $f_T$  we're well on our way to having a faster slew rate, via the formula  $S=0.32mf_T$  as discussed in detail in §4x.9.

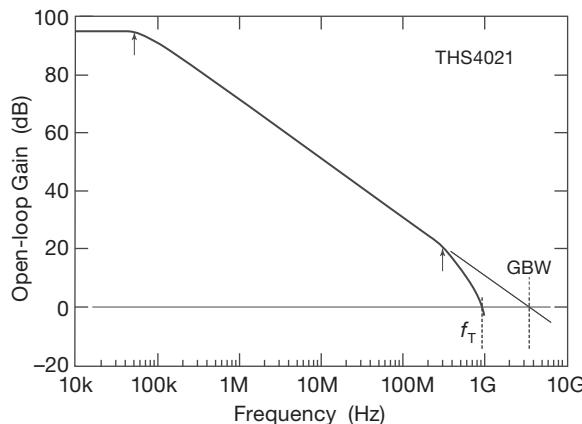
Furthermore, a faster slew rate means a greater full-power bandwidth (FPBW): a sinewave  $V(t)=A\sin\omega t$  has a peak slew rate  $S=\omega A$ , so  $\text{FPBW}=S/\pi V_{pp}$ . Finally, because the first step on the way to waveform settling is the slewing delay  $t=\Delta V/S$ , a higher  $f_T$  is an important step (and often the primary determinant), toward a faster settling-time spec. The data in Tables 5.4 (page 310), and 5.5 (pages 320–321) let you explore the essential question “what price higher bandwidth?”

#### A. An aside: GBW and $f_T$

First, a short riff on “GBW” and “ $f_T$ .” Figure 5.42 shows a plot of open-loop gain versus frequency for a THS4021 wideband op-amp. This is a “decompensated” op-amp, stable for closed-loop gains  $\geq 10$ , with a textbook Bode plot. The term GBW properly describes the product of open-loop gain times frequency in the region in which the gain is dropping at 6 dB/octave (i.e.,  $G_{OL} \propto 1/f$ ). Its extrapolation crosses the  $G_{OL}=0\text{ dB}$  axis at a frequency equal to GBW. At that frequency, however, the gain is less than unity, owing to the effect of additional higher-frequency poles in the amplifier. Strictly speaking, the symbol  $f_T$  is used for the (lower) frequency at which  $G_{OL}=1$ .

<sup>41</sup> Datasheets for the closely similar OP-97 and LT1097 make the same error, evidently corrected in that of the later LT6010 (the recommended successor to the LT1012).

But we *like* the simpler variable  $f_T$ , and so do plenty of other folks; so it's loosely used in place of GBW. Perhaps this is excusable, given that  $f_T$  is pretty close in value to GBW for op-amps that are compensated for stability at unity gain (this describes most op-amps). In any case, unless stated otherwise, we will use  $f_T$  to mean GBW.



**Figure 5.42.** An op-amp's gain–bandwidth product (GBW) is the frequency at which the extrapolation of the open-loop gain curve crosses the unity-gain axis. It's often loosely called “ $f_T$ ,” though the latter is properly the frequency at which the closed-loop gain is unity. The arrows indicate the dominant and second poles. The data here are taken from the THS4021 datasheet, which also shows the phase shift reaching  $180^\circ$  by 400 MHz.

## B. Milking stools and barstools

If offset voltage is one leg of the precision op-amp stool,<sup>42</sup> bandwidth and speed certainly comprise another. Many of the fast op-amps we wanted to include in the precision Table 5.5 (pages 320–321) were passed over because they had too much offset voltage – so some of them have gotten their own Table 5.4 (page 310). For example, in that table we have the LT6200, with 165 MHz GBW,  $0.95 \text{ nV}/\sqrt{\text{Hz}}$ , and 1.0 mV of offset. It's a conventional voltage-feedback (VFB) op-amp offering  $50 \text{ V}/\mu\text{s}$  slewing and 140 ns settling time, so whatever is wrong with it? Only  $50 \text{ V}/\mu\text{s}$ , with  $m = 1$ ? A limited 10 V ( $\pm 5$  V) supply with a high 16.5 mA current? And a *very* high  $40 \mu\text{A}$  (!) max bias current? The point here is that there's a price to pay for high  $f_T$ , and maybe this part is not so attractive after all. But at least the LT6200 has less than  $1 \text{ nV}/\sqrt{\text{Hz}}$  of noise,

1% distortion at 50 MHz, and it's even RRIO to boot.<sup>43</sup> And there's the LT6200-10 variant, with 1.6 GHz of GBW. Nice!

We did include some admirable high-speed op-amps in Table 5.5, (pages 320–321), overlooking in one case an, uh, *underwhelming*  $V_{OS}$ . One such favorite of ours is the OPA656, a member of a small family of highly useful op-amps offered by the Burr–Brown division at TI: it combines 230 MHz GBW,  $290 \text{ V}/\mu\text{s}$  slew rate, and 20 ns settling time. With its 2 pA JFET inputs having less than 3 pF of input capacitance, we readily forgive its 1.8 mV offset voltage, and we're happy enough with its  $7 \text{ nV}/\sqrt{\text{Hz}}$  of input voltage noise. It is excellent for a transimpedance amplifier of the sort used with photodiodes – see §§4x.3 and 4x.9. The OPA656 even has a distortion plot in Figure 5.44 (see next subsection), where we see it has less than 0.1% distortion to 10 MHz and beyond. And it has a 1.6 GHz cousin, the OPA657. When we need higher operating voltage, and crave lower  $e_n$ , we turn to the  $4.5 \text{ nV}/\sqrt{\text{Hz}}$  OPA637; it has somewhat greater capacitance (7 pF), and less bandwidth (80 MHz). And there's the 90 MHz OPA380 shown above. Very fine products from BB/TI.

On the subject of low distortion, Linear Technology offers the LT1468 (90 MHz,  $75 \mu\text{V}$ ,  $\pm 15$  V supply), which claims 0.7 ppm distortion with a 10 V signal; and with its 0.8  $\mu\text{s}$  settling time, it's a good candidate to feed hungry ADCs. Not to be outdone, National Semiconductor offers the LMP7717, a CMOS op-amp with an 88 MHz  $f_T$ , yet drawing just 1 mA, working down to a 1.8 V total supply (!), and offering 1 pA  $I_B$ ,  $6 \text{ nV}/\sqrt{\text{Hz}}$   $e_n$ , and  $150 \mu\text{V}$   $V_{OS}$  input specs, along with rail-to-rail outputs. Parts like these suggest that maybe we can have our speed + precision cake and eat it too.

## 5.10.11 Distortion

Although much of the precision analog-design field concerns dc and low frequencies, there are applications that require accuracy at higher speeds: audio and video, communications, scientific measurement, and so on. With falling op-amp loop gain, input errors are rising, output impedance is rising, and slew-rate limitations may come into play. We need a way to evaluate an op-amp's performance at mid to high frequencies. Some manufacturers help by providing harmonic distortion curves in their datasheets. If

<sup>42</sup> Three-legged, or four? We've found that city folks don't know why it is that milking stools have three legs, whereas barstools have four. Those with rural upbringing can tell you, in a flash.

<sup>43</sup> Yeah, OK, but read carefully: on page 10 of the datasheet you'll see that there are  $\sim 1$  mV shifts of offset voltage when the input is within 1.5 V of the rail! A powerful incentive to use inverting mode!

it's painful going through hundreds of datasheets looking for tabulated specs to compare, it's doubly painful leafing through their back pages looking for distortion curves.

Here we provide some "value added" (hey, this book isn't cheap!) by compiling distortion plots from the datasheets of fifty selected high-performance op-amps: Figure 5.43 (for high-voltage op-amps) and Figure 5.44 for low-voltage and RR-output op-amps (including some HV types). The op-amps listed in Tables 5.4 (page 310) and 5.5 (pages 320–321) have a check in the "dist graph" column if they appear in one of these graphs. We've also measured the distortion and made plots for some popular older op-amps that don't have datasheet plots; see Figure 5.19.

Burr-Brown/TI's OPA134 and OPA627, along with National Semi's LME49990 and other LME49700-series op-amps, are the winners in the high-voltage category. LTC's LT1468 stands out as well. Analog Devices' AD8021 stands out at high frequencies, and they often recommend it for driving ADCs. TI's THS3061, which has an impressive 7000 V/ $\mu$ s slew rate, looks pretty good above 100 kHz, and as a bonus it can deliver 145 mA into 50  $\Omega$ . The OPA1632 and LME49724 are fully differential amplifiers, see §5.17.

The low-voltage and RR op-amp table comprises mostly parts running from  $\pm 5V$  supplies, or lower. Most of these have rail-to-rail outputs, demonstrating that RRO op-amps can compete in the precision arena. Some low-voltage op-amps are at a disadvantage relative to their HV cousins, because they have to use abnormally low signal levels. Consider the graph's OPA1641 JFET winner, at 0.5 ppm. It's a high-voltage part with RR outputs, and it gets tested with 8.5 Vpp signals, a luxury not available to the low-voltage RRO parts. The OPA376 is the low-voltage winner in this category, at 3 ppm, while being tested with 2.8 Vpp. It's interesting that both of these op-amps use the Monticelli output stage (see Figure 5.35 and §4x.11).

### A. Distortion: some caveats

Some caution is advised here: it's tempting to look at the distortion plots and think that you know how the op-amps compare. But some of the distortion measurements need to be taken with more than a grain of salt, and a few caveats are in order. First, there are no de facto standards for op-amp distortion, and manufacturers have chosen different operating conditions.<sup>44</sup> Some use THD, others THD+N (total harmonic distortion plus noise), and still others may

concentrate on specific distortion products, e.g., the 2nd or 3rd harmonic. These hidden choices affect an op-amp's standing on the charts.

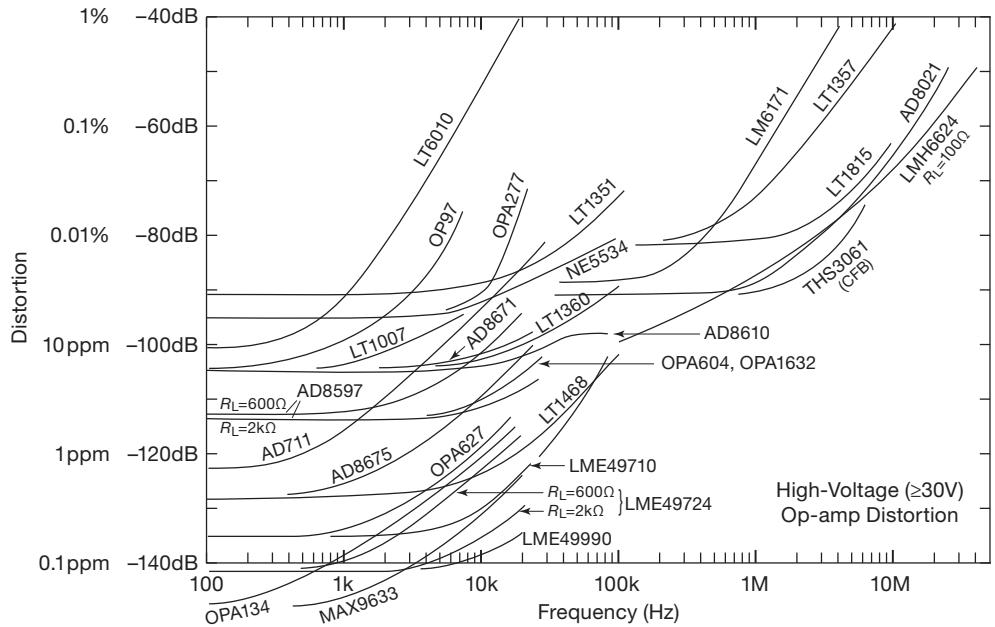
Second, the distortion plots sometimes reveal artifacts of the measurement process – for example, the curves in Figures 5.43 and 5.44 begin with a flat distortion profile starting at dc, which, however, often continues well past the frequency at which we know the op-amp's open-loop gain is falling (i.e., past the primary pole). This is contrary to expectations, and likely reveals a noise-floor instrumental limitation, rather than reality; i.e., the op-amp is better than advertised.

Third, the curves ultimately show the expected rising distortion at higher frequencies; this is due to internal op-amp nonlinearities and to loss of loop gain, both within and outside the op-amp. But this region is heavily dependent on signal size and on load, with varying choices by different manufacturers. At some higher frequency the upward curve may steepen sharply. This is usually due to third-harmonic distortion. We cannot generalize about where second- or third-harmonic distortion will dominate for any given op-amp, but it seems that second-order distortion is the most common culprit. This is surprising, given an effort in many op-amps to fully balance the design.

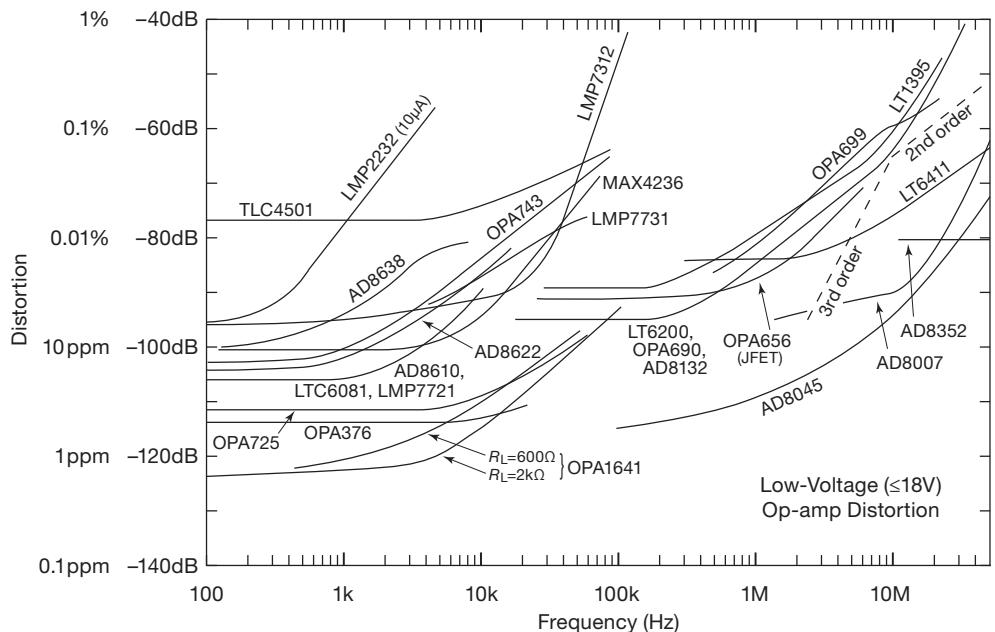
Fourth, when you're in the <10 ppm territory all kinds of strange things can bite you. As wise guru Jim Williams once remarked, "If you think you've measured something to 1 ppm, you're probably wrong." Figure 5.45 illustrates an often-overlooked issue. Here we have a precision OPA1641 op-amp that's capable of distortion performance below 1 ppm at 1 kHz, and in noninverting mode the op-amp has just over 20 ppm of distortion at 100 kHz. It's a JFET op-amp with 8 pF of input capacitance (reasonably low, especially considering the op-amp's low 5 nV/ $\sqrt{\text{Hz}}$  noise spec). In an admirable display of candor, the datasheet warns us that "The *n*-channel JFETs in the FET input stage exhibit a varying input capacitance with applied common-mode input voltage," and it provides plots of increasing distortion arising from an input source resistance driving the op-amp's dynamically-changing capacitance. For example, with  $R_s = 600\Omega$  the 100 kHz distortion increases dramatically, to 100 ppm. They suggest carefully matched input impedances to reduce this type of distortion (an effect that is not confined to this particular op-amp: be warned!). Better yet, use the inverting configuration.

Finally, a look at the test circuit used by many manufacturers to make sub-100 ppm distortion measurements (Figure 5.46). The trick is to reduce op-amp's loop gain by a factor of 100, thus increasing the distortion by the same factor; the reported distortion is then gotten by dividing the

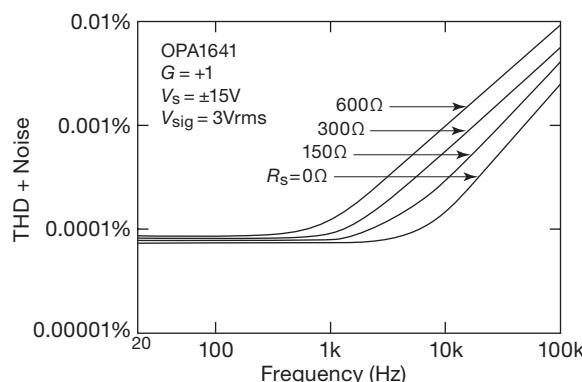
<sup>44</sup> Manufacturers use different voltage levels (2 Vpp, 3 Vrms, 10 V peak, and 20 Vpp), different loads (100  $\Omega$ , 600  $\Omega$ , 2k, 10k and open circuit), different common-mode voltages, different analyzer filters, and even different gains for their measurements.



**Figure 5.43.** Harmonic distortion versus frequency for a selection of “high-voltage” ( $\geq 30V$  total supply) op-amps, from manufacturers’ datasheets.

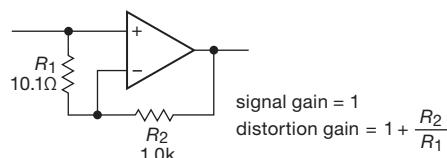


**Figure 5.44.** Harmonic distortion versus frequency for selection of “low-voltage” ( $\leq 18V$  total supply) op-amps, from manufacturers’ datasheets. Most of these have rail-to-rail output stages. See also Figure 5.19.



**Figure 5.45.** The variation of input capacitance with signal voltage causes additional distortion at higher frequencies, dependent on source resistance.

measured distortion by 100. But an immediate concern is the artificial nature of the test, with the op-amp seeing an artificially low source impedance. It's safe to conclude that we all can likely benefit from further work in this area.<sup>45</sup>



**Figure 5.46.** Distortion test circuit. With the values shown, the op-amp's effective loop gain is reduced by  $\times 100$ . Add a series resistor at the output if driving a cable.

### 5.10.12 “Two out of three isn’t bad”: creating a perfect op-amp

We recognized at the outset that there's no such thing as a perfect operational amplifier – but not to fear, there is usually a workaround. If you find the input performance specs you need in one op-amp, and the output specs in another, it may be possible to combine them into a “composite amplifier” circuit that acts as a single op-amp (combining the best features of each) in your feedback loop. Or you can create a composite op-amp by adding a discrete input or output stage to the op-amp IC of your choice. If your overall feedback circuit has very high gain, such as a  $G=10,000$  amplifier, you may not need to worry about compensation (e.g., see Figure 5.61). However, it's not terribly difficult to

deal with the high loop gain of a  $G=1$  circuit, as we show in §4x.5.

In Chapter 4x's discussion of composite amplifiers we show a robust amplifier configuration in which the second op-amp's gain is reduced to unity at a frequency well above the first op-amp's  $f_T$ , allowing loop-gain flexibility. In addition, no restrictions are placed on common-mode voltages or amplifier input connections in this approach. It's a good configuration to consider, although as often as not you'll find plenty of variations in real-world composite amplifier implementations.

### A. Design example: precise high-current piezo positioner

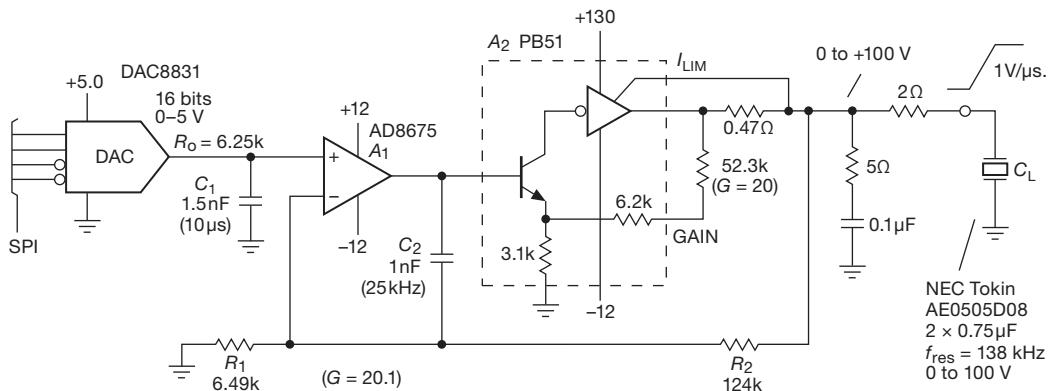
A nice application for a composite amplifier is a precision microscope stage positioner, implemented with a pair of multilayer piezo elements. These devices are both swift and stiff – our chosen part, for example, is good to tens of kilohertz and tens of kilograms – and they provide stable and accurate positioning (at the nanometer scale) over their limited motion (here  $6\ \mu\text{m}$ ). On the down side, they present a difficult load, being highly capacitive (here  $0.75\ \mu\text{F}$ ) and requiring relatively high drive voltage (here  $100\ \text{V}$  full-scale).

A suitable circuit is shown in Figure 5.47. We'd like a moderately fast response, say  $1\ \text{V}/\mu\text{s}$  slewing speed, which requires  $I = CdV/dt = 1.5\ \text{A}$  of drive capability into the piezo pair's  $1.5\ \mu\text{F}$  of capacitance. Our signal source is a 16-bit DAC8831, with a  $+5\ \text{V}$  full-scale output ( $R=2R$  ladder,  $R_{\text{out}}=6.25\ \text{k}\Omega$ ) with a fast SPI interface. Running the DAC with a  $5\ \text{V}$  reference produces a least-significant bit (LSB) step size of  $76\ \mu\text{V}$ . Its  $R_{\text{out}}$  of  $6.25\text{k}$  requires an op-amp whose input current is less than  $12\ \text{nA}$  in order not to add error greater than an LSB. And the op-amp has to swing to  $+100\ \text{V}$  while driving  $1.5\ \mu\text{F}$ .

So, we're evidently looking for a muscular  $150\ \text{V}$ ,  $1.5\text{A}$  op-amp with an input offset less than  $75\ \mu\text{V}$ , and bias current less than  $10\ \text{nA}$ . Looking, looking... no joy! No such part is available, so we'll solve the problem by rigging up a composite amplifier with a gain of 20. We'll specify a frequency response to  $25\ \text{kHz}$  (which is about 20% of the piezo's mechanical self resonant frequency).

For the input op-amp we chose the AD8675 from Table 5.5 (pages 320–321). It has low input errors ( $75\ \mu\text{V}$  and  $2\ \text{nA}$  max), and enough output swing to drive a high-voltage  $G = 20$  stage. For our output amplifier we chose the Apex PB51, a power driver capable of  $300\ \text{V}$  and  $1.5\ \text{A}$  (but subject to a safe-operating-area constraint, for example limited to  $130\ \text{V}$  drop when driving  $2\ \text{A}$  for  $100\ \text{ms}$ ). Its maximum input errors are  $1.75\ \text{V}$  (!) and  $70\ \mu\text{A}$  (!) – that's

<sup>45</sup> OK, we plead guilty-as-charged to that conclusion favored by all academics – “needs more study (and a grant proposal is in the mail).”



**Figure 5.47.** Precision composite amplifier for driving a  $1.5\ \mu\text{F}$  piezo positioner: output to  $100\text{ V}$  and  $1.5\text{ A}$ , with  $75\ \mu\text{V}$  maximum offset,  $2\text{ nA}$  maximum input current, and response to  $25\text{ kHz}$ . The specified piezo actuator moves  $6\ \mu\text{m}$  per  $100\text{ V}$ .

why they call it a *driver* rather than an op-amp! Its gain is set with an external resistor, here  $52.3\text{ k}\Omega$  (for  $G = 20$ ) to match the desired overall gain.

The single feedback path  $R_1R_2$  sets  $G = 20$  for the composite amplifier. A general scheme for compensating composite amplifiers is suggested in §4x.5. Here we use a different scheme, with  $C_2$  isolating the input op-amp  $A_1$  from the output amplifier  $A_2$  at frequencies above  $25\text{ kHz}$ . That way we don't have to worry about  $A_2$ 's response at high frequencies, where it's struggling with its capacitive load. This is really a “custom” configuration, forced on us by the, uh, *challenging* load; what is has in common with other composite amplifier configurations is the single overall feedback path that determines its gain over most of its operating regime.

To get an understanding of circuit operation, imagine a  $2\text{ V}$  input step from the DAC. This causes a  $2\text{ V}$  step at the output of  $A_1$  (which acts like a follower at high speeds, owing to  $C_2$ ), which is approximately the right signal to tell  $A_2$  to drive current into  $C_L$  and move its output by  $40\text{ V}$  (that's why we chose  $G=20$  for the output stage). As amplifier  $A_2$  approaches this goal, op-amp  $A_1$  takes charge and presents correction signals to cause the output to hone in on the precise value.

A few additional circuit details: the DAC's response is slowed by  $C_1$  to a  $10\ \mu\text{s}$  time constant – no point in jolting the amplifier, given its limited bandwidth. The series  $RC$  across the output promotes stability, both by reducing  $A_2$ 's open-loop gain at high frequencies and by providing lossy damping; it's widely used in audio power amplifiers. Note that substantial bypass capacitors (as much as  $100\ \mu\text{F}$ , not shown) are needed with such large signal currents.

Further examples of the composite amplifier technique

can be found in §4x.5 and in Figures 5.58, 5.59, 5.61, 8.49, 8.50, 8.78, 8.80, and 13.48.

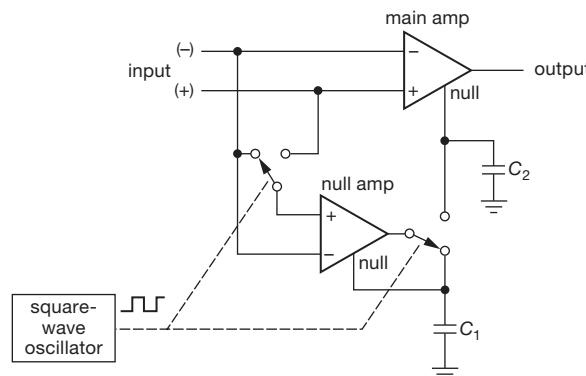
And other examples of techniques that exploit the “two-out-of-three” concept of adding external amplifier blocks are (a) discrete JFET front-end for a BJT op-amp (e.g., Figure 5.58), (b) output unity-gain buffer (§5.8.4), and (c) bootstrapped power supplies to extend voltage range, or to improve CMRR (e.g., Figure 5.79).<sup>46</sup>

## 5.11 Auto-zeroing (chopper-stabilized) amplifiers

Even the best of precision low-offset op-amps cannot match the stunning  $V_{os}$  performance of the so-called “chopper-stabilized” or “auto-zero” (also called “zero-drift”) op-amps. Ironically, these interesting amplifiers are built with CMOS, otherwise famous for its mediocrity when it comes to offset voltage or drift. The trick here is to put a second *nulling* op-amp on the chip, along with some MOS analog switches and offset-error storage capacitors. One of several possible configurations is shown in Figure 5.48. The main op-amp functions as a conventional (imperfect) amplifier. The nulling op-amp's job is to monitor the input offset of the main amplifier, adjusting a slow correction signal as needed in an attempt to bring the input offset exactly to zero. Because the nulling amplifier has an offset error of its own, there is an alternating cycle of operation in which the nulling amplifier corrects its own offset

<sup>46</sup> If you're considering piezo positioners for a precision application, be aware that they exhibit some nonlinearity and hysteresis when driven from a voltage source. These issues are said to be ameliorated when the drive signal is quantified by *charge* instead of voltage. See Chapter 3x for a precision current-drive circuit that circumvents this problem and makes fast linear piezo steps.

voltage. Both amplifiers have a third “nulling” input terminal, analogous to the offset trim seen in some op-amps.



**Figure 5.48.** The original ICL7650 and ICL7652 auto-zero (“chopper-stabilized”) op-amps. Capacitors  $C_1$  and  $C_2$  are external.

The auto-zeroing cycle goes like this. (a) Disconnect the nulling amplifier from the input, short its inputs together, and connect its output to  $C_1$ , the holding capacitor for its correction signal; the nulling amplifier now has zero offset. (b) Now connect the nulling amplifier across the input and connect its output to  $C_2$ , the holding capacitor for the main amplifier’s correction signal; the main amplifier now has zero offset (assuming that the nulling amplifier has not drifted). The MOS analog switches are controlled by an on-board oscillator, typically running in the range of 1–50 kHz.

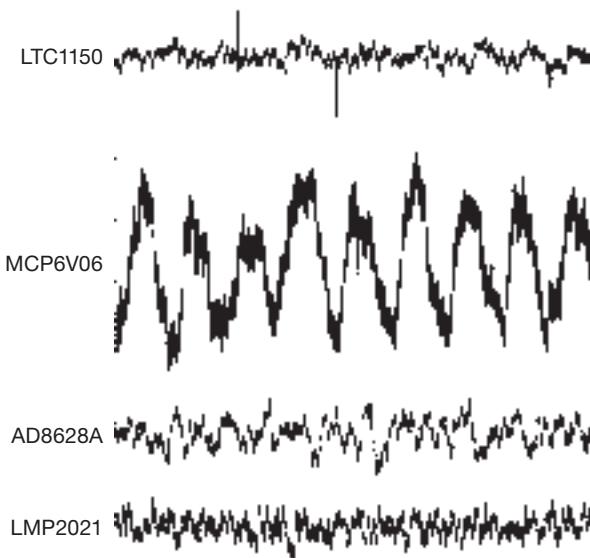
### 5.11.1 Auto-zero op-amp properties

Auto-zero op-amps do best what they are optimized for, namely delivering  $V_{OS}$  values (and tempcos) 5–50 times better than the best precision bipolar op-amp (see Table 5.6 on the next page). What’s more, they do this while delivering full op-amp speed and bandwidth.<sup>47</sup> They also have extraordinarily high open-loop gain at low frequencies (typically 130–150 dB, a consequence of their “composite amplifier architecture”); and, happily, they are inexpensive, particularly when compared with conventional precision op-amps.

That’s the good news. The bad news is that auto-zero amplifiers have a number of diseases that you must watch out for. Being CMOS devices, most of them have a severely limited supply voltage – often 6 V total supply, with a

smaller selection that can run up to 15 V, and just one<sup>48</sup> older generation part (the LTC1150) that can operate at  $\pm 15$  V.

Of greater importance is the problem of clock-induced noise. This is caused by charge coupling from the MOS switches (see §3.4.2E) and can cause wicked spikes at the output. The specifications are often misleading here, because it is conventional to quote input-referred noise with  $R_S = 100\Omega$  and also to give the specification for only very low frequencies. For example, a typical input-referred noise voltage might be  $0.3 \mu\text{Vpp}$  (dc to 1 Hz, with  $R_S = 100\Omega$ ). However, with zero input signal the output waveform might consist of a train of 5  $\mu\text{s}$ -wide 10 mV spikes of alternating polarity!



**Figure 5.49.** Output waveforms from four auto-zero op-amps, configured for  $G = 100$ , with the input connected to ground through a  $100\Omega$  resistor. Vertical: 2 mV/div; horizontal: 100  $\mu\text{s}$ /div.

The internal switching also causes spikes of *input* current, which means that input signals of high source impedance  $R_S$  will exhibit larger input-referred spikes. Figures 5.49 and 5.50 show this behavior, measured with  $R_S$  of  $100\Omega$  and  $1\text{M}\Omega$ , in several auto-zero op-amps configured

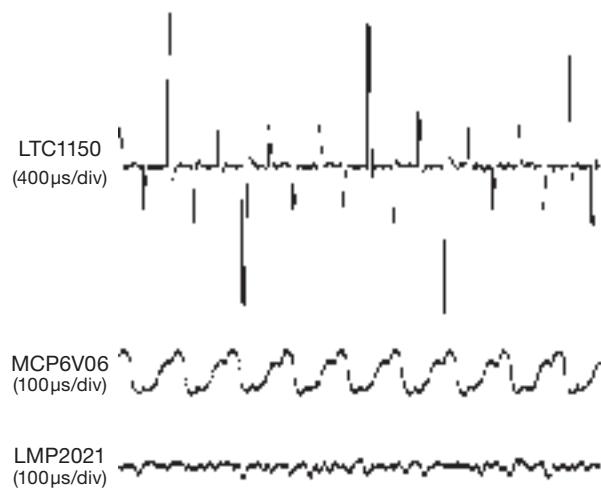
<sup>47</sup> Unlike an earlier generation of synchronous amplifiers that were also called “chopper amplifiers,” but that had bandwidth limited to a fraction of the chopping clock frequency.

<sup>48</sup> Some old high-voltage types that require external (correction-signal) capacitors may still be available.

Table 5.6 Chopper and Auto-zero Op-amps

Part #	Supply (V)	# per Pkg	$I_{bias}$ <sup>a</sup> @25°C		Offset Voltage		CMRR		Noise $\tau_u$		Swing to Supplies?		Cost qty 25 (\$/US)	Comments							
			$I_Q$ typ (pA)	$I_Q$ max (pA)	typ (µV)	max (µV)	Tempco (nV/°C)	min (nV/°C)	$V_{npp}$ dc (µV)	$V_{npp}$ min (µV)	$i_n$ 1kHz (FA/ $\sqrt{Hz}$ )	HF Noise (>kHz) (V/ $\mu$ s)	Slew typ (V/ $\mu$ s)	typ (µs)	IN OUT						
ADA4051-1	1.2	1.8-6	15	20	70	2	15	20	100	110	2	95	100	40	0.13	0.04	110	120	• • •	1.86 auto-zero, -2=dual	
OPA333	1.2	1.8-7	17	20	200	2	10	20	50	106	1.1	55	100	>20	0.35	0.16	45	80	• • •	2.57 auto-cal, dual=2333	
ISL28133	1.2	1.65-6.5	18	30	300	2	8	20	75	118	1.1	65	79	8	0.4	0.1	35	?	• • •	1.69 dual=28233	
OPA330	1.2	1.8-7	21	200	500	8	50	20	250	100	1.1	55	100	>20	0.35	0.16	45	80	• • •	2.56 auto-cal, dual=2330	
MAX9617	1.2	1.6-6	59	10	140	0.8	10	5	120	122	0.42	42	100	50	1.5	0.7	-	? • •	• • •	1.60 RRO, charge pump	
OPA378	1.2	2.2-6	125	150	550	20	50	100	250	100	0.4	20	200	15	0.9	0.4	9	4	c • •	2.16 auto-cal, '2378=dual	
LTC2054	1.2	2.7-6	140	1	150	0.5	3	20	30	120	1.6	85	-	0.5	0.5	-	4000	-	• • •	2.25 dual=2055	
AD8538	1.2	2.7-6	150	15	25	5	13	30	100	115	1.2	52 <sup>b</sup>	-	1.15	0.43	0.4	5	50	• • •	1.80 self-calibrating, duals	
MGP6106	1.2	1.8-6.5	200	1	-	3	-	-	50	140	1.7	82	(0.6)	1	1.3	0.5	300	100	• • •	1.42 low-cost, dual = 6V/0.7	
LTC1049	1	5-16	200	15	50	2	10	20	-	110	3	100	(2)	10	0.8	-	6000	-	• • •	2.85 miniDIP pkg avail	
OPA335	1.2	1.8-7	285	70	200	1	5	20	50	110	1.4	55	20	10	2	1.6	6	50	• • •	2.56 auto-cal, dual=2335	
MAX4238	1	2.7-6	600	1	-	0.1	2	10	-	120	1.2	30	-	?	1	0.35	1000	3300	• • •	1.62 4239=decomp, G=5	
OPA734	1.2	2.7-13	600	100	200	1	5	10	50	115	2.5	135	40	17	2	1	-	8	-	• • •	2.56 auto-cal, dual=2734
AD8551	1.2,4	2.7-6	700	10	50	1	5	5	40	120	1	42	(2)	3.5	1.5	0.4	50	50	• • •	2.34 auto-zero	
AD8572	1.2,4	2.7-6	700	10	50	1	5	5	40	120	1.3	51	(2)	2.3	1.5	0.4	-	50	• • •	3.34 auto-zero	
LTC2050HV	1	2.7-11	750	25	75	0.5	3	-	30	115	1.5	22	-	?	3	2	-	2000	-	• • •	3.20 OK for ±5V supplies
AD8628	1.2,4	2.7-6	850	30	100	1	5	2	20	120	0.5	22	(5)	15	2.5	1	-	50	• • •	1.92 self-cal, dual, quads	
LMP2011	1,2,4	2.7-5.8	950	3	-	0.12	25	15	-	95	0.85	35 <sup>b</sup>	-	25	3	4	-	50	• • •	2.55 copper leadframe	
TLC4501A	1,2	4-6	1000	1	60	10	40	1000	-	90	1.5	12	0.6	none	4.7	2.5	-	2	0	• • •	2.88 self-zero at powerup
AD8638	1,2	5-16	1000	1.5	40	3	9	10,40	60	118	1.2	60	-	8	1.35	2.5	3	50	• • •	2.29 self-calibrating, duals	
LTC1050	1	4.75-16	1000	10	30	0.5	5	10	50	114	1.6	(1.8)	2.5	4	-	3000	-	• • •	2.85 favorite		
LMP2021	1,2	2.5-5.8	1100	25	100	0.4	5	49	20	105	0.2	11	350	30	5	2.5	-	50	• • •	3.38 EMI-rej, '2022=dual	
MAX4236Ab	1	2.4-6	350	1	500	5	20	600	2000	84	0.2	14	0.6	none	1.7	0.3	1	0	• • •	1.78 comp low $V_{os}$ CMOS	
AD8616B <sup>c</sup>	1,2,4	2.7-6	1.7	0.2	1	23	60 <sup>f</sup>	1500	4000	80	2.4	7	2	none	24	12	0.5	0	• • •	1.52 comp CMOS	
LTI028B <sup>d</sup>	1	8-44	7400	25nA	95nA	10	40	200	800	108	0.04	0.9	4700	none	75	15	-	0	-	-	6.31 comp low-noise BJT
<u>high-voltage</u>																					
LTC1150	1	4.8-32	800	10	60	0.5	10	50	110	1.8	-	0.5	1.8	1.5	-	20ms	-	• n n	5.84 OK for ±5V supplies		
LTI012A <sup>e</sup>	1	2.4-40	380	25	100	8	25	200	600	114	0.5	14	20	none	0.2	-	0	-	-	5.11 comp low B bipolar	
LTI007A <sup>e</sup>	1	7-44	3000	10nA	35nA	10	25	200	1000	117	0.06	2.5	1500	none	8	1.7	-	0	-	-	5.83 comp low en bipolar

Notes: (a) check datasheets for plots of  $|I_{bias}|$  vs common-mode input voltage. (b) conventional (not auto-zero) precision op-amps, for comparison. (c) crossover region. (d) 0.01Hz to 10Hz. (f) at  $V_{CM}=0.5V$  and 3.0V (with  $V_S=5.5V$ ), but as much as 400µV near  $V_S=2.5V$ . (g) 1mV/°C at  $V_S=2.5V$ . (h) 150mV/V/Hz hump above 2kHz. (i) near to rails w/o load. (j) typical. (k) current noise values indicated with light italics should not be relied upon; measured values appear to be greater by factors of 100 or more, see discussion in Chapter 8.



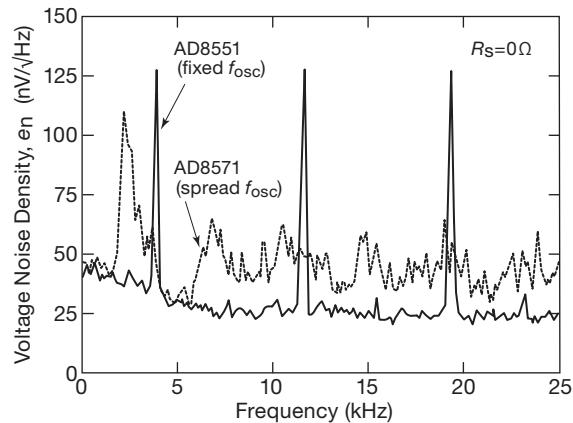
**Figure 5.50.** Output waveforms from three auto-zero op-amps, configured for  $G = 100$ , with the input connected to ground through a  $1\text{ M}\Omega$  resistor. Vertical: 100 mV/div.

for a voltage gain of 100.<sup>49</sup> There is considerable variation among these parts, with the conventional auto-zero configuration (Figure 5.48, used in the LTC1150 and MCP6V06) exhibiting greater clock feedthrough compared with that of alternative designs (as in the AD8628A and LMP2021) that are intended to reduce these undesirable effects.<sup>50</sup>

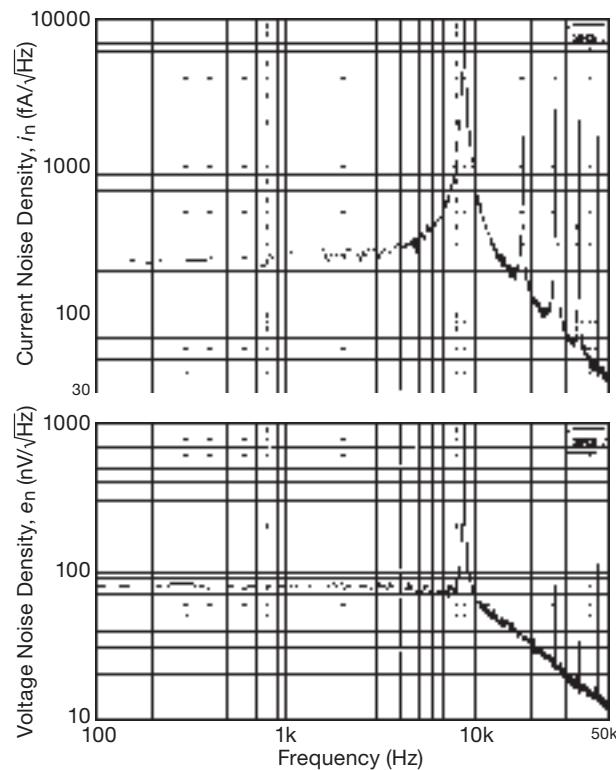
The datasheets do reveal this unseemly behavior, indirectly, in plots of voltage noise versus frequency.<sup>51</sup> Figure 5.51 shows a pair of such plots for two auto-zero products from Analog Devices: the AD8551 has a  $\sim 4\text{ kHz}$  fixed-frequency oscillator, whereas their AD8571 has a deliberately variable (spread-spectrum) oscillator to eliminate sharp spectral lines (which can create undesirable intermodulation with nearby signal frequencies). Note, by

the way, that these plots specify an input signal of *zero* source impedance.

It's always instructive to make some actual measure-



**Figure 5.51.** Spectra of noise voltage, adapted from their datasheets, for a pair of auto-zero op-amps. The AD8571 varies its oscillator frequency in order to suppress sharp spectral features.

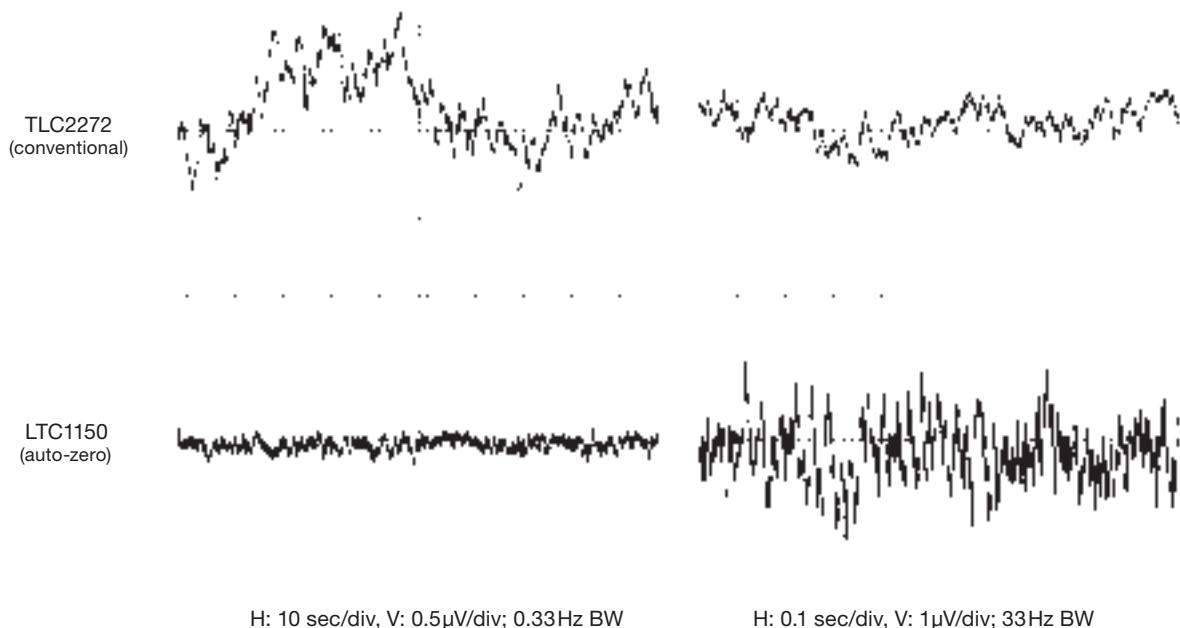


**Figure 5.52.** Measured voltage-noise density (bottom) and current-noise density (top) for an MCP6V06 auto-zero amplifier. Switch-induced clock noise at 9 kHz (and harmonics) is prominent.

<sup>49</sup> The waveforms in the latter show 8 nApp current spikes for the LTC1150, 1 nApp noise for the MCP6V06 (despite its impressive spec:  $0.6\text{ fA}/\sqrt{\text{Hz}}$  at 10 Hz), and 0.2 nApp “rumble” for the LMP2021 (which sports a  $0.35\text{ pA}/\sqrt{\text{Hz}}$  current-noise spec).

<sup>50</sup> From the AD8628A datasheet: “The AD8628/AD8629/AD8630 family uses both auto-zeroing and chopping in a patented ping-pong arrangement to obtain lower low frequency noise together with lower energy at the chopping and auto-zeroing frequencies, maximizing the signal-to-noise ratio for the majority of applications without the need for additional filtering. The relatively high clock frequency of 15 kHz simplifies filter requirements for a wide, useful noise-free bandwidth.”

<sup>51</sup> Watch out, though, for claimed values of noise *current* – the low values listed in many datasheets are completely incorrect, sometimes by factors of  $\times 10$  to  $\times 100$ , evidently having been calculated *a priori* as the shot noise corresponding to the dc input current; see the discussion in §5.10.8 and in §8.9.1F.



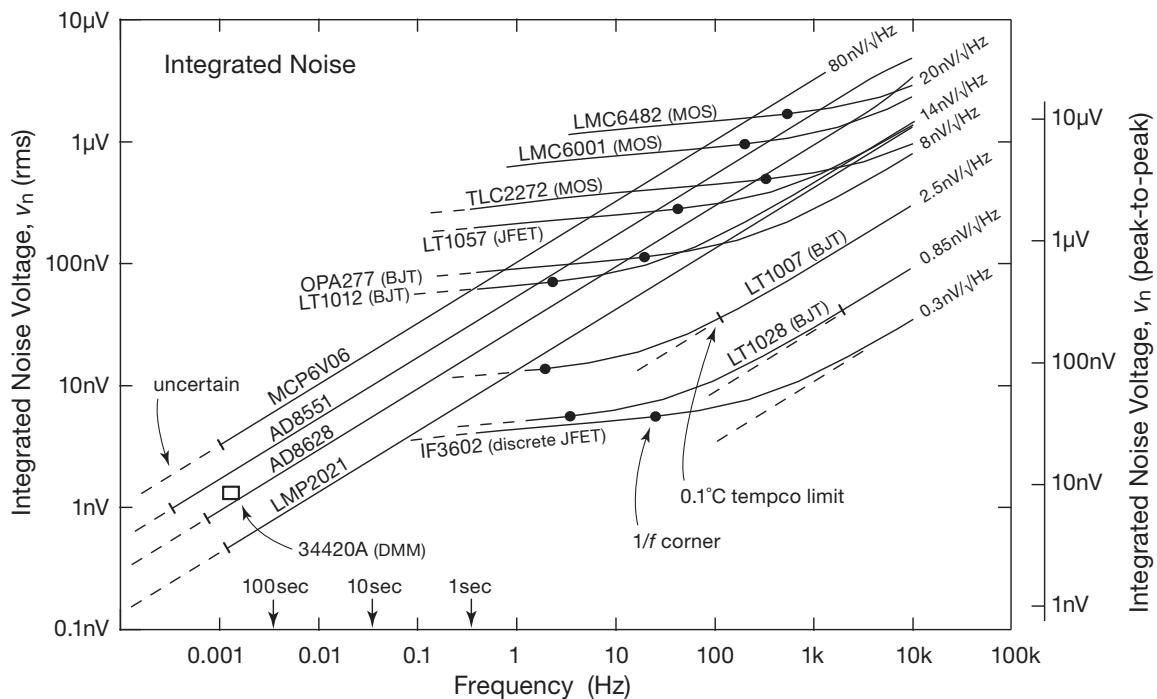
**Figure 5.53.** At very low frequencies a chopper-stabilized op-amp has lower noise than a conventional op-amp, but with 100× greater bandwidth it has more noise, as seen in these measured traces. See also Figure 5.54

ments yourself, if for no other reason than for “fact checking” the manufacturer. We ran some spectral noise plots for a half-dozen auto-zero amplifiers, with a particular interest in chopper-induced narrowband noise at the clock frequency and its harmonics. For these measurements we took data with  $R_s = 0$  (to reveal the input voltage noise  $e_n$ ), and then with  $R_s=1M\Omega$  (to reveal input current noise  $i_n$ ). Figure 5.52 shows the results, for one specimen from our collection of auto-zero amplifiers. The low-frequency measured  $e_n$  agrees well with the datasheet’s value of  $82\text{ nV}/\sqrt{\text{Hz}}$ , but, as noted above, the measured current-noise density  $i_n$  is far greater than the specified value of  $0.6\text{ fA}/\sqrt{\text{Hz}}$  – a factor of ×400 in this case.

For low-frequency applications you can (and should)  $RC$ -filter the output to a bandwidth of a few hundred hertz, which will suppress output spikes. This spiky input-current noise is also of no importance in applications with low input impedances, in integrating applications (e.g., integrating ADCs; see §13.8.3), or in applications in which the output is intrinsically slow (e.g., a thermocouple circuit with a meter at the output). In fact, if you want only very slow output response, and therefore lowpass-filter the output to extremely low frequencies (below 1 Hz), a chopper amplifier will actually have *less* noise than a conventional low-noise op-amp; see Figures 5.53 and 5.54.

Another way to put it is that auto-zero amplifiers have lots of *wideband* voltage noise ( $\sim 50\text{ nV}/\sqrt{\text{Hz}}$  at 1 kHz, compared with just a few  $\text{nV}/\sqrt{\text{Hz}}$  for a good low-noise op-amp), but their noise density holds constant at very low frequencies, as contrasted with the  $\sim 1/f$  (“flicker-noise”) divergence of conventional op-amps (and everything else; see Chapter 8). For example, a conventional low-noise BJT op-amp like the LT1007 has  $e_n = 2.5\text{ nV}/\sqrt{\text{Hz}}$  (typ) at 1 kHz, but its noise power density rises as  $1/f$  below its “corner frequency” of 2 Hz, thus  $e_n \sim 100\text{ nV}/\sqrt{\text{Hz}}$  at 0.001 Hz. Compare that with an auto-zero like the AD8551, with roughly flat  $e_n = 42\text{ nV}/\sqrt{\text{Hz}}$ : the latter will have far smaller fluctuations on time scales of minutes. In fact, the AD8551’s datasheet even specifies a peak-to-peak noise voltage from “0 Hz to 1 Hz” of  $0.32\text{ }\mu\text{V}$  (typ); no conventional op-amp would dare to project its drift out to infinite time!

A final problem with auto-zero amplifiers is their unfortunate overload recovery. What happens is this: the auto-zeroing circuit, in attempting to bring the input difference voltage to zero, implicitly assumes there is overall feedback operating. If the amplifier’s output saturates (or if there is no external circuit to provide feedback), there will be a large differential input voltage, which the nulling amplifier sees as an input offset error; it therefore blindly



**Figure 5.54.** Integrated RMS voltage noise versus amplifier bandwidth. The “zero-drift” aspect of auto-zero amplifiers causes their low-frequency integrated noise voltage  $v_n$  to fall at low frequencies, proportional to the square root of the lowpass bandwidth. In contrast, the rising noise density  $e_n \propto 1/\sqrt{f}$  of conventional op-amps causes a plateau in the integrated noise voltage  $v_n$  below the  $1/f$  corner frequency, as seen in these computed curves. (We chose a lower limit of 0.01 Hz, when integrating the latter because the unbounded integral is divergent. This plot lets you see the region where the auto-zero op-amp wins or loses, compared with the op-amp of your choice. You can draw in estimated plots for other parts if you know  $e_n$  and the  $1/f$  corner frequency. See §8.13.4.) See Figure 5.37 for the spectral noise-density plots behind these curves, and see Figure 8.63 for graphs of three dozen more op-amp types.

generates a large correction voltage that charges up the correction capacitors to a large voltage before the nulling amplifier itself finally saturates. Recovery is slow –  $t_r$  can extend up to several milliseconds. One “cure” is to sense when the output is approaching saturation, and clamp the input to prevent it. You can prevent saturation in chopper amplifiers (and in ordinary op-amps, as well) by bridging the feedback network with a bidirectional zener (two zeners in series), which clamps the output at the zener voltage, rather than letting it limit at the supply rail; this works best in the inverting configuration.

Alternatively, you can do an end run around this problem by choosing a part with fast recovery time, for example the OPA378 or OPA734 (with  $t_r = 4 \mu\text{s}$  and  $8 \mu\text{s}$ , respectively).

### 5.11.2 When to use auto-zero op-amps

- Slow but accurate measurements from transducers: weigh scales, thermocouples, current shunts

- Accurate in-circuit dc conditioning, for example creating precise sets of voltages from a voltage reference
- “Normal-bandwidth” applications that want low-voltage and low- $I_B$  CMOS, can tolerate broadband noise, require low offset voltages ( $\ll 1 \text{ mV}$ ), and don’t want to pay the cost premium of precision CMOS op-amps.

### 5.11.3 Selecting an auto-zero op-amp

Table 5.6 on page 335 lists a nice selection of currently available auto-zero op-amps, plus a few conventional op-amps for comparison. This is a good place to go first when you need an auto-zero amplifier. It’s also a good place to learn about some of the common properties, and some of the quirks, of these amplifiers. Here are some comments, to get you started.

**Supply voltage** All but one of the parts are low voltage, 5.5–6 V max, and many will operate down to 2 V or less. Five can operate from  $\pm 5 \text{ V}$  supplies. The supply

currents range from  $15\ \mu\text{A}$  to  $1.1\ \text{mA}$ . The parts are listed approximately by supply current.

**Input current** Auto-zero amplifiers are built with CMOS, so the input currents are typically in the picoamps. We might expect the input currents to be in the single picoamp territory like other CMOS op-amps. Although there are a few parts for which this is true (e.g., MAX4238, MCP6V06 and LTC2054), most have considerably higher currents, up to  $0.5\ \text{nA}$  max, no doubt because of input-switch charge coupling. Even most conventional JFET op-amps do better except at high temperatures, where auto-zeros are usually much better. For example,  $I_B$  of the auto-zero LMP2021 typically stays below  $75\ \text{pA}$  at  $125^\circ\text{C}$  (for any common-mode voltage), compared with the (conventional) JFET OPA124 ( $0.5\ \text{nA}$ ) and LF412 ( $10\ \text{nA}$ ). Auto-zero input currents are not as low as the best conventional CMOS parts (with their femtoamp currents), but considerably better than conventional precision BJT-input parts like the LT1028 or LT1007 on the table.<sup>52</sup>

**Offset voltage** This is where auto-zero amplifiers really shine, with maximum offset voltages ranging from  $0.1\text{--}5\ \mu\text{V}$  typical (and  $2\ \mu\text{V}$ – $25\ \mu\text{V}$  maximum – in the precision design business you've often got to pay serious attention to "maximum" specifications). A few conventional (not auto-zero) parts can approach this figure ( $20\ \mu\text{V}$  for the CMOS MAX4236A,  $25\ \mu\text{V}$  for the BJT LT1012A and LT1007), but they cannot begin to match the excellent tempco of the auto-zero parts (generally in the range of  $5\text{--}20\ \text{nV}/^\circ\text{C}$ ), gained, of course, by continual offset correction. Conventional op-amps also suffer from the devastating effects of  $1/f$  noise, which sets performance floors in the  $10\text{--}100\ \text{nV}$  region; see Figure 5.54 and associated discussion.

The auto-zero amplifiers have typical offset-voltage tempco drifts from  $4\text{--}100\ \text{nV}/^\circ\text{C}$ . The maximum specs range up to  $250\ \text{nV}/^\circ\text{C}$  (and beyond? many parts don't list a maximum spec). The AD8628 and LMP2021 are the winners in this category. But these parts consume about  $1\ \text{mA}$  and thus can be expected to have more self-

<sup>52</sup> Some parts warn that for high source impedances the bias current may change dramatically as a function of input capacitance! For example, the input current of an LMP2021 with  $R_s = 1\ \text{G}\Omega$  varies from  $-25$  to  $+25\ \text{pA}$  for an input shunt capacitance  $C_s$  ranging from  $2$  to  $500\ \text{pF}$ . Note that such input currents create large offsets with such high source resistances:  $25\ \text{pA}$  into  $1\ \text{G}\Omega$  is  $25\ \text{mV}$ . A graph in the datasheet shows that the input current  $I_B$  goes through zero for  $C_s = 22\ \text{pF}$ . Other manufacturer's parts show similar effects. In a transimpedance amplifier with high  $R_F$ , using a large feedback capacitor  $C_F$  can dramatically reduce the bias-current error.

heating of the die than say the  $60\ \mu\text{A}$  MAX9617 with its  $5\ \text{nV}/^\circ\text{C}$  spec. No manufacturer can afford to perform temperature tests on production parts, so these specs should be taken with a grain of salt.

Is this performance believable? At the level of  $\text{nV}/^\circ\text{C}$  you have to worry seriously about thermocouple effects in external connections, and even within the chip's lead frame itself: typical thermal EMFs are of order  $5\text{--}40\ \mu\text{V}/^\circ\text{C}$  – that's  $1000\times$  (or more) the specified tempco of these auto-zero op-amps!

**Voltage noise** Auto-zero amplifiers exhibit higher broadband noise than conventional op-amps, owing to their CMOS inputs and associated switching elements. The voltage noise density  $e_n$  at  $1\ \text{kHz}$  (the usual benchmark) is of order  $50\text{--}100\ \text{nV}/\sqrt{\text{Hz}}$ , bested by many conventional CMOS parts, and all BJT parts. But, unlike conventional op-amps, the noise density does not rise at low frequencies, so the low-frequency integrated noise voltage (which you can think of as fluctuations, or drift) is better than even the best low-noise op-amps (as seen in Figures 5.53 and 5.54). Speaking approximately, the integrated-noise voltage falls as the  $1/\sqrt{t}$  (or proportional to the square root of lowpass frequency).

In addition to  $e_n$ , a useful parameter in the table is the  $0.1\text{--}10\ \text{Hz}$  peak-peak voltage noise ( $v_n$ ) specification. The AD8628, MAX9617, and OPA378<sup>53</sup> do very well with  $0.5\ \mu\text{V}$  to  $0.4\ \mu\text{Vpp}$  noise specs, but the LMP2021 is the clear winner with an  $e_n$  of  $11\ \text{nV}/\sqrt{\text{Hz}}$  and a  $v_n$  of  $0.26\ \mu\text{Vpp}$ . This part is available in a convenient SOT23 package.

A note of caution, here: the very low-frequency extrapolation (i.e., long-term drift) must eventually become dominated by other drift sources (e.g., diffusion of impurities); see Bob Pease's column, "What's All This Long-Term Stability Stuff, Anyhow?" (published in *Electronic Design*, 20 July 2010).

**Noise current** The noise current density  $i_n$  must be at least the shot noise value (given by  $i_n = \sqrt{2qI_B}$ , with  $q = 1.6 \times 10^{-19}\ \text{C}$ ; thus  $1.8\ \text{fA}/\sqrt{\text{Hz}}$  for a bias current of  $I_B = 10\ \text{pA}$ ) corresponding to the input current  $I_B$ , generally in the range of a few  $\text{fA}/\sqrt{\text{Hz}}$ . In fact, for most of these parts it is much larger – by as much as factors of  $10\text{--}100$ .

The current noise is  $125\times$  higher than shot noise for

<sup>53</sup> A favorite of Phil Hobbs, who gushes "The OPA378 is a really beautiful zero-drift chopper-CAZ mutation that doesn't exhibit switching noise, and has constant  $35\ \text{nV}$  [per root hertz] noise down to DC. I've used it in an etalon-locked diode laser for downhole applications, and it's a thing of great charm."

the LMP2021, which was the winner of the voltage-noise competition. Parts that claim to do well in this regard (i.e., with input noise current approximately equal to the calculated shot noise<sup>54</sup>) include the AD8572, AD8551, and LTC1050. The MCP6V06 is the winner of the current-noise competition, with  $0.6 \text{ fA}/\sqrt{\text{Hz}}$ . This spec predicts  $2 \mu\text{V}$  from current noise through a  $1 \text{ G}\Omega$  resistor in a 10 Hz bandwidth, about equal to the part's  $1.7 \mu\text{V}$  voltage noise  $v_n$ . The TLC4501A does equally well because it does its auto-zero only once, at powerup. But we know that the TLC4501A, unlike the MCP6V06, will do well at higher frequencies and with wider bandwidths because it has no busy auto-zero oscillator and switches. But it will do poorly at long time scales because of  $1/f$  noise and multiple sources of drift.

A caution, loyal reader, as you make your choices – seven otherwise-attractive parts in the table, like the AD8538 and LMP2011, don't have any current-noise specs or plots. You may need to go to the bench to get your answer.

**Slew rate and settling time** For the listed parts the slew rates range from 0.04 to  $2.5 \text{ V}/\mu\text{s}$ , and the gain-bandwidths range from 0.13 to 4.7 MHz. The faster parts are meant to compete for use in ordinary op-amp sockets. For these parts, the settling time  $t_s$  is dominated by slew rate. But there are anomalies, for example the MCP6V06<sup>55</sup> and MAX4238, whose settling times are one or two orders of magnitude longer than the competition. This may be related to recovery time – the parts with recovery times in the milliseconds have very long settling times (the MAX4238), or they aren't willing to say (five other parts).

**Input voltage range** Most auto-zero op-amps do not support input voltages to the positive rail (though they all are rail-to-rail *output*). The MCP6V06, OPA333, ISL28133, MAX9617, and most of the Analog Devices parts are notable for full rail-to-rail input operation, without  $V_{OS}$  or CMRR degradation. The MAX9617 achieves this using an internal above-the-rail charge-pump power supply. Note also that the  $V_{OS}$  specifications may be conditioned on a restricted range of input voltage – most of the way to  $V_+$  for some, others only partway. *Be sure to read the fine print in the specification!* For example, the OPA335 datasheet says “ $(V_-) - 0.1\text{V} < V_{cm} < (V_+) - 1.5\text{V}$ ” next to its 130 dB

CMRR spec, and “ $V_{cm} = V_s/2$ ” next to its  $1 \mu\text{V}$  offset-voltage spec.

**Packages** A few of the (older) Linear Technology types are available in DIP-8 packages for easy breadboarding. Otherwise you can use a SOIC-to-DIP or SOT23-to-DIP adapter (check out offerings from Aries or Bellin Dynamic Systems).

### 5.11.4 Auto-zero miscellany

#### A. ac-coupled “chopper amp”

When considering auto-zeroing chopper amplifiers, be sure you don't confuse this technique with another “chopper” technique, namely the traditional low-bandwidth chopper amplifier in which a small dc signal is converted to ac (“chopped”) at a known frequency, amplified in ac-coupled amplifiers, then finally demodulated by multiplying with the same waveform used to chop the signal initially (Figure 5.55). This scheme is quite different from the full-bandwidth auto-zeroing technique we've been considering, in that it rolls off at signal frequencies approaching the clock frequency, typically just a few hundred hertz. You sometimes see it used in chart recorders and other low-frequency instrumentation.

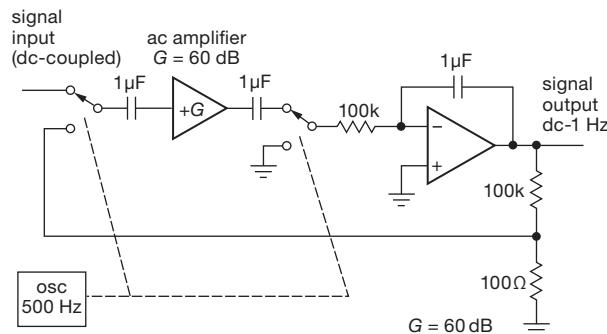


Figure 5.55. An ac-coupled chopper amplifier.

#### B. Thermal offsets

When you build dc amplifiers with submicrovolt offset voltages, you should be aware of *thermal offsets*, which are little thermally driven batteries produced by the junction of dissimilar metals. You get a Seebeck-effect “thermal EMF” when you have a pair of such junctions at different temperatures. In practice you usually have joints between wires with different plating; a thermal gradient, or even a little draft, can easily produce thermal voltages of a few microvolts. Even similar wires from different manufacturers can produce thermal EMFs of  $0.2 \mu\text{V}/^\circ\text{C}$ , 10 to a 100 times the

<sup>54</sup> At 10 Hz, who knows about higher frequencies?!

<sup>55</sup> The curious part number reminds us old timers of a favorite vacuum tube of yesteryear.

typical drift spec of the auto-zero amplifiers in Table 5.6 on page 335! The best approach is to strive for symmetrical wiring and component layouts, and then avoid drafts and gradients.

Here, for approximate guidance when worrying about parasitic thermocouples, Peltier shifts, and the like, are some thermoelectric-pair voltages (from Agilent AN-1389-1):

Copper-to-	Approx $\mu\text{V}/^\circ\text{C}$
Copper	<0.3
Cd-Sn solder	0.2
Sn-Pb solder	5
Gold	0.5
Silver	0.5
Brass	3
Be-Cu	5
Aluminum	5
Kovar	42
Silicon	500
Copper oxide	1000

### C. Power-up self-calibrating op-amps

Texas Instruments has an interesting approach to circumventing clocking noise in auto-zero amplifiers, namely to do it only once! Their TLC4501 family of “Self-calibrating (Self-Cal™) precision CMOS rail-to-rail output operational amplifiers” comes to life at power-up, performing an auto-zero and holding the correcting offset in an on-chip DAC. The good news is that you get no chopping noise and pretty good offset specs ( $10\ \mu\text{V}$  typ,  $40\ \mu\text{V}$  max), at least compared with those of typical CMOS op-amps. The less good news, as you might expect, is that the drift of offset with temperature is unspectacular (you might say “worst in class”), at  $\pm 1000\ \text{nV}/^\circ\text{C}$  typ, compared with  $\sim 20\ \text{nV}/^\circ\text{C}$  for true auto-zero op-amps (see Table 5.6 on page 335).

### D. The non-chopper competition

You can't beat the  $\sim 1\ \mu\text{V}$  (typ) offset specs of these auto-zero and chopper amplifiers, but you can do pretty well with the very best factory-trimmed precision op-amps. Bipolar op-amps like the LT1007 or LT1012 do quite well, at  $10\ \mu\text{V}$  (typ), and the remarkable CMOS MAX4236A has an offset spec of  $5\ \mu\text{V}$  typ (and  $20\ \mu\text{V}$  max). Note, however, that these factory-trimmed op-amps cannot come close to the drift specs of a true auto-zero:  $\pm 200\ \text{nV}/^\circ\text{C}$  (typ) for the bipolar types, and  $\pm 600\ \text{nV}/^\circ\text{C}$  for the CMOS, compared with  $\sim 20\ \text{nV}/^\circ\text{C}$  for true auto-zero.

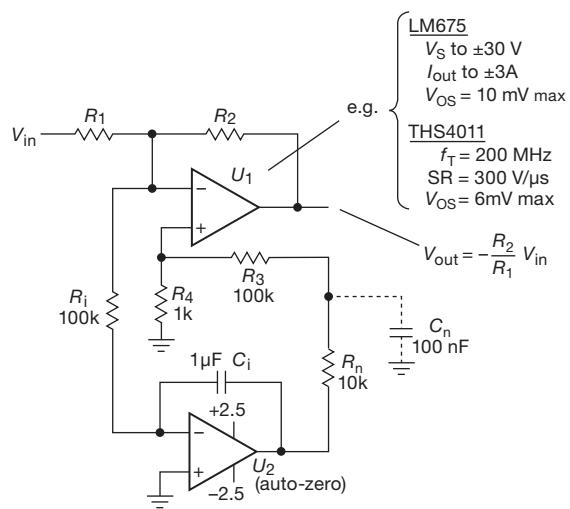


Figure 5.56. External auto-zeroing. But watch out for  $U_1$ 's high bias current.

### E. External auto-zero

You can use an auto-zero op-amp as an external offset trim for a conventional op-amp. This can be handy when you need a high-voltage, high-power, or high-speed op-amp whose input offset is too large. Figure 5.56 shows the scheme, which works most naturally with the inverting configuration, as shown.

The (low-voltage) auto-zero op-amp  $U_2$  is configured as an integrator, looking at the error voltage at the inverting input of the conventional op-amp ( $U_1$ ). The integrator's output is attenuated by  $\times 100$  to trim the voltage at the noninverting accordingly. With the values shown, the integrator's output responds according to  $dV/dt = -\Delta V/R_i C_i$ ; thus a  $10\ \mu\text{V}$  error produces a  $-100\ \mu\text{V}/\text{s}$  integrator output and a  $-1\ \mu\text{V}/\text{s}$  correction at the noninverting terminal of  $U_1$ . A long time constant is both desirable (op-amp offsets drift only very slowly), and necessary (to prevent loop oscillation). Here the correction range is set by divider  $R_3 R_4$ , so that a  $\pm 1\ \text{V}$  integrator output produces a  $\pm 10\ \text{mV}$  correction. The LM675 is a nice high-power op-amp (3 A output current, supply voltage to  $\pm 30\ \text{V}$ , with sophisticated on-chip safe-operating-area and thermal protection), but with a maximum offset voltage of  $\pm 10\ \text{mV}$ . The auto-zero reduces that by a factor of 1000. Similarly, the THS4011 is a fast op-amp ( $f_T = 200\ \text{MHz}$ ,  $\text{SR} = 300\ \text{V}/\mu\text{s}$ ) with a maximum offset voltage of  $\pm 6\ \text{mV}$ . An additional noise filter  $R_n C_n$  at the output of the auto-zero, as shown, may be needed to suppress switching noise in the (slow) correction loop when this technique is used with small signals and low-noise parts like the THS4011 ( $7.5\ \text{nV}/\sqrt{\text{Hz}}$ ). You can think

of this technique as a discrete implementation of the integrated scheme of Figure 5.41.

### F. Auto-zero instrumentation amplifiers

In §5.13 we discuss *instrumentation amplifiers*, which are differential-input amplifiers with very high input impedance ( $10\text{ M}\Omega$ – $10\text{ G}\Omega$ ), wide gain range ( $G_V=1$ –1000, set by internal or external gain-setting resistors), and very high CMRR at higher gains (110–140 dB at  $G_V = 100$ ). These are largely built with conventional (not auto-zero) circuitry; but some are of the CMOS auto-zeroing flavor, with very low offset voltage and drift (to  $10\text{ }\mu\text{V}$  and  $20\text{ nV}/^\circ\text{C}$  max). Table 5.8 (page 363) lists a good selection of instrumentation amplifiers, which include auto-zero types such as the AD8553, AD8230, AD8293, INA333, LTC2053, and MAX4209. Some *conventional* instrumentation amplifiers that compete in this low-drift arena are the LTC1167/8 ( $40\text{ }\mu\text{V}$ ,  $50\text{ nV}/^\circ\text{C}$  max) and the AD8221 ( $25\text{ }\mu\text{V}$ ,  $300\text{ nV}/^\circ\text{C}$  max).

### G. Do it yourself

If you like to get down into the guts of circuits, take a look at the LTC1043 “Precision Instrumentation Switched Capacitor Building Block.” It lets you make your own high-CMRR differential amplifier. That’s just one of its many tricks, which include switched-capacitor filters, oscillators, modulators, lock-in amplifiers, sample-and-hold, frequency-to-voltage conversion, and “flying capacitor” voltage inversion, multiplication, and division. The datasheet makes great bedtime reading.

## 5.12 Designs by the masters: Agilent's accurate DMMs

This is another in the series of featured “Designs by the Masters,” in which we take a close look at some exemplary circuit designs. Think of these as *master classes* in circuit design. You can learn a lot by cracking open a well-designed instrument. A fine example is provided by the excellent digital multimeters from Agilent, specifically their 34401A (6.5-digit) and 34420A (7.5-digit) benchtop meters. In Chapter 13 (§13.8.6) we discuss the precision “multislope ADC” technique they use. Here, in the context of precision *analog* design, we look in some detail at the clever front ends they’ve designed, from schematics help-

fully provided in their service manuals.<sup>56</sup> Let’s see how the real pros do it!

### 5.12.1 It's *impossible*!

At first glance, the task is impossible. Here's why.

**Accuracy** We need accuracy and linearity at the parts-per-million level, in a meter whose full-scale ranges go down to a fraction of a volt (100 mV for the 34401A, 1 mV for the 34420A). That's far down in the nanovolt range.

**Low noise** Accuracy is useless if the instrumental noise makes successive measurements bounce around by many LSBs. So we need input voltage noise levels down at the nanovolt level for the most sensitive ranges.

**High input impedance** A voltmeter should have high input impedance to minimize circuit loading. So, for measurements at the ppm level, you'd like  $R_{in}$  to be something like a million times as large as typical circuit impedances. That puts you in the gigaohm range, with input currents down in the picoamps.

Hence the quandary: gigaohms and picoamps means FETs. Conventional FET op-amps won't deliver this performance, though, owing to relatively large offset voltage, drift, and voltage noise. AZ op-amps (see Table 5.6 on page 335) are considerably more accurate, but they suffer from plentiful current noise. And discrete JFETs (those with large area can have very low voltage noise, less than  $1\text{ nV}/\sqrt{\text{Hz}}$ ) with their uncertain  $I_D$  versus  $V_{GS}$  characteristics (§3.1.5) would seem to be hopeless at the fractional microvolt level. End of discussion.

### 5.12.2 Wrong – it *is* possible!

But it *can* be done. The trick is to realize that a digital instrument (with its on-board microprocessor brain) can calibrate away offsets (with a “zero” measurement) and scale errors (with a “full-scale” measurement), so that what matters is not the presence of offsets *per se*, but their stability (drift) during the measurement time.<sup>57</sup> That allows the use

<sup>56</sup> In the good ol' days, manufacturers proudly published their circuits. Nowadays it's far less common – for example, circuit diagrams are absent from the service manual for the Agilent 34410A (successor to the 34401A). Happily, some manufacturers (e.g., Stanford Research Systems) continue to display their circuit ingenuity, with full schematics and parts lists.

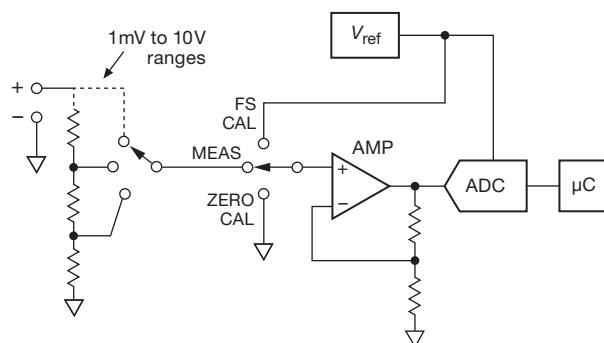
<sup>57</sup> A further trick is to average many such calibrate-measure cycles (up to 2 minutes' worth, in the 34420A) to beat down the scatter.

of discrete dual JFETs, with their unbeatable combination of low  $e_n$  and low  $I_B$ , in a JFET-enhanced op-amp configuration. That way you can get the high loop gain you need for linearity, particularly in the sensitive ranges where the front-end gain is 1000 or 10,000.

That's not the end of the story. You need precise resistor networks with low voltage coefficient, a circuit configuration that maintains its accuracy over a large common-mode input range (to  $\pm 10$  V), and of course a voltage reference whose stability determines the instrument's overall accuracy.

### 5.12.3 Block diagram: a simple plan

These instruments leverage the power of “embedded control” (an on-board microcontroller) to deliver great performance from an architecture of great simplicity. The basic scheme (Figure 5.57) is simplicity itself: it consists of a single amplifier, configured in the familiar noninverting op-amp connection, with floating ground referenced to the (–) input jack. The microcontroller is the boss, here: its code implements the high-accuracy ADC (§13.8.6), and takes care of the multiple on-the-fly calibrations that are needed to wring part-per-million (or better) performance out of a collection of inexpensive parts. Let’s dive into the guts of these two DMMs to see how it all works.

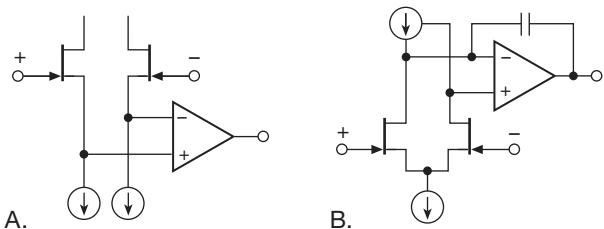


**Figure 5.57.** The Agilent DMMs: extreme simplicity... at the block diagram level.

### 5.12.4 The 34401A 6.5-digit front end

The 34401A made its debut in 1991, surprising the T&M (test and measurement) world with astonishingly good performance (resolution to 6.5 digits, measurements to 1000/s, accuracy to 20 ppm) at an affordable price (~\$1k). The input amplifier (preceded by protection circuitry, and attenuators for the 100 V and 1000 V ranges<sup>58</sup>) provides gains of  $\times 100$  (100 mV range),  $\times 10$  (1 V range), and  $\times 1$  (10 V range) with  $R_{in} > 10\text{ G}\Omega$ ; the input attenuator kicks in for the 100 V and 1000 V ranges, for which  $R_{in} = 10\text{ M}\Omega$ .

The basic structure is a low-noise precision op-amp (an OP-27), driven by a JFET source-follower pair, as shown in Figure 5.58A. (The configuration in Figure 5.58B, where a JFET common-source differential amplifier replaces the follower, is used in the 34420A to provide the additional loop gain and lower voltage noise needed for its more sensitive 1 mV and 10 mV full-scale ranges). The BJT-input op-amp provides lots (120 dB) of stable ( $0.2\text{ }\mu\text{V}/^\circ\text{C}$ ) low-noise ( $3\text{ nV}/\sqrt{\text{Hz}}$ ) gain, but at a price: an unacceptable input current of  $\pm 15\text{ nA}$ , with correspondingly high input current noise ( $1.7\text{ pA}/\sqrt{\text{Hz}}$ ). The JFET follower cures problems of input current and noise, at the expense of offset stability ( $40\text{ }\mu\text{V}/^\circ\text{C}!$ ) and significant added voltage noise ( $10\text{ nV}/\sqrt{\text{Hz}}$ ). The tradeoff sounds bad – but it’s good enough for this instrument. (It’s *not* good enough for the more accurate and sensitive 34420A, as we’ll see presently.)

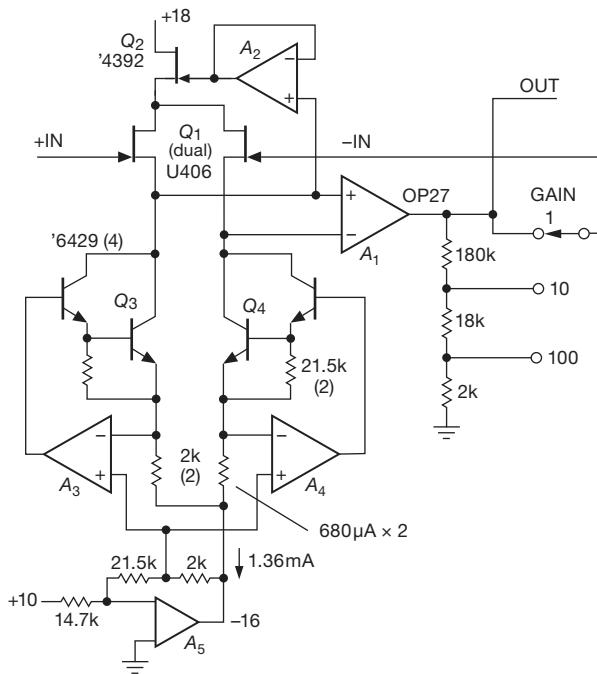


**Figure 5.58.** Basic JFET-enhanced op-amp configurations for the Agilent DMMs: A. source follower, used in the 34401A; B. common-source differential amplifier, used in the 34420A.

The full circuit is shown in Figure 5.59. Note first the bootstrapped drain supply for the JFET pair:  $Q_2$  maintains a constant drain-source voltage across  $Q_1$  (equal to  $Q_2$ 's  $V_{GS}$  at the operating current, the latter held constant by the complicated-looking current sinks on  $Q_1$ 's source terminals). This is essential, because the JFET pair  $Q_1$  is anything but precise (would you believe,  $V_{os(max)}=40\text{ mV}!$ ), and so the variation of this mediocre offset voltage with varying input signal (thus varying  $V_{DS}$ ) would surely torpedo any expectation of accuracy. But by bootstrapping the drains to follow the sources, the transistors don’t even know that there’s any input signal variation; they don’t know, so they can’t mess things up. Furthermore, the low operating voltage (1–2 V) keeps the gate leakage small

<sup>58</sup> All have 20% “overrange,” e.g.,  $\pm 12$  V on the “10 V” range.

and unchanging with input voltage variations over the full  $\pm 15$  V input signal range. Clever!<sup>59</sup>



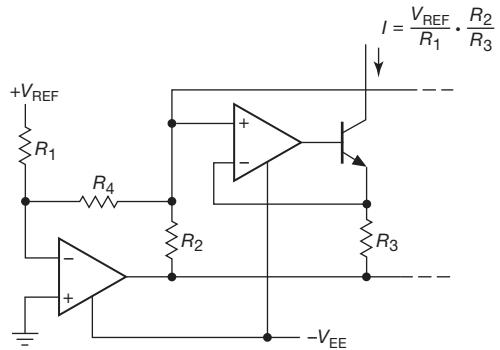
**Figure 5.59.** Agilent 34401A frontend amplifier, capable of measurements with  $0.1 \mu\text{V}$  resolution. The input is single-ended, amplified and measured with respect to the instrument's input common terminal.

The circuit's voltage gain is set accurately by the analog switch and matched resistor network, implemented in a custom gain-switching IC.

The source pull-down circuitry is a current sink pair, based on the stable +10 V reference that is used also for the downstream ADC (see §13.8.6). It's easier to understand in the redrawn form of Figure 5.60, in which only one of the current sink pair is shown and the Darlington is replaced with a single *n*p*n* transistor. The left-hand op-amp generates a voltage across  $R_2$  of  $V_{\text{REF}}R_2/R_1$ ; hence the sink current shown in the figure. In its DMM, Agilent uses a matched network for  $R_2$  and the pair of  $R_3$ 's (one for each source pull-down). The extra resistor  $R_4$  offsets the emitter voltage downward, to  $V_E = -V_{\text{REF}}R_4/R_1$ , to provide the needed compliance for input signals that range over  $\pm 15$  V ( $\pm 12$  V operating range, plus an additional 3 V to accommodate ripple and noise). If you plug in the resistor values

<sup>59</sup> It's necessary that  $Q_1$ 's  $V_{GS}$  at 0.7 mA be less than  $Q_2$ 's  $V_{GS}$  at 1.4 mA, because the difference is  $Q_1$ 's  $V_{DS}$  operating voltage. It's likely that Agilent has an incoming batch inspection to ensure that this condition is met.

from Figure 5.59, you'll find that the compliance extends down to  $-14$  V (emitter is at  $-14.6$  V), and that the individual source pull-down currents are  $680 \mu\text{A}$ . The designers used Darlington transistors to keep the base current error small (roughly  $I_c/4500$ , assuming a transistor beta of 200).



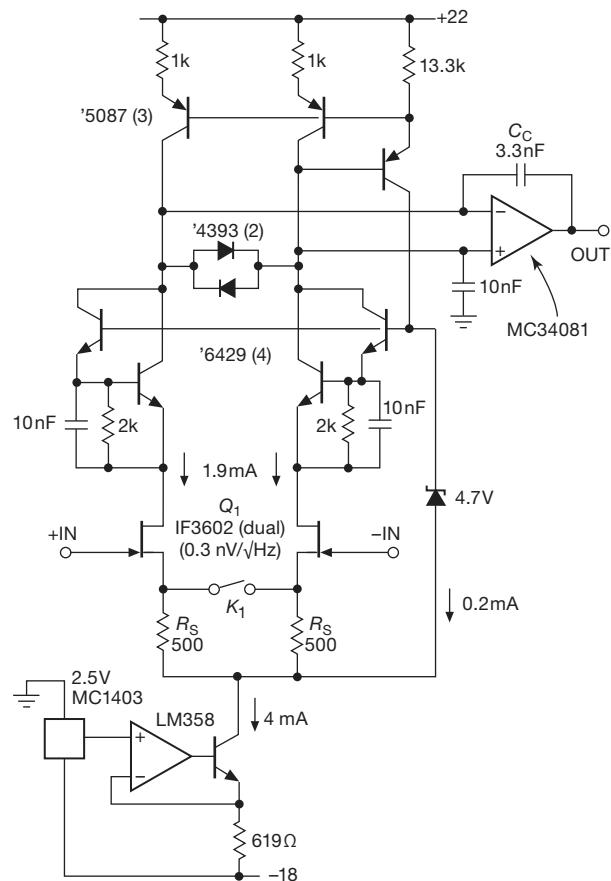
**Figure 5.60.** Agilent 34401A reference-based current sink.

### 5.12.5 The 34420A 7.5-digit frontend

With the 6.5-digit 34401A as a warmup, let's look at its wiser sibling, the 34420A 7.5-digit DMM. It boasts both improved resolution and greater sensitivity (1 mV fullscale), putting real demands on the accuracy, stability, and noise of the front end. On its most sensitive range the frontend amplifier has a gain of 10,000 (to bring the  $\pm 1$  mV input to the  $\pm 10$  V ADC span), requiring lots of open-loop gain to maintain accuracy and linearity. With sensitivity and resolution comes a demand for low noise; for example, the specifications list a "DC Voltage Noise" (with 2-minute averaging) of  $1.5 \text{nV}(\text{rms})$  on the 10 mV range – that's 0.15 ppm.

To meet these demands, the designers used the configuration of Figure 5.58B, in which the JFET pair is configured as a common-source differential amplifier for greater loop gain and reduced noise. The full amplifier circuit is shown in Figure 5.61.

Once again the JFETs are operated at constant current (2 mA each), with bootstrapped drains (held  $V_Z - 2V_{BE} - 1$  V above the source, i.e.,  $V_{DS} \approx 2.5$  V). They chose JFETs with much larger geometry for greatly reduced noise voltage (an impressively low  $e_n = 0.4 \text{nV}/\sqrt{\text{Hz}}$  at 10 Hz). These are *monster* JFETs:  $I_{DSS} = 50 \text{ mA}$  min,  $1000 \text{ mA}$  max (how's that for a parameter spread?!), with an input capacitance of order  $500 \text{ pF}$ , and an unenviable offset voltage spec of  $\pm 100 \text{ mV}$  (with no tempco

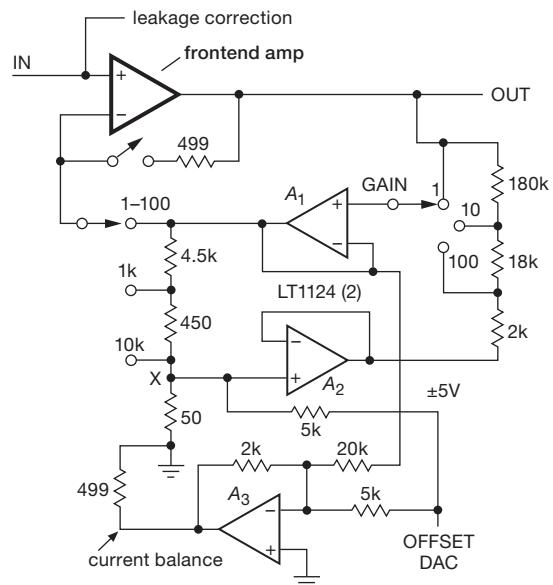


**Figure 5.61.** Agilent 34420A frontend gain block, used for measurements of 0.1 nV resolution with  $G=10,000$  (gain-switching feedback shown in Figure 5.62).

specified). This latter parameter would not appear to bode well for nanovolt measurements! (As we'll see presently, there's provision for continually trimming the measured offset.) These things are brutes, but they sure are quiet. We'll return to this circuit shortly, to deal with issues of gain, bandwidth, and noise. First, though, a look at the overall gain-setting loop.

#### A. Two-stage gain-setting loop

What's shown in Figure 5.61 is the bare amplifier, which sits in the feedback and trimming circuit of Figure 5.62. The gain selection is made with two stages of precision attenuators of high stability (these are custom modules), isolated by precision op-amp  $A_1$  whose offset can be read and cancelled as part of the measurement cycle (by comparing the amplifier's output with the two possible  $G=1$  configurations). The offset of  $A_2$  can also be measured, from the



**Figure 5.62.** Agilent 34420A range-switching feedback circuit, wrapped around the “amplifier” of Figure 5.61. See §13.3.3 and Figure 13.15 for the leakage-correction circuit.

amplifier's output when the offset DAC is set to zero and the gain to  $\times 100$ .

The “offset DAC” is used to dispatch the big elephant in the room (the JFET pair's offset, which can range to  $\sim 50$  mV). This it does by generating a voltage output ( $\pm 5$  V range), which offsets the node marked “X” over a range of  $\pm 50$  mV. Follower  $A_2$  replicates that offset at the bottom of the right-hand gain-setting divider, establishing an effective ground-reference point for the first-stage divider. Here's a subtle point: with a circuit of this precision (in the 1 mV range, the LSB is just 1 nV) you have to worry about stuff you normally ignore – for example the effect of voltage drops through ground path resistance. Here the offset DAC may sink or source up to 1 mA (5 V across 5 k $\Omega$ ), which would push “ground” around by an unacceptable 100 nV if its path had, say, 0.1 m $\Omega$  of resistance. That's why the designers added  $A_3$ , whose output pushes a balancing current of opposite sign into the same ground node; if that current is matched to 1%, the error is reduced to 1 nV.

This raises a related point: when talking nanovolt stability, don't we have to worry about  $A_1$ 's drift? Quite so – but the effect of splitting the gain-setting attenuator into two sections is to reduce that drift by the attenuation of the left-hand divider, i.e.,  $\times 100$  in the most sensitive range.

## B. Care and feeding of the JFETs

An important rule to follow when doing low-distortion precision design using discrete transistors: use a circuit configuration that keeps the transistor's operating conditions ( $V_{ds}$  and  $I_d$ ) unchanged as the input signal changes. Both of these amplifier designs carefully follow this rule, but in different ways, to achieve their good performance. Both designs also operate the JFETs at low drain-source voltages to reduce gate leakage and to minimize self-heating (see §3.2.8).

This same rule is followed (in the second design) with respect to the MC34081 JFET op-amp's input voltages, which are both pinned to  $2V_{BE}+1.9$  V below the +22 V rail. Likewise the mirror transistors see no change in voltage for a signal-input swing from -15 to +15 V. Only the feedback capacitor  $C_c$  sees a change in voltage.

Finally, despite the low  $V_{DG}$  operating voltage for the  $Q_1$  JFETs, there is still an issue of small gate leakage currents to worry about. Agilent has added an input-bias-current correction circuit, with an 8-bit DAC, to solve this problem. We discuss how this interesting circuit works in §13.3.3.

## C. Amplifier gain: $\times 1$ to $\times 10,000$ , stable to 0.1 ppm

Turning back to the amplifier (Figure 5.61), we can understand some nice subtleties. A closed-loop gain of  $\times 10,000$  is needed in the 1 mV range, for which lots of open-loop gain is needed. The JFET differential amplifier provides gain ahead of the op-amp; though the gain is not easily calculated at dc (it depends on the impedance of the current-mirror drain load), we can estimate its gain-bandwidth product  $f_T$  by noting that compensation capacitor  $C_c$  makes the differential gain roll off according to  $G = g_m X_C / 2$ , where  $g_m$  is the transconductance of each JFET at the operating current (relay  $K_1$  is closed for the 1 mV and 10 mV ranges, removing the  $500\Omega$  source degeneration resistors).  $f_T$  is the frequency at which the gain of the composite amplifier has fallen to unity, i.e.,  $f_T = g_m / 4\pi C_c$ . To estimate  $g_m$ , we note that these JFETs are operating well down in the subthreshold region (their  $I_{DSS}$  is typically 300 mA), where FETs behave more like BJTs ( $I_D$  exponential in  $V_{GS}$ ; see Figure 3.15), with their transconductance proportional to drain current and with  $g_m$  only somewhat less than a BJT operating at the same current. For the IF3602 JFET running at  $I_D=2$  mA, then, we can estimate  $g_m \approx 60$  mS (a BJT would have  $g_m = 40I_c$  mS), thus  $f_T = 1.5$  MHz.

Running this backward, we find that the open-loop gain is about  $10^6$  at 1 Hz, as seen in Figure 5.63. Source degeneration is enabled for the low-gain ranges, to maintain stability. The rolloff is easy to calculate, because the differ-

ential transconductance is reduced to  $1/2R_s$  (1 mS), thus  $f_T = 50$  kHz.

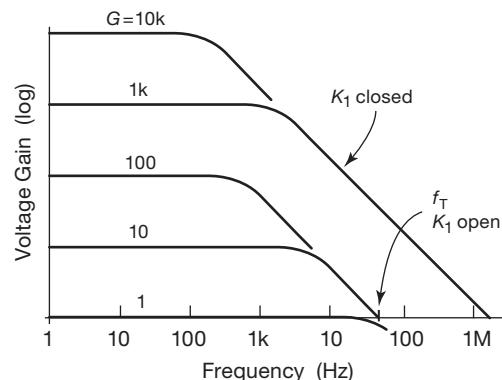


Figure 5.63. Differential gain for the amplifier of Figure 5.61.

## D. Sub-nanovolt amplifier noise

Finally, the important issue of noise. This is a big deal when you're talking nanovolts; it's the reason the designers chose huge-geometry JFETs, in spite of their, uh, less-than-ideal characteristics (offset voltage, input capacitance). Noise matters most in the most sensitive range, where full scale is 1 mV (1.2 mV, to be precise, owing to the 20% overrange), and the 6.5-digit LSB is 1 nV.

There are several noise sources here. The JFETs contribute about 1 nVrms in a 3 Hz band around 1 Hz ( $e_n = 0.4$  nV/ $\sqrt{\text{Hz}}$  each, multiplied by 1.4 for uncorrelated noise). To explore further the measurement fluctuations, take a look at Figure 5.54, where we show the voltage noise of various op-amps, of both conventional and chopper-stabilized varieties. The traces show the integrated rms noise voltage up to a cutoff frequency (x-axis), including the effect of the component's  $1/f$  noise. The IF3602 is the lowest-noise part on the graph. If we assume an integration time of 100 PLC (powerline cycles) or 1.67 s, to achieve 7.5-digit performance, that interval corresponds to a 0.6 Hz cutoff frequency, and about 3 nV rms of noise. If we average 64 such measurements over a two-minute period, we could hope for the rms fluctuations to be reduced by  $8\times$ , to about 0.4 nV. Agilent claims 1.3 nV in their datasheet, evidently allowing for some nonrandom variations and other errors.

Noise in the current sink is of lesser importance because the differential stage cancels it to a high degree; that's a good thing, because this design uses a noisy voltage reference! (the MC1403 is an early bandgap design,

with unspecified noise voltage).<sup>60</sup> The 34420A's digitizing capability drops from 7.5 to 6.5 digits when operating faster than 20 PLC, or 1.5 readings/sec, and further drops to 5.5 digits above 25 rd/s, and 4.5 digits above 250 rd/s, so rising high-frequency amplifier noise wouldn't be noticed.

### E. Going beyond the specifications

When pushing the limits of the possible, you often find that the job cannot be done while respecting worst-case component specifications. Here, for example, the critical JFET transistor pair has a worst-case specified gate leakage current of 500 pA (at 25°C), whereas the instrument specifies a maximum input current of 50 pA. What to do?

If you are a major manufacturer, you can often persuade the supplier to screen parts to a tighter specification. In any case, you can do the job yourself. Be aware, though, that there's generally no guarantee of process continuity, and the availability of better-than-specified parts; worse still, the special parts you need may be discontinued altogether! One possibility, if you're willing to hazard a guess as to an instrument's long-term popularity, is to buy a lifetime supply of a critical part.

## 5.13 Difference, differential, and instrumentation amplifiers: introduction

These terms describe a class of dc-coupled amplifiers that accept a differential signal-input pair (call them  $V_{in+}$  and  $V_{in-}$ ), and output either a single-ended signal or a differential-output pair that is accurately proportional to the difference:  $V_{out} = G_V \Delta V_{in} = G_V (V_{in+} - V_{in-})$ . Their shared claim to fame is high common-mode rejection, combined with excellent accuracy and stability of voltage gain. Here are their distinctive features, as commonly understood among circuit designers.

**Difference amplifier** differential in, single ended out; op-amp plus two matched resistor pairs (Figure 4.9, §4.2.4, and Figure 5.65); CMRR 90–100 dB; accurate but low gain ( $G_V=0.1\text{--}10$ ); input impedance 25–100k, intended

<sup>60</sup> For single-ended amplifiers we would want the current-source noise  $i_n$  to be less than  $e_n(\text{amp})g_m$ . We can use the expression  $e_n(\text{ref})/e_n(\text{amp}) = g_m R_S$  to determine the allowable voltage noise in the current-source reference. For this circuit that ratio is 37, thus only 11 nV/ $\sqrt{\text{Hz}}$  for a noise contribution comparable to that of the 0.3 nV/ $\sqrt{\text{Hz}}$  JFET. The MC1403 reference is about 20× worse than this. Evidently the Agilent engineers are relying on the matched noise currents in the two JFETs to cancel to better than 5%, enforced by the 1% current-mirror resistors. At frequencies above about 10 Hz, however, the 10 nF capacitor defeats this cancellation.

to be driven by a low impedance; inputs typically can go beyond rails.

**Instrumentation amplifier** differential in, single ended out; very high input impedance ( $10\text{ M}\Omega\text{--}10\text{ G}\Omega$ ), wide gain range ( $G_V=1\text{--}1000$ ), and very high CMRR at higher gains (110–140 dB at  $G_V=100$ ); §5.15, e.g., Figure 5.77.

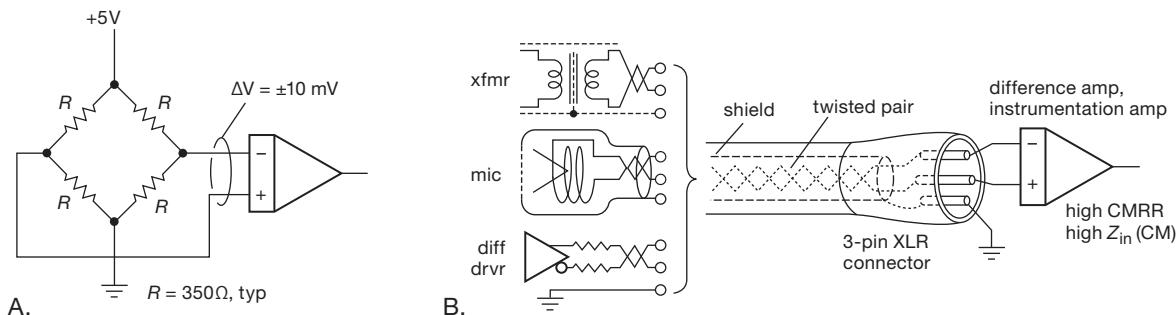
**Differential amplifier** differential or single ended in, differential out; most are low voltage, fast settling, and wideband; ideal for twisted-pair cable drivers and fast differential-input ADCs; §5.17, e.g., Figure 5.95.

An obvious application is the recovery of a signal that is inherently differential, but that rides on some common-mode level or that is afflicted by common-mode interference. Figure 5.64 shows an example of each.

The first example is the strain gauge we saw earlier (§5.4), a bridge arrangement of resistors that converts strain (elongation) of the material to which it is attached into resistance changes; the net result is a small change in differential-output voltage when powered by a fixed dc-bias voltage. The resistors all have approximately the same resistance, typically  $350\Omega$ , but they are subjected to differing strains. The full-scale sensitivity is typically  $\pm 2\text{ mV}$  per volt, so that the full-scale output is  $\pm 10\text{ mV}$  for 5 V dc excitation. This small differential-output voltage (proportional to strain) rides on a +2.5 V dc level. The differential-input amplifier must have extremely good CMRR in order to amplify the millivolt differential signals while rejecting the  $\sim 2.5\text{ V}$  common-mode signal and its variations. For example, suppose you want a maximum error of 0.1% of full scale. That's  $\pm 0.01\text{ mV}$ , riding on 2500 mV, which amounts to a CMRR of 250,000:1, or 108 dB. This overestimates the needed CMRR: in practice you would perform a "zero calibration," so that the CMRR need be adequate only to reject *variations* in the +5 V bridge bias; something like 60 dB would suffice here.<sup>61</sup>

The second example (Figure 5.64B) comes from the world of professional audio, where you encounter some pretty impressive challenges. In a concert recording situation, for example, you may have microphones hanging from a high ceiling, with connecting cables 100 m or more in length. The peak signal levels may be around a volt, dropping to a millivolt during quiet portions of the music. But you've got to keep powerline pickup and

<sup>61</sup> We'll see the strain gauge again, in connection with analog-to-digital converters, in §13.9.11C (Figure 13.67). A similar biased bridge arrangement is used in the platinum resistance temperature detector (RTD), which is the sensor used in the microcontroller-based thermal controller in §15.6.



**Figure 5.64.** Inherently differential signals, for which good common-mode rejection is required. (a) Strain gauge. (b) Audio balanced line pair.

other annoyances (e.g., switching noise from lighting dimmers) another 40 dB below that – the human ear is distressingly sensitive to extraneous sounds. Add up the dBs – we need 100,000:1 suppression of pickup ( $<10 \mu\text{V}$ )! Seems impossible; but recording engineers have been doing this successfully for decades by the simple expedient of transporting audio signals on a balanced differential pair (with a standard signal impedance of  $150 \Omega$  or  $600 \Omega$ ). For this they use a well-shielded twisted-pair cable, terminated in the legendary 3-pin XLR latching connector (which can take plenty of abuse – it's got a tough metal shroud, good strain relief, and so on). And to keep the signal fully balanced they use either a high-quality audio transformer or a well-designed differential-output driver (at the transmit end), and another transformer or a well-designed differential-input amplifier (at the receive end).

Lest we leave a misleading impression, we hasten to note that differential-input amplifiers are helpful also in situations in which the signals themselves are not inherently differential. Two common examples are accurate low-side current sensing (Figure 5.68a), and the use of differential-input amplifiers when sending signals between instruments (Figure 5.67). In the latter, the flexibility provided by a difference amplifier's output REF pin enables us to avoid ground loops while transporting a signal between a pair of instruments whose local grounds are not identical.

The tricks involved in making good instrumentation amplifiers and, more generally, high-gain differential amplifiers are similar to the precision techniques discussed earlier. Bias current, offsets, and CMRR errors are all important. Let's begin by discussing the design of difference amplifiers for less critical applications, working up to the most demanding instrumentation requirements and their circuit solutions.

## 5.14 Difference amplifier

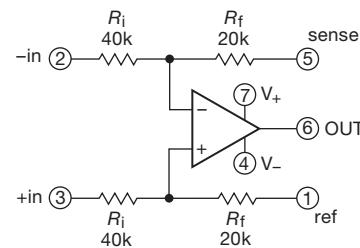
Let's look first at the difference amplifier: its basic operation, some applications, a closer look at its performance parameters, and finally some clever circuit variations.

### 5.14.1 Basic circuit operation

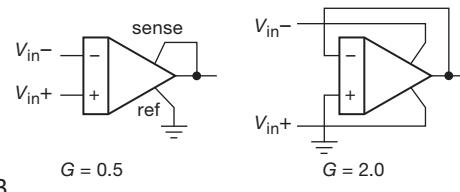
The classic difference amplifier (Figure 5.65) consists of an op-amp with matched resistor pairs  $R_f$  and  $R_i$ , for which the differential gain is

$$G_{\text{diff}} \equiv \frac{V_{\text{out}}}{V_{\text{in+}} - V_{\text{in-}}} = R_f / R_i.$$

Assuming an ideal op-amp, the common-mode rejection

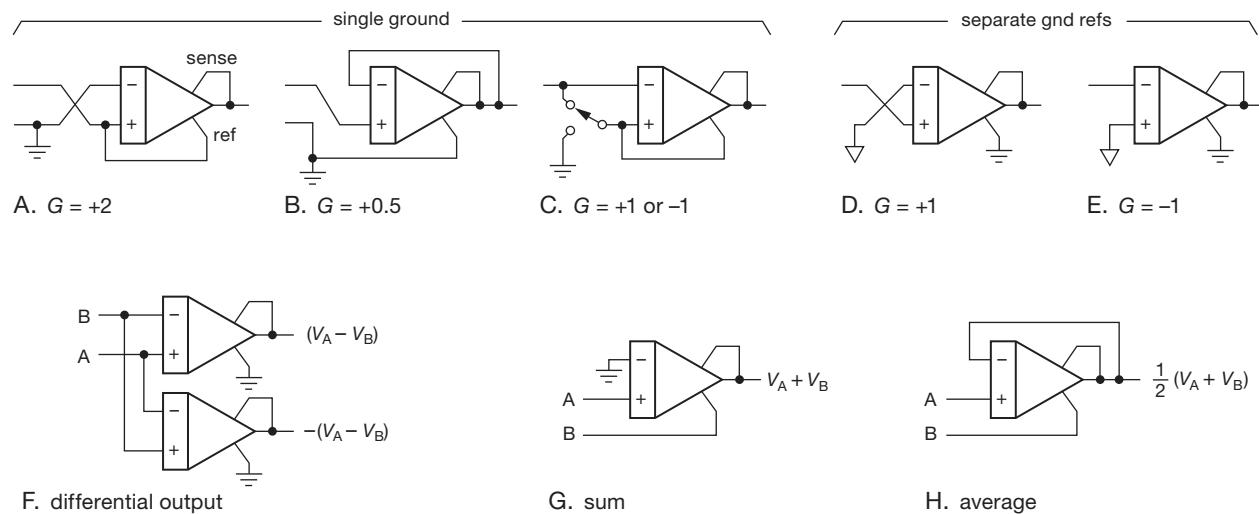


A.



**Figure 5.65.** Classic difference amplifier, with the resistor values used in the AD8278/9 ( $G=0.5$  or 2).

is limited by the matching of the  $R_f/R_i$  ratio in the two paths; with discrete resistors of 1% tolerance, for example,



**Figure 5.66.** The tricks that a unity-gain difference amplifier can play. Note the separate ground symbols for (D) and (E).

you could expect a CMRR of  $\sim 40$  dB at low frequencies (where the op-amp itself has a higher CMRR and where effects of capacitive imbalance are negligible). That's adequate for situations in which only a modest CMRR is needed, for example low-side current sensing. For better performance you could trim one of the resistors or use resistors of tighter tolerance (for example the commonly available 0.1% types, typically \$0.10–\$0.20 in 100-pc quantity, or the inexpensive Susumu RG-series, with 0.05% tolerance and 10 ppm/ $^{\circ}\text{C}$  tempco, about \$1); or, better, use matched resistor pairs (e.g., the not-inexpensive Vishay MPM-series, with ratio tolerances down to 0.01% and ratio tempcos to 2 ppm/ $^{\circ}\text{C}$  or the LT5400 matched quad, with similar ratio tolerance and with ratio tempco of 1 ppm/ $^{\circ}\text{C}$  max).

But don't get carried away... because there are plentiful offerings of complete integrated difference amplifiers of excellent performance, costing a lot less than you'll spend rolling your own. We've listed many of these in Table 5.7 on page 353. In the "normal" configuration the SENSE line is connected to the output and the REF line is connected to circuit ground. But you can run it in reverse, as shown in Figure 5.65B. To give a sense of the performance you get with these integrated difference amplifiers, the worst-case specifications for the AD8278B in the figure are gain accuracy of  $\pm 0.02\%$  (for  $G=0.5$  or  $G=2$ ); gain tempco of 1 ppm/ $^{\circ}\text{C}$ ; offset voltage and tempco of  $100 \mu\text{V}$  and  $1 \mu\text{V}/^{\circ}\text{C}$ ; and a CMRR of 80 dB. It costs about \$3.

## 5.14.2 Some applications

### A. Single-ended input

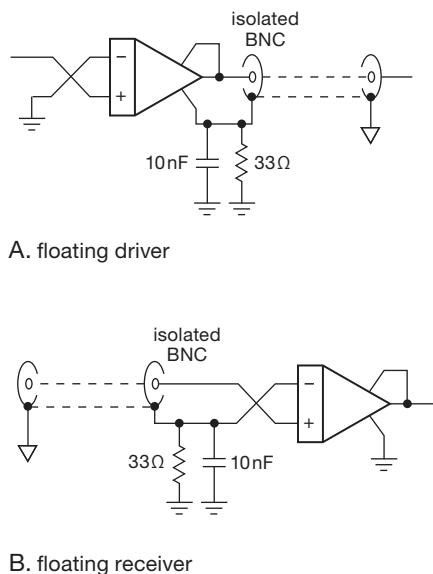
It's perfectly OK to use a difference amplifier with a single-ended input, for example, to get a precise and stable gain. Figure 5.66 shows some simple configurations. Note that differential nature of its inputs is not "wasted" when a difference amplifier is used as in (D) and (E), because its independent REF pin accommodates small differences in ground potential at input and output. Put another way, the output voltage, relative to *its* ground, is precisely  $\pm 1.0$  (in this case) times the input voltage, relative to *its* ground.

### B. Ground-loop isolation

This property is just what you need for the ground-loop isolating application of Figure 5.67. In the first circuit, the driver side permits its output reference to float to the potential of the (non-floating) receive side. The small-value resistor and bypass capacitor allow a small voltage difference when forced by the receive end. Both sides are happy. In the second circuit we've allowed the receive side's input common to float, as forced by the (non-floating) driver. These circuits resolve minor ground-loop problems in single-ended cable connections; but they are no substitute for the fully isolated and/or balanced approach that's needed in demanding applications like professional audio or video.

### C. Current sensing

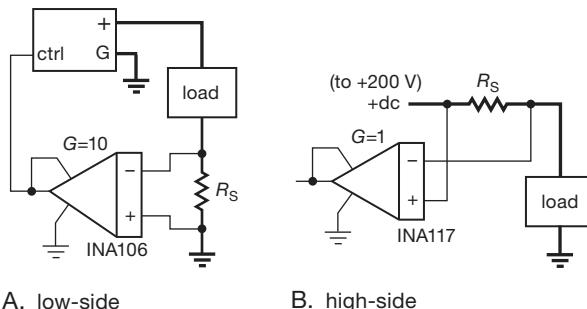
Figure 5.68 shows difference amplifiers used for low-side and high-side current sensing, perhaps as part of a constant-



**Figure 5.67.** Exploiting the REF pin to prevent powerline-frequency ground loops in connected instruments.

current control, or simply for precise load-current monitoring. At first glance it may seem unnecessary to use a difference amplifier for the low-side sensing configuration, because the sense resistor returns to circuit ground. But imagine that we're dealing with lots of power, say up to 10 A load current. We'd use a low-value precision sense resistor, perhaps  $0.01\ \Omega$ , to keep its power dissipation below 1 W. Even though one side of it is connected to ground, it would be unwise to use a single-ended amplifier, because a connection resistance of just a milliohm would contribute 10% error! A differential-input amplifier is the solution, connected as shown to a 4-wire "Kelvin-connected" sense resistor. Note that the difference amplifier need not have particularly good CMRR, because the low side isn't moving far from ground.

The same is not true of the high-side configuration (Figure 5.68B), where the common-mode voltage is far greater than the differential voltage. Here we've specified a unity-gain difference amplifier intended for high-voltage applications, which permits common-mode input voltages to  $\pm 200\text{ V}$  (its internal circuit uses a 20:1 resistive divider at the front end, see §5.14.3). Let's look at the numbers: the dc voltage can range from 0 to 200 V; so the specified minimum CMRR of 86 dB (1:20,000) would interpret that as equivalent to a differential input variation of 10 mV. Is that bad? You bet! To maintain 1% accuracy of sensed current we'd need to size  $R_s$  to drop 1 V at full load current. That's a lot of "voltage burden," and it's a whopping 10 W



**Figure 5.68.** Current sensing for measurement or control. Low-side sensing (A) is forgiving of CMRR, unlike high-side sensing (B), for which the difference amplifier shown would introduce significant error in a high-voltage and high-current supply (where only a small drop across  $R_s$  can be tolerated).

of dissipated power, if this were a powerful 10 A dc supply as in the other example. So this scheme is adequate for a low-current 200 V supply; but there are better ways to do high-side sensing, for example by floating the amplifier and relaying the output down to earth as a current or (via an opto-isolator) as a digitized quantity.

#### D. Current sources

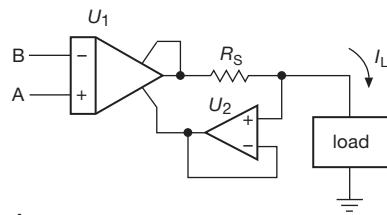
Figure 5.69 shows how to rig up a difference amplifier so that the (differential) input signal controls the voltage drop across a series sense resistor  $R_s$ ; in other words, it's a current source. You're welcome to use a single-ended control input, if you want. The output current can be of either polarity, and these circuits don't know, or care, whether the load returns to ground or to some other potential.

The op-amp follower  $U_2$  should be chosen for an input bias current that is small compared with the minimum load current, and an offset voltage that is small compared with the drop across  $R_s$  at minimum load current. You might begin by choosing  $R_s$  for  $\sim 1\text{ V}$  drop at  $I_L(\text{max})$ , or smaller for high  $I_L(\text{max})$  or for low supply voltage, then see if  $R_s I_L(\text{min})$  is at least  $100\ \mu\text{V}$  or so. For a large dynamic range, say  $I_L(\text{max})/I_L(\text{min}) \geq 10^4$ , you'll need low-offset amplifiers for both  $U_1$  and  $U_2$ ; candidates are listed in Tables 5.2 (page 302), 5.3 (page 303), 5.5 (pages 320–321) and 5.6 (page 335).

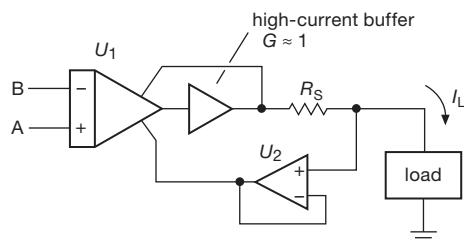
The circuit in Figure 5.69A is good for low currents. For load currents greater than  $\sim 5\text{ mA}$  use a power buffer (Figure 5.69B); that can be a wideband (for stability) unity-gain IC buffer like the LT1010, or (if you need only one polarity of output current) a MOSFET or BJT follower. A favorite trick is to use a 3-terminal adjustable regulator (like an LM317) as a "power buffer," thus taking advantage of

its internal thermal and overcurrent protection. To use it this way, you drive the ADJ pin, and the OUT pin “follows” 1.25 V higher (should this be called a voltage *leader*?!). As always, feedback takes care of the offset.

We like using  $G = 10$  difference amplifiers, such as INA106 or INA143, connected “backwards” for  $G = 0.1$ , because the sense voltage is then one-tenth of the programming voltage and thus doesn’t eat up so much of the output compliance range. In common with other current sources whose output comes from an op-amp, these configurations become more like *voltage* sources at high frequencies, where op-amp compensation and slew-rate effects dominate (see §§4.2.5 and 4.4.4).<sup>62</sup> Better performance at high frequencies can be gotten with an active current source based on an instrumentation amplifier, configured with an inherently high impedance output terminal, as shown in Figure 5.87 in §5.16.9.



A. low current,  $I_L \lesssim 5\text{mA}$



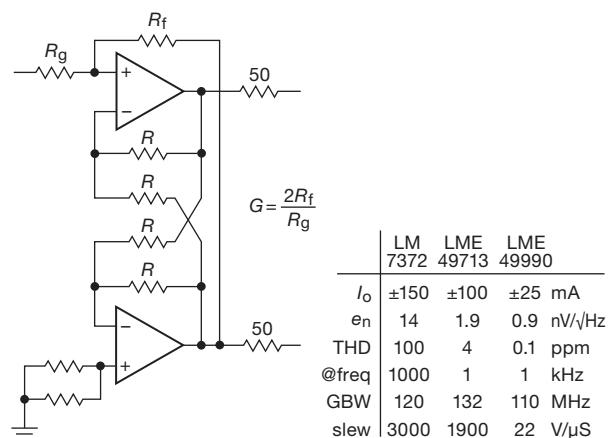
B. “any” current,  $R_S : \text{m}\Omega \text{ to } \text{G}\Omega$

**Figure 5.69.** Precision current sources:  $I_L = G_{\text{diff}}(V_A - V_B)/R_S$ . The (bipolarity) output current in (A) is limited to  $U_1$ ’s  $I_{\text{out}}$  (max). Adding a unity-gain power buffer (an IC wideband buffer, or a transistor follower) in (B) allows large output currents (follower  $U_2$  can be omitted if  $R_S$  is less than  $0.2\Omega$  or so). These circuits don’t know, or care, where the load returns.

<sup>62</sup> You can define an effective current-source output capacitance  $C_{\text{eff}} = I_{\text{out}}/S$  (where  $S$  is the output slew rate) as a way to characterize this shortcoming.

## E. High-level line driver

Professional audio lives and breathes *differential* analog signaling, in the form of balanced lines terminated (usually) in a nominal  $600\Omega$  bridging resistance. And the levels are substantial: pro-audio equipment adopts a “0 dB” standard of  $1.23\text{ Vrms}$ ,<sup>63</sup> and you’ll generally see additional specified headroom of  $16\text{ dB}$  to  $20\text{ dB}$  without clipping. So a  $+20\text{ dB}$  level is  $12.3\text{ Vrms}$ , or a differential amplitude of  $17.4\text{ V}$  ( $34.8\text{ Vpp}$ ). That requires some serious attention to line drivers, which should not compromise the low noise and distortion qualities of the program material.



**Figure 5.70.** Differential high-level line driver for professional audio.

Figure 5.70 shows a nice  $G = 2$  circuit for the job, based around a pair of unity-gain difference amplifiers. These are implemented here with wideband op-amps of substantial output-current capability, allowing for an overall gain (differential-output voltage divided by single-ended input voltage) other than  $\times 2$ . The particular listed op-amps all run from supply voltages to  $\pm 18\text{ V}$ , producing output swings (on each line of the pair) to  $\pm 15\text{ V}$  or so (significantly greater than the  $\pm 9\text{ V}$  corresponding to a  $20\text{ dB}$  pro-audio headroom). Three op-amp choices are shown, of comparable bandwidth ( $\sim 100\text{ MHz}$ ) but aimed at different applications. The LM7372 is well specified out to  $10\text{ MHz}$  and intended for video and broadcast applications, whereas the LME49xxx parts are optimized for audio bandwidths, with pretty impressive specifications out to  $20\text{ kHz}$ . It’s not often (more like *never!*) that you see amplifiers with total harmonic distortion (THD) specifications of 0.1 ppm

<sup>63</sup> The base unit for audio level is “0 dBu,” an rms voltage corresponding to  $1\text{ mW}$  into a  $600\Omega$  load; that works out to  $0.775\text{ Vrms}$ . Pro-audio’s 0 dB level is  $+4\text{ dBu}$ , hence  $1.23\text{ Vrms}$ ; home audio is considerably less muscular, at  $-10\text{ dBu}$ , or  $0.25\text{ Vrms}$ .

(−140 dB), here combined with very low voltage noise ( $1.4 \text{ nV}/\sqrt{\text{Hz}}$  at 10 Hz).<sup>64</sup>

Some interesting alternatives are the DRV134 and the LME49724, which integrate a fully differential output circuit, capable of comparable performance, into one IC. The latter are examples of *fully-differential amplifiers*, with differential inputs as well; one input can be grounded to work with single-ended signal sources. We'll see these later, in §5.17.

#### F. Wideband analog over twisted pair

Twisted-pair network cable ("Cat-5e," etc.) is ordinarily used for digital data transmission on local area networks (LANs), but it can be used successfully for analog signals as well. The ubiquitous Cat-5e and Cat-6 cables contain four unshielded pairs (hence "UTP" – unshielded twisted pair), which interestingly are twisted with differing (incommensurate) pitches to minimize normal-mode coupling.

However, there's plenty of common-mode coupling, both between pairs and to the outside world. So you need to use a differential driver (§5.17) combined with a difference amplifier at the far end. And, for wideband transmission over anything longer than a few inches you need to terminate the pair in its characteristic impedance of  $100 \Omega$  (see Appendix H).

Figure 5.71 shows how to use a difference amplifier as an analog "line receiver" for such signals (we'll show the driving end later, in §5.17.2). This circuit comes from the AD8130's datasheet, illustrating the use of some "peaking" to compensate for the signal attenuation at high frequencies in the rather long (300m!) cable. The amplifier has  $G = 3$  at low frequencies, with  $R_4C_1$  kicking in around 1.6 MHz (where  $C_1$ 's reactance equals  $R_3$ ) to boost the sagging response. The result of this simple "equalizer" is to produce an overall response flat to  $\pm 1 \text{ dB}$  from dc to 9 MHz. The  $\times 3$  low-frequency gain is needed (a) to compensate for the signal that is lost because of the resistive divider consisting of  $R_1$  and the cable pair's  $50.5 \Omega$  round-trip resistance (a factor of 1.5), and (b) to double the output signal so it can drive a "back-terminated" video cable (see Appendix H).

#### 5.14.3 Performance parameters

In the preceding sections we glossed over some important issues: input impedance (both differential and common

mode) and its effect on gain when driven with finite source impedance or its effect on CMRR when driven from an unbalanced source impedance; common-mode input range; and amplifier bandwidth and its effect on CMRR. It's time to peel the onion.

#### A. Input impedance

From Figure 5.65 it might seem that the input impedance of a difference amplifier is  $R_i$  (or maybe some multiple of that), and that all is well if the signal driving it has a source (Thevenin) impedance  $R_S$  that is much less, perhaps our usual seat-of-the-pants criterion of  $R_S \lesssim 10R_i$ .

Not so! These amplifiers revel in their precise gain (the selections in Table 5.7 on the next page have worst-case gain errors of 0.1% or better), and that gain accuracy is compromised unless  $R_S$  is smaller than the amplifier's  $R_i$  by at least that factor. That's because the signal's source impedance is effectively in series with  $R_i$ , reducing the gain by the factor  $R_i/(R_i + R_S)$ .

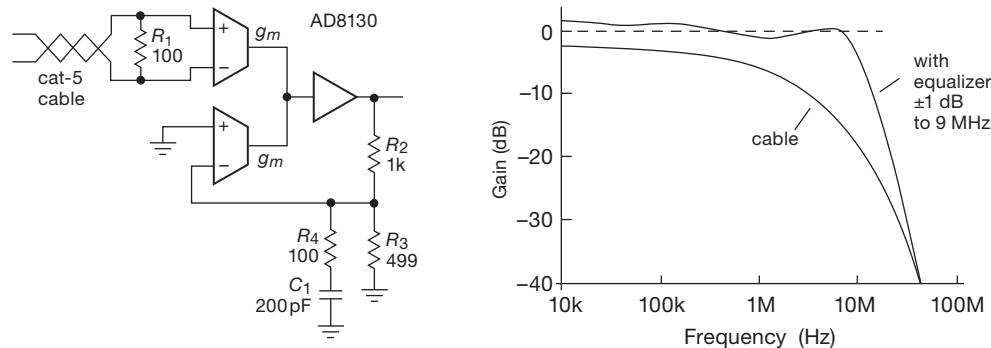
That gain reduction is actually the lesser of two problems. If you read the fine print on the datasheet, you'll discover that the excellent CMRRs of these amplifiers (surely their greatest claim to fame) is invariably specified for (drumroll)  $R_S = 0 \Omega$ ! You could reasonably expect that the CMRR would be maintained if you drive both inputs with equal source impedances, because that should maintain the resistor ratio match. You could expect that, but you would be disappointed: the CMRR degrades rapidly with rising source impedances, even if they are precisely matched.

Why is that? During manufacture the internal resistors are laser trimmed, so that the ratio  $R_F/R_i$  of the upper resistor pair precisely matches the corresponding ratio of the lower pair. The ratios are what are matched, at the expense of the absolute values; the two input resistors  $R_i$  may differ somewhat.<sup>65</sup> So, if you drive it with a signal pair of matched source impedances  $R_S$ , you get a mismatch in the feedback ratios, hence degraded CMRR. Bottom line: drive these puppies from an op-amp output, or from a signal source of very low  $R_S$  (e.g., a low-value current-sense resistor).

That's not always possible, of course. What can be done to raise the input impedance? The first thought might be simply to raise all the resistor values by some large factor. That has several drawbacks, the most serious of which are (a) the Johnson noise contribution

<sup>64</sup> You've got to keep source impedances quite low in order not to degrade such a low  $e_n$ ; even a  $100 \Omega$  resistor has an open-circuit voltage noise of  $1.3 \text{ nV}/\sqrt{\text{Hz}}$ . See Chapter 8.

<sup>65</sup> And the overall resistor scaling is typically good to only  $\pm 20\%$  of the nominal value on the datasheet: absolute resistance value is sacrificed on the altar of resistor ratio matching.



**Figure 5.71.** Differential line receiver for wideband analog over twisted pair. The “treble boost” provided by  $C_1R_4$  compensates for the high-frequency rolloff of a 300m length of Cat-5 cable, as shown. See also Figure 5.101.

**Table 5.7 Selected Difference Amplifiers**

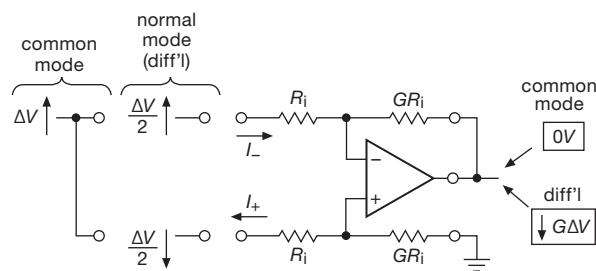
Part #	Config	Gain	Offset Voltage			CMRR		Diff'l Curve	Noise <sup>k,t</sup>			BW	Settle 0.01%	Filter?	Supply			Comments	
			$\Delta G_{\text{max}}$ (%)	$V_{\text{CM}}^{\text{a}}$ (V <sub>pp</sub> )	typ	max	typ	min	$Z_{\text{in}}$ (kΩ)	$V_{n(\text{pp})}$ (µV)	$e_{n^r}$ (nV/√Hz)				Range (V)	$I_S$ (mA)	Cost (\$US)	DIP?	
INA105K	A	1.0	0.025	20	50	500	100	72	1	50	2.4	60	1	5	-	10-36	1.5	8.88	● legacy
AMP03G	A	1.0	0.008	20	25	750	95	75	1	50	2	20 <sup>c</sup>	3	1	-	10-36	2.5	5.86	●
INA134	A	1.0	0.075	26	75	250	90	74	4	50	7 <sup>d</sup>	52	3.1	3	-	8-36	2.4	2.36	● audio, <5ppm dist
INA154	A	1.0	0.1	25	75	1500	90	74	4	50	2.6	52	3.1	3	-	8-36	2.4	2.36	- low cost
INA132P	A	1.0	0.075	28	75	250	90	76	5	80	1.6	65 <sup>c</sup>	0.3	88	-	2.7-36	0.16	4.62	●
AD8271B	A <sup>m</sup>	0.5,2	0.02	>18	300	600	92	80	10	20	1.5	38	15	0.55	-	5-36	2.6	3.50	- $G_{\text{diff}} = 0.5,1,2$
AD8273	A	0.5,2	0.05	>18	200	1400	86	77	7	24	7 <sup>d</sup>	52	20	0.75	-	5-36	2.5	3.13	- dual audio, 6ppm dist
THAT1206	F	0.50	0.5	26	-	10mV	90	70	-	48	-	28 <sup>n</sup>	34	-	-	24-40	4.7	4.75	● $Z_{\text{CM}}=10\text{M}$ , <6ppm dist
AD8278B	A	0.5,2	0.02	>18	50	100	100	80	1	80	1.4	47	1	9	-	2-36	0.2	2.72	- $G_{\text{diff}} = 0.5,2$ ; dual '79
INA106	A	10	0.025	11	50	200	100	86	1	20	1	30	5	10	-	10-36	1.5	11.00	● 0.10, legacy
INA143U	A	10	0.1	15	100	500	96	86	4 <sup>h</sup>	20	1	30	0.15	9	-	4.5-36	1.0	3.36	- 0.10, dual=INA2143
LT1991A	A <sup>m</sup>	1,4,10	0.06	27	15	50	100	77	8 <sup>h</sup>	90	0.25	46	0.11 <sup>f</sup>	48	f	2.7-40	0.10	2.50	- 3,9,12, includes filter
LT1996A	A <sup>m</sup>	9-117	0.07	27 <sup>g</sup>	15	50	100	80	-	33 <sup>g</sup>	0.25	18	0.04	85	f	2.7-40	0.10	5.72	- 9,27,81, includes filter
LT1995	A <sup>m</sup>	1-7	0.2	31 <sup>g</sup>	600	4000	87	75	-	4 <sup>g</sup>	-	14	32	0.1	-	2.7-40	7	3.78	- 1,2,4,6
INA146	B	0.1 <sup>o</sup>	0.1	100	1000	5000	80	70	11	200	10	550	0.55	80	●	4.5-36	0.57	4.25	● high-voltage inputs
INA117P	C	1.0	0.02	200	120	1000	94	86	8	800	25	550	0.2	10	-	10-36	1.5	5.54	● high-voltage inputs
AD629B	C	1.0	0.03	270	100	500	96	86	6	800	15	550	0.5	15	-	5-36	1.2	7.21	● high-voltage inputs
INA149	C	1.0	0.02	275	350	1100	100	90	1	800	20	550	0.5	7	-	4-40	0.8	6.00	- high-voltage inputs
AD8479	C	1.0	0.01	600	500	1000	96	90	u	2000	30	1600	0.13	11	-	5-36	0.55	9.66	● highest CM voltage
AD628A	C ext	0.1	120	-	1500	-	75	220	15	300	0.6	40	●	4.5-36	1.2	3.36	- HV inputs, filter+gain low-voltage		
AD8275B	D	0.2	0.024	27	150	500	96	80	2	108	1 <sup>c,e</sup>	40 <sup>e</sup>	15	0.45	-	3.3-15	1.9	4.23	- 5, with $V_{\text{ref}}/2$ offset
AD830	E <sup>g</sup>	1-10	0.6	24	1500	3000	100	90	12	370	-	27	85	0.025	p	8-36	14	4.86	- $G=1+R_2/R_1$ , $V_{\text{in}} < 2\text{V}$
AD8129	E <sup>g</sup>	10-100	0.6	20	200	800	105	92	12	1000	-	4.6	200	0.02	p	4.5-25	10	2.90	- $G=1+R_2/R_1$ , $V_{\text{in}} < 2\text{V}$
AD8130	E <sup>g</sup>	1-20	0.6	20	400	1800	105	90	12	6000	-	12.3	250	0.02	p	4.5-25	11	2.90	- $G=1+R_2/R_1$ , $V_{\text{in}} < 2\text{V}$
EL5172	E <sup>g</sup>	1-20	1.5	20	7mV	25mV	95	75	-	300	-	26	250	0.01	p	4.7-12	5.6	1.27	- $G=1+R_2/R_1$ , $V_{\text{in}} < 4\text{V}$
INA152EA	A	1.0	0.1	18	250	1500	94	80	3	40	2.4	87	0.8	25	-	2.7-20	0.5	3.60	- 0.5,1,2, RRIO
MAX4198	A	1.0	0.1	0.5	30	500	90	74	9 <sup>h</sup>	50	7.8	58	0.175	34	-	2.7-7	0.05	2.38	-

Notes: (a) maximum common-mode voltage. (b) 0.01–10Hz or 0.1–10Hz. (c) at 100Hz. (d) 20 to 20kHz. (e) RTO (referred to output). (f) includes 4pF filter caps. (g) circuit E from Fig. 5.89. (h) CMRR curve flattens at BW frequency. (k) RTI (referred to input) unless noted. (m) multiple resistors. (n) noise =  $-107\text{dBu}$ , 20kHz BW. (o) includes extra opamp filter stage. (p) for  $G>1$ . (q) depends on gain. (r) at 1kHz. (t) typical. (u) between 5 & 6.

( $e_n=0.13 R^{\frac{1}{2}} \text{nV}/\sqrt{\text{Hz}}$ , thus a devastating  $130 \text{nV}/\sqrt{\text{Hz}}$  for resistor values of  $\sim 1 \text{ M}\Omega$ ; see Chapter 8), and (b) the bandwidth penalty caused by distributed stray capacitances. The second thought would be to put a pair of precision op-

amp followers at the inputs. That's OK, but there's a better way still, in circuit configurations such as the “three-opamp instrumentation amplifier,” in which a front-end differential stage of high CMRR and high impedance drives a

difference amplifier output stage. We'll see these wonderfully useful amplifiers a bit later.



$$\text{common mode: } Z_{in}(\text{each}) = (G + 1)R_i \\ Z_{in}(\text{combined}) = \frac{1}{2}(G + 1)R_i$$

$$\text{differential mode: } Z_{in}(-) \equiv \frac{\Delta V}{\Delta I_-} = 2 \frac{G + 1}{2G + 1} R_i \\ Z_{in}(+) \equiv \frac{\Delta V}{\Delta I_+} = \frac{G + 1}{2} R_i \\ "Z_{in}" \equiv \frac{\Delta V}{\Delta(I_- - I_+)} = \frac{G + 1}{G} R_i$$

**Figure 5.72.** The several input impedances of the difference amplifier.

Circling back around to the initial question, what exactly is the “input impedance”? There are several answers – look at Figure 5.72. By *common-mode* input impedance we mean the incremental impedance seen at either input<sup>66</sup> when both are driven together. With common-mode drive the two impedances are equal (apart from minor mismatching, as described above), because the output is fixed.

The *differential*-input impedance is a longer story. Taken individually (by grounding the other input), the two inputs exhibit different  $R_{in}$ : the inverting input connects to a virtual ground by way of  $R_i$ , so its  $R_{in} = R_i$ , whereas the noninverting input sees  $R_{in} = R_i + R_f$ . For a difference amplifier with  $G=10$ , for example, these differ by a factor of eleven. This is a useful result, particularly if you plan to use a difference amplifier in a single-ended configuration like Figures 5.66D and E. A purist could argue, though, that we've done it wrong: a single-ended input change is really a combination of a purely differential input plus a common-mode offset of half that value. To satisfy such a person we've calculated the expressions in Figure 5.72, based upon a “pure” (symmetrical) differential input. Even when defined this way the input impedances seen at the two inputs are different. This leads us to a final definition of

differential input impedance based on the *net* input-current change, as shown. Datasheets most often list this value, without explanation: you've been warned!

A final observation: you won't generally see difference amplifiers with gains greater than 10, because  $R_i$ , and therefore the differential input impedance, becomes unmanageably small. For example, to get  $G = 1000$  you might choose  $R_i = 100 \Omega$  and  $R_f = 100k$ . OK, maybe you can live with  $\sim 100 \Omega$  input impedance; but you'd have to match the signal source impedances to  $0.001 \Omega$  in order not to degrade the CMRR. The take-away: use an *instrumentation amplifier* (§5.15), not a difference amplifier, for high-gain differential-input applications.

## B. Common-mode input range

The voltage divider formed by  $R_i$  and  $R_f$  permits the standard difference amplifier to accept input signals *beyond* the supply rails: protection diodes are at the internal op-amp inputs, so the input signals could in principle swing as far as  $\pm V_S(G + 1)/G$  with supply voltages of  $\pm V_S$  (figure out why). For example, the AD8278 has  $R_i = 40k$  and  $R_f = 20k$ , so it can be connected for  $G = 0.5$  (“normal”) or  $G = 2$  (reversed). It specifies the common-mode input range for both gains:  $-3(V_S + 0.1)$  to  $+3(V_S - 1.5)$  for  $G = 0.5$  (that is, about  $\pm 40V$  with  $\pm 15V$  supplies, which can be very useful indeed!), and  $-1.5(V_S + 0.1)$  to  $+1.5(V_S - 1.5)$  for  $G = 2$  (i.e.,  $\pm 20V$  with  $\pm 15V$  supplies). But check the specs – not all difference amplifiers let you go that far.

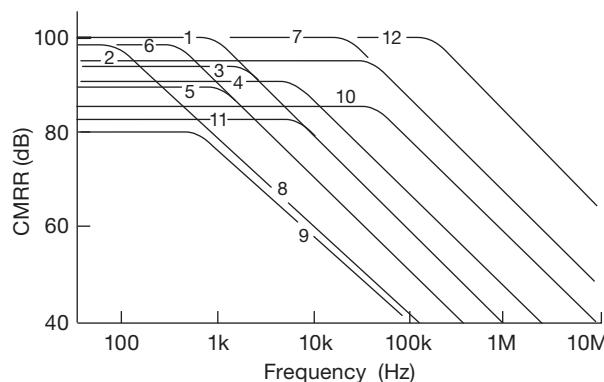
A few of the difference amplifiers in Table 5.7 on the preceding page do better still; for example, the INA117 has a common-mode range of  $\pm 200V$  while maintaining a differential gain of unity. This it does by using a pair of 200:1 voltage dividers at the input to bring the  $\pm 200V$  signal within the op-amp's common-mode range of  $\pm 10V$  (the circuit is shown in Figure 5.75C). The price it pays is degraded offset and noise: typical values of  $120 \mu V$  and  $550 nV/\sqrt{Hz}$ , compared with  $25 \mu V$  and  $20 nV/\sqrt{Hz}$  for the conventional AMP03.<sup>67</sup>

An important point: when using difference amplifiers with large input voltage ranges, beware of the large equivalent input errors created by imperfect common-mode rejection. For example, the AD629B specifies a typical dc CMRR of 96 dB, but you've got to consider the worst-case (minimum) value of 86 dB. With that CMRR, a 200 V

<sup>66</sup> Some manufacturers specify half that value, i.e., the impedance with both terminals tied together; the datasheet usually tells you which they mean.

<sup>67</sup> There's another way to boost  $V_{CM}$  without such compromise, by using a second op-amp to cancel the common-mode signal; see Figure 7.27 in the previous edition of this book. We are unaware of any commercial difference amplifiers that use this trick.

common-mode input has an input-referred differential error of 10 mV, completely swamping the 0.5 mV specified maximum offset voltage. Put another way, the error that is due to finite CMRR is larger than the specified  $V_{OS}$  for input  $|V_{CM}| > 10$  V. And the situation is worse still at signal frequencies: you might imagine using such a difference amplifier for powerline current monitoring. At 60 Hz the CMRR of the INA117 (similar to that of the AD629B at dc) degrades to 66 dB (min). So the 160 V peak powerline signal produces a huge 80 mV input error. And ideally you'd want to monitor currents on powerlines at higher voltages still, perhaps to 400 V. There are better ways to do this; take a look at §13.11.1 if you're curious.



**Figure 5.73.** Common-mode rejection ratio versus frequency for the difference amplifiers in Table 5.7 on page 353 (identified in the “Curve” column).

### C. Bandwidth

Difference amplifiers are built with op-amps, frequency compensated for stability with the by-now familiar  $1/f$  open-loop gain rolloff. As with any op-amp circuit, loop gain is responsible for good behavior and the loss of loop gain at higher frequencies not only limits the bandwidth of a difference amplifier (and its linearity, constancy of gain, low output impedance, etc.), it also degrades the all-important CMRR. Figure 5.73 shows this behavior, for the difference amplifiers listed in Table 5.7 (page 353). Not surprisingly, amplifiers with greater closed-loop bandwidth (therefore with op-amps of greater  $f_T$ ) maintain high CMRR to higher frequencies.

Note that some amplifiers perform well near dc but poorly at high frequencies, for example classic parts like the INA105, AD829 and LT1991, whereas others not so well rated at dc may look better at high frequencies, such as the AD8271 and the MAX4198. A few are outstanding performers from dc to high frequencies, according to the

plots, such as the AD8273<sup>68</sup> and our favorite AD8275. The transconductance-balancing AD8129 does very well, but is limited to small inputs and high gains. This information is not shown in the datasheet specifications; it's necessary to look for performance plots to make the comparisons.

Common-mode rejection at higher frequencies is degraded also by the effects of stray inductance and by asymmetrical capacitive loading. It is necessary to balance the circuit capacitances to achieve good CMRR at high frequencies. This may require careful mirror-image placement of components. Even when so symmetrized, the decreasing input capacitive shunt reactance at higher frequencies creates an increasing sensitivity to any unbalance in signal source impedance.<sup>69</sup>

### 5.14.4 Circuit variations

#### A. Filter node

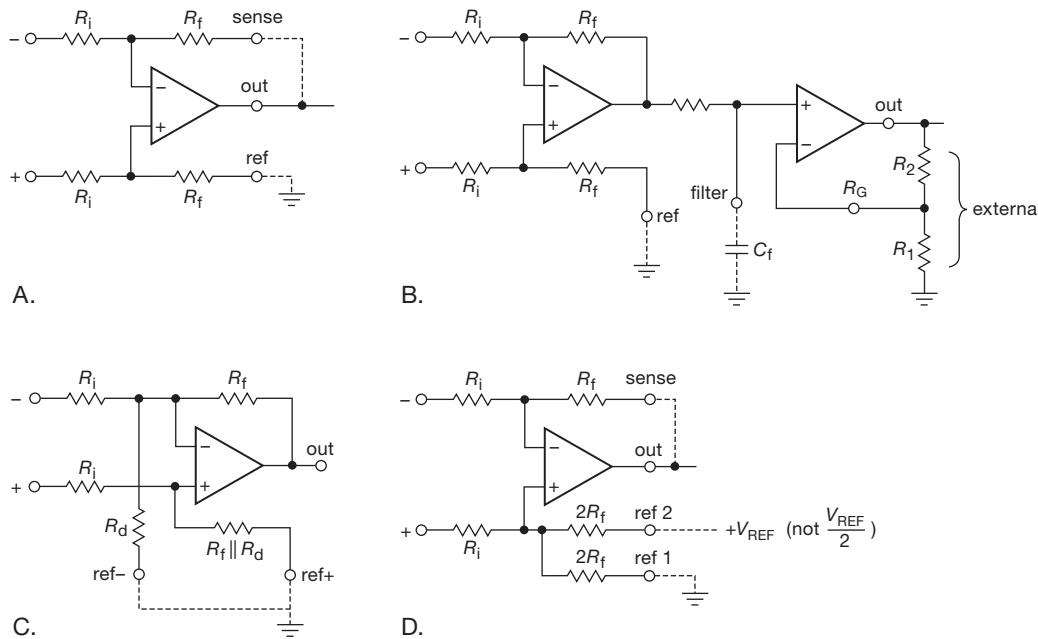
The difference amplifiers in Table 5.7 (page 353) include one (the INA146) with a node brought out for noise (low-pass) filtering of the difference stage ( $G = 0.1$ ) by a capacitor to ground (Figure 5.75B). It includes a second single-ended stage, with gain set by a pair of external resistors; so you can have overall gains from 0.1 to 100. The low first-stage gain gives you a large  $\pm 100$  V common-mode range, though at the expense of noise and offset.

#### B. Offset trim

Integrated difference amplifiers are factory trimmed to pretty good precision, with typical values in the 25–100  $\mu$ V range (but with worst-case offsets an order of magnitude larger). As with any op-amp circuit, you can rig up an external trim, as in Figure 5.74A. Here  $R_2R_3$  divides the trimmer's  $\pm 15$  V range to  $\pm 1$  mV at the REF pin;  $R_1$  balances the  $10\Omega$  added resistance to ground, to preserve the CMRR. The amplifier's gain is slightly reduced by the ratio  $R_f/(R_f + R_2)$ ; for a typical  $R_f$  of 25k, that's 0.04%, in the same range as the amplifier's specified gain accuracy. If this bothers you, use a smaller  $R_2$ .

<sup>68</sup> Although the datasheet graphs for the AD8273 show a best-in-class CMRR of 100 dB below 40 kHz, the tabulated typical CMRR is shown as only 86 dB. We like the plot better, but the designer may be stuck with the worst-case spec of 77 dB. Some users may want to perform incoming inspections to settle the issue.

<sup>69</sup> The professional audio people are keenly aware of these effects, and they don't mince words. As stated pithily by Whitlock and Floru in an Audio Engineering Society paper, “**Noise rejection in a balanced system has absolutely NOTHING to do with signal symmetry** (equal and opposite signal voltage swings). It is the balance of common-mode impedances that defines a balanced system!” [emphasis in the original].



**Figure 5.75.** Circuit configurations for the difference amplifiers in Table 5.7 on page 353 (identified in the “Config” column). The “E” form is shown in Figure 5.89. The “C” form is used for high voltages (e.g.,  $\pm 270\text{ V}$  for the AD629B).

### C. CMRR trim

Likewise, you can trim out residual CMRR (caused by slight mismatch of resistor ratios  $R_f/R_i$  in the two paths) with the circuit of Figure 5.74B. It’s important to limit the trim range, to permit an accurate and stable trim to something considerably better than the off-the-shelf 80 dB (worst-case) CMRR specification. You can’t get any old value of trimmer, and it’s best not to use values less than  $100\Omega$  (even if you can find them) if you care about stability. Here we’ve chosen standard resistor values and a  $100\Omega$  trimmer to produce a resistance range of  $20\text{--}30\Omega$  from the REF terminal to ground, providing a  $\pm 5\Omega$  symmetrical variation around  $R_1$ . For the 25k resistor values of this unity-gain difference amplifier (rather typical; see Table 5.7 on page 353), this corresponds to a trim range adequate to null an initial CMRR of 75 dB. You can, of course, add an offset null to this circuit, as indicated.

### D. Single-supply offset

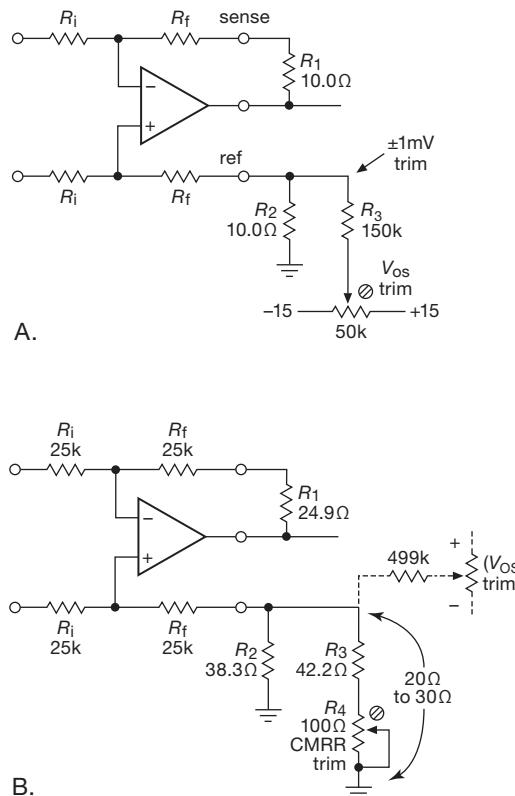
One of the difference amplifiers in Table 5.7 (page 353) helpfully splits the reference feedback resistor into a parallel pair (Figure 5.75D), so it’s easy to offset the output voltage range. For example, you could run the amplifier on a single  $+5\text{ V}$  supply, with REF2 driven from a clean reference of that same voltage. With no difference signal the output will be  $+2.5\text{ V}$ . The amplifier can accommodate in-

put signals over a  $\pm 10\text{ V}$  range, and its gain of 0.2 takes a  $\pm 10\text{ V}$  differential input to a  $0\text{--}4\text{ V}$  output. You can, of course, use a lower reference voltage. It’s often convenient to use  $V_{ref} = 4.096\text{ V}$  when driving an ADC; this makes the step size come out in round numbers, e.g., 1 mV/step for a 12-bit conversion.

## 5.15 Instrumentation amplifier

The difference amplifiers of the previous section are inexpensive, and fine for many applications; and they have the nice feature of accepting inputs beyond the rails. But they have limited gain ( $\leq 10$ ) and CMRR ( $\lesssim 85\text{ dB min}$ ), their resistors make them somewhat noisy ( $20\text{ nV}/\sqrt{\text{Hz}}$  to  $50\text{ nV}/\sqrt{\text{Hz}}$ ), and their relatively low input resistance ( $\sim 10\text{k}$  to  $\sim 100\text{k}$ ) limits their utility to situations where the driving signals are of low impedance (op-amp outputs, low-Z balanced lines, low- $R$  sense resistors).

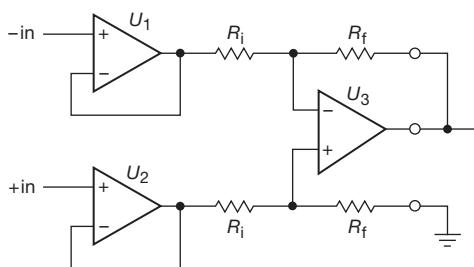
If you need lots of gain, or a high input impedance, or superior CMRR, you need something different. It’s called an *instrumentation amplifier*. These impressive devices have input impedances upward of  $10^9\Omega$ , gains from unity to 1000 or so, low voltage noise (down to  $\sim 1\text{ nV}/\sqrt{\text{Hz}}$ ), and worst-case CMRRs of 100–120 dB (look ahead to Table 5.8 on page 363).



**Figure 5.74.** Trimming the offset and CMRR of a difference amplifier.

### 5.15.1 A first (but naive) guess

High input impedance – that’s easy, just add op-amp followers to the difference amplifier (Figure 5.76); and then the resistors  $R_i$  and  $R_f$  can be smaller, reducing their Johnson noise contribution.



**Figure 5.76.** A first stab at improving the difference amplifier.

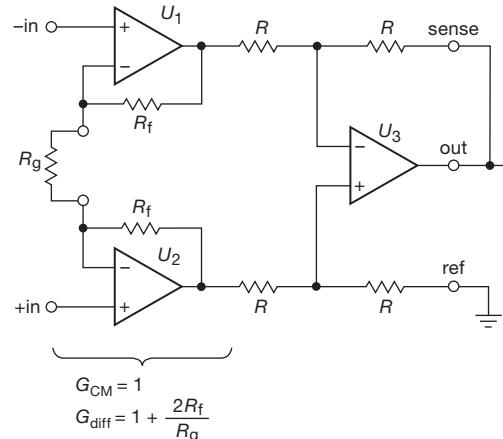
Indeed, this circuit has the enormous input impedance we expect from an op-amp follower, so there is no longer a problem from any reasonable source impedance.<sup>70</sup> But it

<sup>70</sup> At least at dc. At higher frequencies it again becomes important to

does not improve the CMRR, which is still limited by the resistor ratio matching of  $R_f/R_i$ : it’s really hard to do better than 100,000:1 with on-chip laser trimming (both initial trim and stability with time and temperature). In fact, this circuit degrades the CMRR somewhat, with two more amplifiers in the signal path.

### 5.15.2 Classic three-op-amp instrumentation amplifier

The circuit in Figure 5.77 is much better. It is the standard “three-op-amp instrumentation amplifier,” one of several configurations that provide the desirable combination of high CMRR, high  $R_{in}$ , low  $e_n$ , and plenty of gain when you need it. The input stage is a clever configuration of two op-amps that provides high differential gain and unity common-mode gain without requiring close resistor matching. Its differential output represents a signal with substantial reduction in the comparative common-mode signal (when configured for  $G_{diff} \gg 1$ ), and it is used to drive a conventional differential amplifier circuit. The latter is usually arranged for unity gain and is used to generate a single-ended output while removing the common-mode signal.



**Figure 5.77.** The classic three-op-amp instrumentation amplifier.

It’s worth looking more closely at this circuit. We’ve hinted that it can deliver very high CMRR and very low  $e_n$ . But that is true *only when configured for large differential*

have matched source impedances relative to the common-mode signal, because the input capacitance of the circuit forms a voltage divider in combination with the source resistance. “High frequencies” may even mean 60 Hz and its harmonics, because common-mode ac powerline pickup is a common nuisance.

*gain.* To see why, imagine we configure it for  $G_{\text{diff}} = 1$ , by omitting the gain-setting resistor  $R_g$ . Then we've just got the previous circuit (Figure 5.76), i.e., a buffered unity-gain difference amplifier. It has the same limitations of CMRR (set by resistor matching) and noise (from  $U_3$ 's resistors).

Now imagine we set  $G_{\text{diff}} = 100$ . We would do this by choosing  $R_g$  so that  $1 + 2R_f/R_g = 100$ , i.e.,  $R_g = 2R_f/(G - 1)$ . For the INA103, for example,  $R_f = 3\text{k}\Omega$ , so we'd use  $R_g = 60.6\Omega$ . The INA103 conveniently includes a resistor of that value,<sup>71</sup> so for  $G_{\text{diff}} = 100$  you need only tie a pair of pins together. Let's look again at the CMRR and noise scene.

First, the CMRR: the front end has a differential gain of 100 and a common-mode gain of unity. In other words, it passes on to the difference amplifier stage a signal that has received 40 dB of CMRR blessings. Another 80 dB of CMRR in the output stage, and we've got the promised 120 dB CMRR. These numbers are typical of available instrumentation amplifiers, as listed in Table 5.8 (page 363) and graphed in Figure 5.82. For the INA103, for example, the datasheet lists CMRR=86 dB/72 dB (typ/min) for  $G = 1$ , and 125 dB/100 dB (typ/min) for  $G = 100$ .

Second, the noise voltage: the output stage still contributes Johnson noise from its resistor array, along with noise inherent in its amplifier. That cannot be helped. But that noise is combined with the *already-amplified* input signal, so the effect, *relative to the input signal* (RTI), is  $100\times$  less. For the INA103, for example, the datasheet lists the noise density (typical, at 1 kHz) as  $e_n=65\text{nV}/\sqrt{\text{Hz}}$  for  $G = 1$  and  $1.6\text{nV}/\sqrt{\text{Hz}}$  for  $G = 100$ .<sup>72</sup>

### 5.15.3 Input-stage considerations

Several preliminary comments here on the all-important input stage (about which a well-known circuit guru remarked “instrumentation amplifiers are all about their inputs”), with more to come later in §5.16.

#### A. Resistor matching

The circuit looks handsome with its symmetrical matched  $R_f$ 's, but that requirement did not intrude in the discussion above. What is the effect of mismatched feedback resistors

in the first stage? The common-mode gain remains unity (if you tie both inputs together, both outputs follow); and the differential gain expression is the same as before, but with  $2R_f$  not surprisingly replaced with the sum  $R_{f1} + R_{f2}$ . What changes, though, is that a purely differential input causes a differential output (amplified by  $G_{\text{diff}}$ , as before) combined with some common-mode output.

You can see how this goes by imagining that  $U_2$ 's feedback resistor is replaced with a short and  $U_1$ 's resistor with  $2R_f$ , and a symmetrical dc input signal  $\pm\Delta V$  is applied to the inputs:  $U_2$ 's output goes down  $\Delta V$ , while  $U_1$ 's goes up  $(1 + 4R_f/R_g)\Delta V$ . That's the correct differential output, but with a common-mode offset of  $(2R_f/R_g)\Delta V$ . This is not of great concern, providing that the  $R_f$ 's are reasonably matched; they do not require the precise matching needed for the output stage.

#### B. The input amplifiers

It is essential that  $U_1$  and  $U_2$  have excellent CMRRs in order that a purely common-mode input signal is not converted to a differential signal (which would then be passed on to the output). Stated more precisely, they must have *matched* CMRRs, so that the voltage across  $R_g$  accurately tracks the differential input voltage. Viewing the circuit operation more generally in this light, the input amplifiers need not have extremely low individual offset voltages – what matters is that their offset voltages are accurately *matched* and remain so with changes in common-mode voltage. This gives rise to several circuit variants in which the “op-amps”  $U_1$  and  $U_2$  are configurations with well-matched base-emitter drops between each input and the corresponding  $R_g$  pin; see for example Figure 5.88C below.

#### C. Input-stage overload

The input stage amplifiers  $U_1$  and  $U_2$  will clip if their outputs are forced close to their supply rails, even though the output of the full circuit ( $U_3$ 's output) would be expected to stay within safe bounds. Put another way,  $V_{\text{CM}} \pm 0.5V_{\text{diff}}(1 + 2R_f/R_g)$  must not reach either supply rail.

#### D. Signal guards

Instrumentation amplifiers are used with low-level signals, often conveyed by shielded cables to minimize noise. This adds input capacitance, thereby limiting the bandwidth (particularly with signals of moderate to high source impedance). Of perhaps greater importance, it degrades the CMRR at signal frequencies: the cable's shunt capacitance forms a voltage divider with the signal's source impedance, separately for each input; so if you have a

<sup>71</sup> More precisely, it includes an on-chip resistor, ratio matched to  $R_f$ , to produce an overall guaranteed gain of  $100.0 \pm 0.25\%$ .

<sup>72</sup> The datasheet separates out the front-end and second-stage contributions,  $1\text{nV}/\sqrt{\text{Hz}}$  and  $65\text{nV}/\sqrt{\text{Hz}}$ , respectively. From these you can calculate the input-referred noise  $e_n(\text{RTI}) = \{e_n(\text{in})^2 + [e_n(\text{out})/G]^2 + 4kTR_g\}^{1/2}$ . The last term is the square of the Johnson noise voltage  $e_n = 0.13\sqrt{R_g}\text{nV}/\sqrt{\text{Hz}}$ .

signal pair with unbalanced source impedances (a common situation), common-mode signal variations will create some differential signal input.<sup>73</sup> Finally, leakage currents become significant with signals of very high ( $M\Omega$  to  $G\Omega$ ) source impedance. A nice technique for greatly reducing both the effective cable capacitance and any leakage current is to drive the shield actively with a “guard” voltage (Figure 5.78).

If there is a common shield surrounding the signal pair, the idea is to drive it with a buffered replica of the common-mode signal, as in circuit (A); a small series resistor is generally a good idea, for stability. A grounded outer shield can be used to eliminate any noise coupling to the guard, if that becomes necessary. This circuit requires access to the first-stage outputs, which you rarely get in an integrated instrumentation amplifier. Some amplifiers oblige by including this circuitry internally and providing a “data-guard” output pin, as in (B). If not, you can derive a common-mode signal yourself, as in (C).

Common-mode guarding reduces greatly the capacitive loading of the signal pair and therefore improves the CMRR (by minimizing the conversion of common-mode to normal-mode signals). But it does not reduce the effects of cable capacitance (and leakage) on *normal-mode* (differential) signals themselves. To accomplish that you need to shield the signals individually, driving each shield with a replica of the signal it shields, as in circuit (D). This is the familiar “bootstrap,” acting here to reduce both the capacitance and leakage current seen by each signal. It thus minimizes high-frequency rolloff, and dc error in signals of high source impedance.<sup>74</sup> And, as with common-mode guarding, it also minimizes the degradation of CMRR by effectively eliminating the cables’ capacitances.

Some instrumentation amplifiers provide these individual guard outputs, for example the INA116 shown, which is evidently intended for very low current measurements (it boasts a 3 *femtoamp* typical input current). If not, you can roll your own, as in circuit (E), exploiting the fact that the  $R_g$  nodes follow the inputs. A caution, though: there is no guarantee that the signals at the  $R_g$  pins are not offset from the input signals, as for example in the configurations we’ll see later in Figures 5.88C and 5.89F. With such an offset the bootstrap would be effective in minimizing capacitance, but much less so for leakage currents.

<sup>73</sup> Put another way, the cable’s capacitance degrades the CMRR by creating differential phase shifts between the two signals, owing to their unbalanced source impedances.

<sup>74</sup> Instruments for measuring very low currents – “electrometers” and “source measure units” – include guard outputs and (usually) special BNC-like “triax” connectors for use with triaxial shielded cables.

## E. Bootstrapped power supply

Figure 5.79 shows a trick that is analogous to signal guarding (you could call it “power-supply guarding”), occasionally helpful if you need to enhance further the CMRR of an instrumentation amplifier.  $U_3$  buffers the common-mode signal level, driving the common terminal of a small floating split supply for  $U_1$  and  $U_2$ . This bootstrapping scheme effectively eliminates the input common-mode signal from  $U_1$  and  $U_2$ , because they see no swing (due to common-mode signals) at their inputs relative to their power supplies.  $U_3$  and  $U_4$  need not be bootstrapped if this is a discrete implementation. This scheme can do wonders for the CMRR, at least at dc. At increasing frequencies you have the usual problems of presenting matched impedances to the input capacitances.

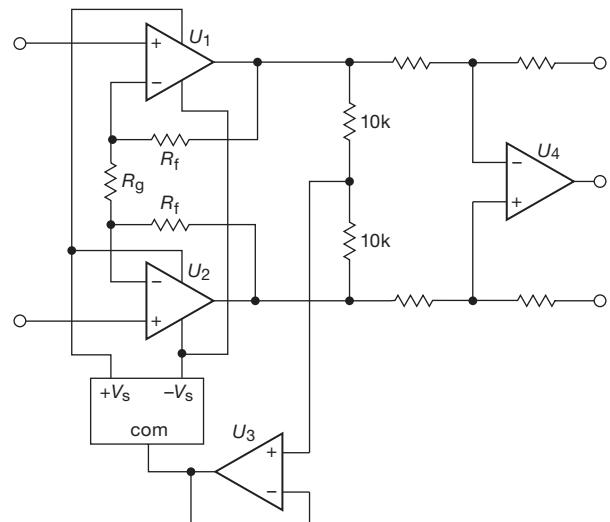
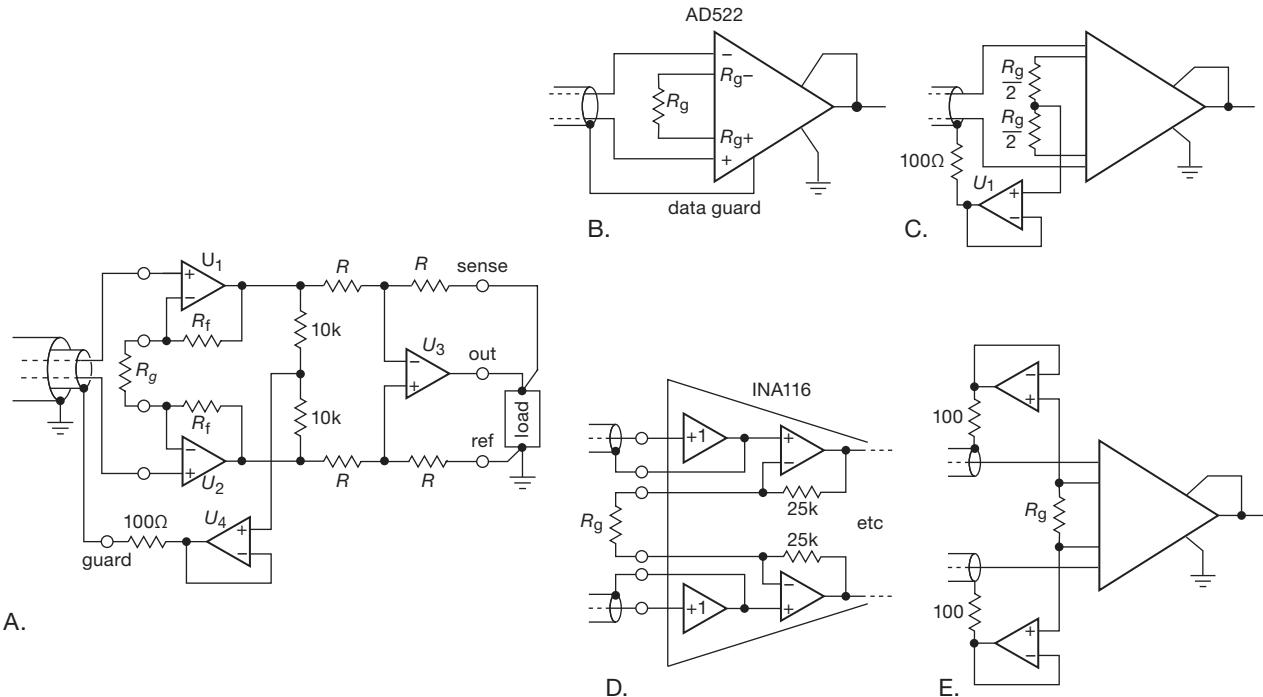


Figure 5.79. Instrumentation amplifier with bootstrapped input power supply for high CMRR.

## 5.15.4 A “roll-your-own” instrumentation amplifier

Integrated instrumentation amplifiers are excellent performers, and you usually can save a lot of work (and expense, and PCB real estate) by taking advantage of the broad selection of available parts, for which Table 5.8 (page 363) is a good starting point. But sometimes you need additional capability, for example a wider range of gains, or precise trimming of offsets and CMRR, or protection against outrageous abuse visited upon the input terminals.

Figure 5.80 is an example (another in the series of Designs by the Masters), based upon a commercialized



**Figure 5.78.** Signal guarding for greatly reduced cable capacitance. (A)–(C) are common-mode guards; (D) and (E) are individual signal guards.

design<sup>75</sup> from the talented John Larkin, to which we've added a few decorations. He needed a flexible front end that combined (a) overvoltage protection to  $\pm 250$  V, (b) logic-switchable gains from 1/16 to 256 by factors of four, (c) precision low offsets, (d) common-mode input range to  $\pm 10$  V for  $G \geq 1$  and  $\pm 140$  V for  $G < 1$ , and (e) common-mode rejection of 120 dB at high gain.

The overall structure is the familiar three-op-amp configuration, with  $U_1$  and its symmetrical twin (not shown) as the differential front end, which drives the unity-gain difference amplifier output stage  $U_3$ . The gain is set by analog switches  $U_2$  and its twin, which select a tap on the resistor string  $R_6$ – $R_{10}$ . For example, when the “ $\times 64$ ” position is selected,  $R_g = 201.1\Omega$  and  $R_f = 6.411\text{ k}\Omega$ , thus a gain  $G_{\text{diff}} = 1 + 2R_f/R_g = 64.8$ .<sup>76</sup>

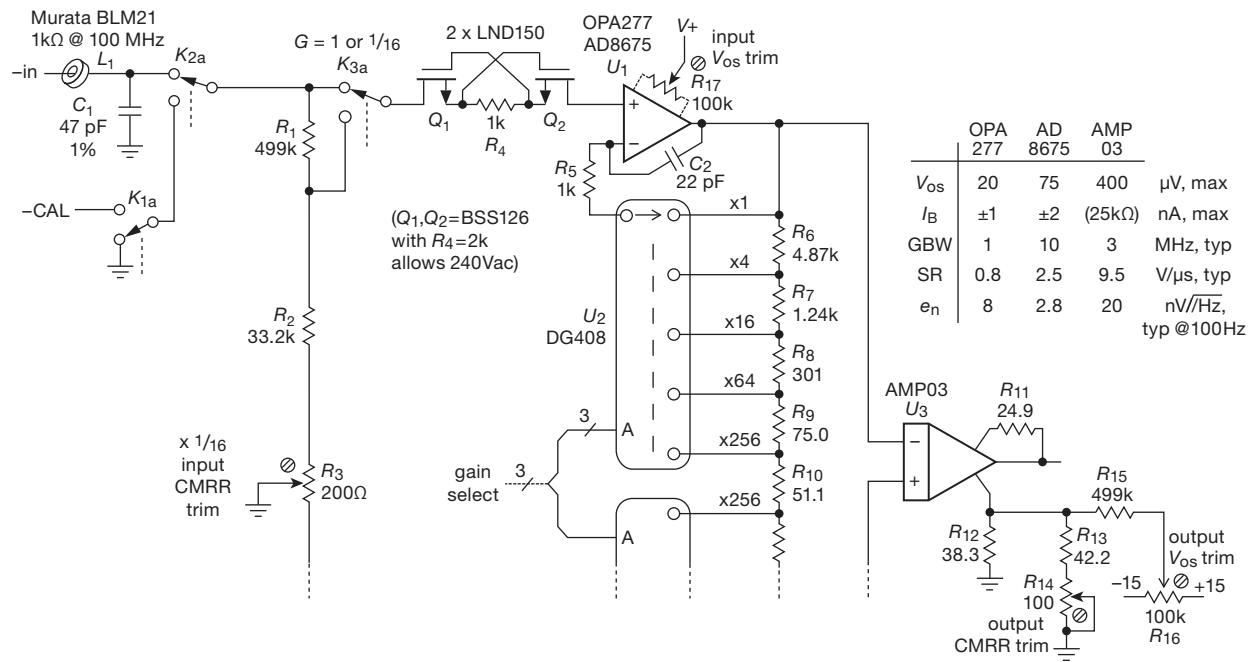
There's plenty more going on, here. Let's take it from

left to right. The lossy chip inductor  $L_1$  (these are often rated by their impedance, mostly resistive, at 100 MHz) combined with capacitor  $C_1$  suppresses high-frequency interference that is outside the amplifier's bandwidth but able to cause nonlinear mischief in op-amp input stages.  $C_1$ 's value is not critical, but its tight tolerance keeps the input impedance balanced in order not to compromise high-frequency CMRR. Relays  $K_1$  and  $K_2$  are used for frequent in-system offset and gain calibrations, essential for establishing and maintaining gain accuracies better than 0.1% (with a circuit using 1% resistor values) and zero offsets of  $10\text{ }\mu\text{V}$  or better.<sup>77</sup> Relay  $K_3$  switches in a  $\times 16$  gain attenuation for input signals beyond  $\pm 10$  V. The input impedance is then set by  $R_1+R_2$ , with  $R_3$  balance trim for good CMRR.  $R_1$ 's value is a compromise between high  $R_{\text{in}}$  and bandwidth: here the  $33.2\text{ k}$ , loaded by a downstream capacitance of  $\sim 10\text{ pF}$ , puts the 3 dB rolloff at 500 kHz, about right for the product's 200 kHz specified bandwidth. But a higher input impedance would be nice, perhaps a nice round-number  $1\text{ M}\Omega$  (preferably with  $R_1$  as a series pair

<sup>75</sup> Highland Technology's V490 VME Multi-Range Digitizer.

<sup>76</sup> The gain-setting resistors are chosen from standard “E96” 1% resistor values, so the actual gains differ from round-number values (with  $\pm 1\%$  tolerance they would never be perfect, anyway). This front end would form part of a data-acquisition system, with overall gain and offset data held in software, from a calibration procedure carried out by relays  $K_1$  and  $K_2$ .

<sup>77</sup> Larkin sings the praises of Fujitsu FTR-B3GA4.5Z DPDT relays, with their sub-picofarad capacitances.



**Figure 5.80.** A “discrete” instrumentation amplifier design that combines precision, high CMRR, large common-mode voltage range, digital gain selection, and protection against  $\pm 250\text{ V}$  inputs. The symmetrical path to the noninverting input of  $U_3$  is omitted to save paper.

of 464k, to accommodate input overvoltages without damage).

The curious  $Q_1Q_2$  twins are a nice touch: these are *depletion-mode* high-voltage MOSFETs (see Table 3.6) from Supertex or Infineon,<sup>78</sup> used here as a  $\sim 0.5\text{ mA}$  bidirectional current limiter to protect the op-amps (whose internal clamp diodes are unfazed by a few milliamps of input current; the AD8675, for example, specifies a maximum  $I_{in}$  of  $\pm 5\text{ mA}$ ).<sup>79</sup> Here  $R_4$  serves to reduce the saturation current from its maximum  $I_{DSS}$  of 3 mA (thus 750 mW dissipation at 250 V input, too much for a little SOT23 transistor), the tradeoff being an additional 1 kΩ of noise-adding input resistance in series with the unavoidable  $\sim 1.7\text{k}$  on-resistance of the series pair of zero-biased MOSFETs. This circuit provides a reasonable level of overdrive protection; but an abrupt input step to +500 V, say, could well couple a large enough current transient (through the

capacitance of  $Q_1$  and  $Q_2$ ) to damage the amplifier, or produce gate breakdown in the MOSFETs themselves.

The op-amps are precision bipolar types with input-current cancellation; note the tradeoff, for the two choices, of precision versus speed and bias current. These are laser trimmed for low offset voltage and are probably good enough without human intervention; but they include offset-trimming terminals to which you can attach a trim-pot as shown. That sounds like a good idea – but be cautious, because the external trimming range you get when you do that is generally too much! For the AD8675, for example, the trimmer adjusts  $\pm 3500\text{ }\mu\text{V}$  (nearly 50 times the maximum untrimmed offset). So it may be a touchy (and unstable) adjustment to improve upon the already-trimmed  $V_{os}$ .

The analog multiplexer selects a tap on the gain-setting string, which makes the gain insensitive to the on-resistance of the switches (here  $\sim 100\Omega$ ). That’s the right way to do it; the *wrong* way would be to use a separate feedback resistor to each switch position, with  $R_{10}$  connected between the common terminals of the multiplexers. The small feedback capacitor  $C_2$  ensures stability, given the lagging phase shift within the loop contributed by the  $\sim 40\text{ pF}$  of switch capacitance to ground.

Finally, we arrive at our destination via the difference

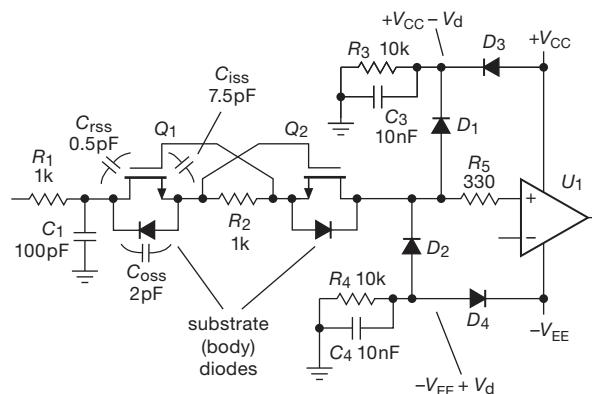
<sup>78</sup> They’re rated at 500 V, and come in three small package styles (TO-92, SOT23, and SOT89 with tab). The alternative BSS126 from Infineon is rated at 600 V, and costs less (\$0.15 in small quantities). It comes only in the SOT-23 package, whereas the LND150 is available in three package styles, including a 1.5 W TO-243 small power package.

<sup>79</sup> Be sure to check that the op-amp does not suffer from phase reversal (see, e.g., §4.6.6), if you care about the output during input overdrive. A robust cure is to use a pair of input clamp diodes, as in Figure 5.81.

amplifier  $U_3$ , chosen for its combination of gain stability (an impressive  $\pm 0.0008\%$  over temperature, typ), fast settling time ( $1 \mu\text{s}$  to  $0.01\%$ , typ), and low noise ( $20 \text{nV}/\sqrt{\text{Hz}}$ , typ). It's easy to add CMRR and offset trim as shown (see earlier discussion in §5.14.4), which matters most at low gain settings.

### 5.15.5 A riff on robust input protection

As we remarked above, a really nasty transient can ruin your whole day. Consider, for example, an input voltage step of  $1 \text{ ns}$  rise time and  $350 \text{ V}$  amplitude (the input accidentally touches a  $240 \text{ Vac}$  powerline): the  $350 \text{ GV/s}$  slew rate would force  $350 \text{ mA}$  through  $1 \text{ pF}$ ! That's bad news for the op-amp  $U_1$ , as well as for the gate insulation of  $Q_1$  and  $Q_2$ . How can the input protection in Figure 5.80 be improved to ensure robust protection against such egregious insults? The input choke  $L_1$  helps, providing a series impedance that rises at high frequencies. But we can do better – see Figure 5.81.



**Figure 5.81.** Bulletproofing the amplifier frontend with slew-rate control, diode clamping, and input-current limiting.

The first step is to limit the incoming slew rate, with  $R_1 C_1$ : a  $500 \text{ V}$  step causes a maximum  $dV/dt = 5 \text{ V/ns}$  slew rate across  $C_1$ ,<sup>80</sup> with no significant degradation of bandwidth or noise. Looking at it most simply, that maximum slew rate can produce a transient current through the drain-source capacitance ( $C_{oss}$ ) of at most  $I = C_{oss} dV/dt = 10 \text{ mA}$ , thus a drop across  $R_2$  of at most  $10 \text{ V}$ . That is well below the gate-source rating of  $\pm 20 \text{ V}$ ; and it's a conservative estimate, because the actual gate-source voltage is

<sup>80</sup> And a momentary  $500 \text{ V}, 250 \text{ W}$  surge in  $R_1$ , which should be a bulk-composition type, or several SMT resistors in series, to handle both the voltage and the energy transient; see Chapter 1x.

reduced further by the relatively larger gate-source capacitance ( $C_{iss}$ ).

So  $Q_1$  and  $Q_2$  are safe. So is the amplifier, because the worst-case  $10 \text{ mA}$  transient current is clamped by diodes  $D_1$  and  $D_2$  to stay strictly within the amplifier's rails: those diodes clamp to bypassed voltages that are a diode drop inside the rails, set by diodes  $D_3$  and  $D_4$ . Note the  $\sim 1.5 \text{ mA}$  dc forward bias of the latter (we're assuming  $\pm 15 \text{ V}$  rails), maintaining their diode drop even with a continuous dc input overdrive (with corresponding current limited to less than  $1 \text{ mA}$ , owing to  $Q_1 Q_2$ ).

Finally, the small input resistor  $R_5$  is added for extra insurance, to limit any input current to the op-amp if the clamped voltage were to exceed the rails. This is more important with the usual arrangement of external clamp diodes (i.e., to the rails themselves), which allows the op-amp inputs to go a diode drop *beyond* the rails. Without the added series resistor it's then a current-divider contest between the external and the internal diodes.

This circuit looks pretty good, to us. Before shipping these to customers, though, you would be well advised to put it on the bench and beat it up, brutally. Circuits can surprise you.

## 5.16 Instrumentation amplifier miscellany

There's much to love about these amplifiers, whose nice variety is evident in the collection of Table 5.8 (with additional "programmable gain" types in Table 5.9 on pages 370–371). When you need some real performance, those tables are worth serious study! Here we collect some important advice on their care and feeding; refer often to Table 5.8 for enrichment as you read the following sections.<sup>81</sup>

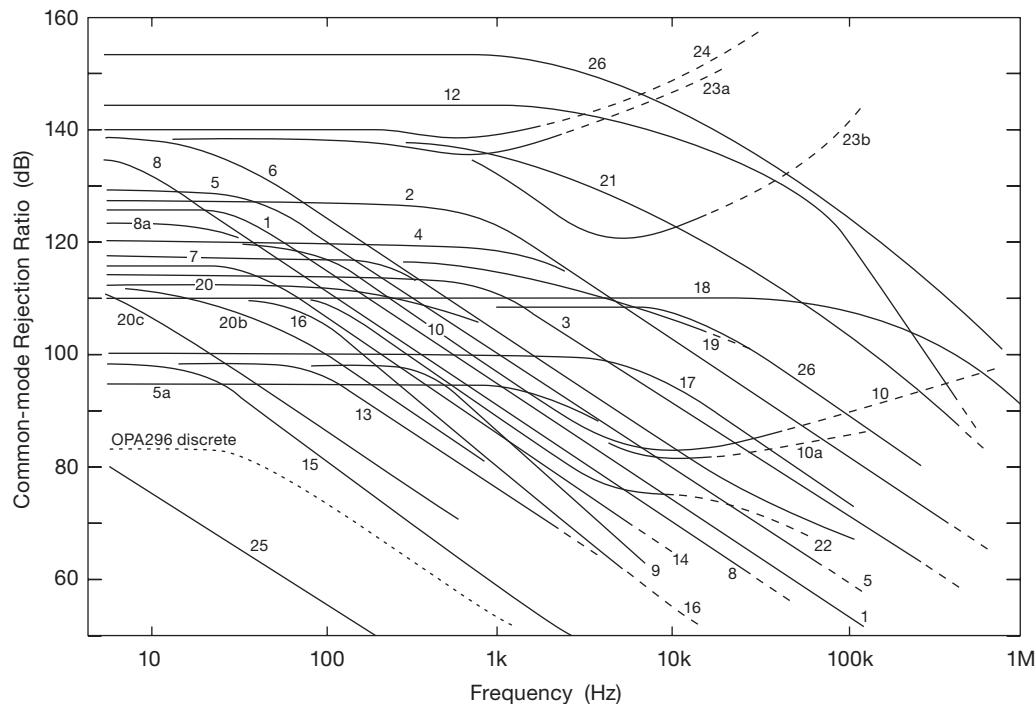
### 5.16.1 Input current and noise

Instrumentation amplifiers (INAs) must condition the input signal without disturbing it; thus high input impedance  $Z_{in}$ , low input current  $I_{in}$ , and a low level of current noise  $i_n$ . There's the usual tradeoff – the lower voltage offset and voltage noise of BJTs versus the lower input current and current noise of FETs. Some bipolar-input INAs (for example the LT1167/8) use the trick of bias-current cancellation to achieve sub-nanoamp input currents. Conversely, the auto-zero INAs do best in terms of offset voltage, but they pay a price in current noise and chopping artifacts.

For some amplifiers the input current noise is close to

<sup>81</sup> See also related material in the second edition, pp. 422-428.

**Table 5.8 Selected Instrumentation Amplifiers**



**Figure 5.82.** Common-mode rejection ratio versus frequency, for the instrumentation amplifiers listed in Tables 5.8 and 5.9. The “OPA296 discrete” curve is measured data of a two-op-amp discrete implementation (Figure 5.88, circuit B) shown in the AD627 instrumentation amplifier (curve 16) datasheet as testimonial to the superiority of integrated amplifiers. Curves 23a and 23b show the effect of filter choices (1 kHz and 10 kHz, respectively) for the same amplifier (the AD8553). The dashed line extensions indicate a region well past the amplifier’s cutoff frequency.

the shot-noise limit (as in §5.10.8), but this is greatly exceeded for auto-zero amplifiers, and BJT-input amplifiers that use bias-current cancellation.

### 5.16.2 Common-mode rejection

Instrumentation amplifiers must often deal with small difference signals riding on much larger common-mode voltages, requiring a high CMRR.

To get a sense of the problem, consider an INA used with a strain-gauge bridge powered from 5 V: the INA’s common-mode signal input is 2.5 V, with a typical full-scale output of 10 mV (i.e., 2 mV/V). So a signal that is 0.1% of full-scale is just  $10 \mu\text{V}$ . That’s  $-108 \text{ dB}$  with respect to the 2.5 V common-mode signal! Just seven of the INAs in Table 5.8 meet this spec for their minimum CMRR (though all but five meet it for their *typical* CMRR). An amplifier that fails to meet the CMRR spec will simply exhibit an output error larger than 0.1% of full scale, entirely down to insufficient CMRR. Keep this in perspective, though: most of the INAs in the table have more than

$10 \mu\text{V}$  of input offset voltage, anyway. So we can say for many applications the listed INAs do reasonably well (and some exceptionally well), at least at dc.

#### A. CMRR versus frequency

A much more severe test of an INA’s CMRR is its ability to reject common-mode signals at high frequencies. The plots in Figure 5.82 show this degradation, beginning somewhere in the range of 100 Hz to 5 kHz. In contrast to the (dc) strain-gauge application, imagine measuring the voltage across low-value current-sense resistors monitoring the windings of a three-phase motor. If the ac drive frequency is low enough (say 50 to 60 Hz), the INA may be up to the task. But if the motor windings are driven from a pulse-width-modulated (PWM) controller, with say 40 kHz pulses, high-frequency CMRR degradation may render an INA unusable for the task. So Figure 5.82’s curves may be useful not only for choosing the best INA for the job, but in fact to determine whether an INA may be used at all.

Some of the curves of CMRR versus frequency level off after falling at the usual 6 dB/octave. This happens because

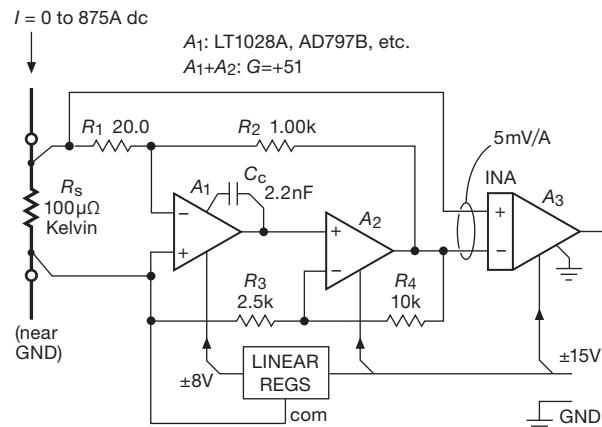
the INA includes provision for bandwidth filtering at an internal node (with an external capacitor) that attenuates the common-mode feedthrough by the same 6 dB/octave, thus cancelling out the input-stage CMRR degradation. For example, we see this for the AD8293 and AD8553 (curves 24 and 23a) with 1 kHz filtering, and for the latter and the INA337 (curves 23b and 10) with 10 kHz filtering. This is seen also for some micropower parts, with their very limited bandwidth, for example the MAX4194 (curve 10a), which runs at 90  $\mu$ A and has a 1.5 kHz bandwidth. That's where its response falls with  $G = 100$ . You can use this principle to advantage by adding a post filter in applications where fast response isn't necessary, for example the three-phase motor current monitor.

### B. Case study: increasing CMRR with upstream gain

Here's a nice example from our research lab, where we needed to control the current through a magnet coil with a high degree of stability, to bring about a "Bose-Einstein condensate" of cold atoms, and (in a first experiment of its kind) slow the speed of light to that of a bicycle.<sup>82</sup> The currents ranged up to 875 A (!), and we wanted something close to  $\sim 10$  ppm stability in the controlled current, which was sensed with a low-side 4-wire current shunt of 100  $\mu\Omega$  resistance (thus 87.5 mV full-scale signal<sup>83</sup>). With such high currents flowing, we had to deal with as much as a volt or so of common-mode signal, upon which the difference signal (87.5 mV maximum, but often much less) was riding. To control that current to 10 ppm thus required some 140 dB of CMRR, and of course very low input offset drift. It also required a low input noise voltage, ideally less than 0.1  $\mu$ Vpp integrated low-frequency noise in order to achieve 10 ppm.

The combination of very high CMRR and very low noise and offset is a tall order for any instrumentation amplifier. The solution here is to put some low-noise and low-drift gain upstream of the INA, as shown in Figure 5.83. A precision low-noise op-amp like the LT1028A (typical drift and low-frequency noise of 0.1  $\mu$ V/ $^{\circ}$ C and 35 nVpp) is wired as a  $G = +51$  composite amplifier ( $\S 4.x.5$ ), within which conventional op-amp  $A_2$  is configured with  $G = 5$ . The compensation capacitor  $C_C$  limits the bandwidth (and therefore the noise) to about 10 kHz. The precision op-amp

is run at lower voltage to reduce thermal effects (offset-voltage drift, thermal-gradient errors), with its  $\pm 8$  V rails regulated down (with floating 3-terminal linear regulators) from the ground-referenced  $\pm 15$  V that powers  $A_2$  and  $A_3$ .



**Figure 5.83.** Current sensing with low drift and low noise. Op-amps  $A_1$  and  $A_2$  form a composite amplifier, with a precision low-noise first stage and a noncritical second stage. The output signal provides feedback to the high-current power supply.

### 5.16.3 Source impedance and CMRR

Instrumentation amplifiers excel in high input impedance, but that does not confer automatic immunity from the effects of mismatched signal source impedances (which so severely degrade CMRR in *difference* amplifiers, with their relatively low input impedance; see  $\S 5.14.3A$ ). Most datasheets are bashful about displaying their dirty laundry in this regard, so we must applaud Analog Devices' candor, as shown in Figure 5.84. Note the greater effect at higher gain settings (where there's more to lose, because the CMRR is so good), and at higher frequencies (where the amplifier's input impedance is falling, owing to capacitance).

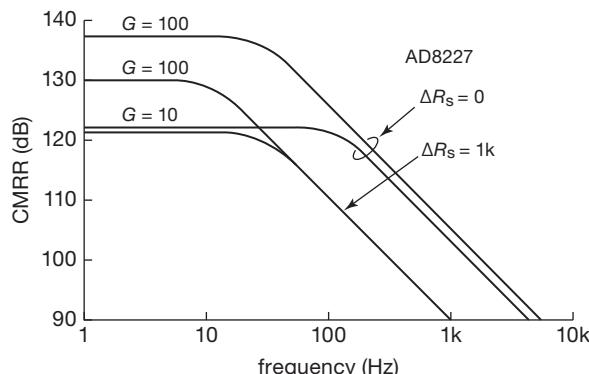
### 5.16.4 EMI and input protection

Whether you "roll your own" or use an integrated instrumentation amplifier, you've got to think about protection both from overloads and from electromagnetic interference (EMI). An example from real life: a colleague uses temperature-monitoring thermocouples in his lab, connected with the usual unshielded wire pairs to an INA front end running at high gain. All was well until a particular switching power supply was energized, at which point the thing went haywire.

The problem, of course, was common-mode EMI, coupled onto the long unshielded cable. And the solution, in

<sup>82</sup> L. Vestergaard Hau, S. E. Harris, Z. Dutton, and C. H. Behroozi, "Light speed reduction to 17 metres per second in an ultracold atomic gas," *Nature* **397** 594–598 (1999).

<sup>83</sup> A smallish voltage, but already an uncomfortably large power dissipation ( $\sim 75$  W), requiring a temperature stabilized oil bath.



**Figure 5.84.** Instrumentation amplifier CMRR, best at high gain settings, is degraded by source-impedance mismatch.

a situation like this where you don't need bandwidth, is to filter the input aggressively, as in Figure 5.85. The diode clamps are optional, but probably a good idea if you want the amplifier to survive an unusual input event. In situations where you *do* need bandwidths that include interfering signals, you're unlikely to get away with unshielded cable; use shielded wiring, and pay attention to grounding paths.

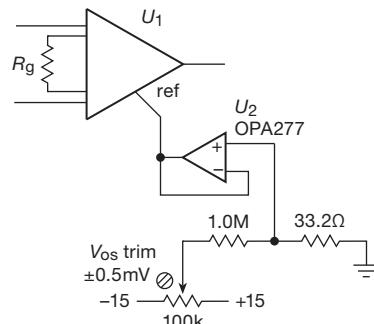
It is hard to overstate the seriousness of EMI/RFI: RF signals leaking into your inputs cause rectification inside BJT op-amps, and thus dc offsets. Cables or PCB traces can exhibit narrowband (high- $Q$ ) resonances, enhancing these effects. If you notice changes of offset voltage when you grip a cable, or touch a circuit node with a pencil, or just wave your hands around, you're probably seeing RF coupling (the other possibility is a circuit oscillation). Lossy ferrite beads are excellent for attenuating RF, as well as for reducing the  $Q$  of unwanted wiring resonances; but they're not a panacea, and often you need to resort to additional filtering.

### 5.16.5 Offset and CMRR trimming

Instrumentation amplifiers that provide both SENSE and REF pins (e.g., the INA103) can be externally trimmed, if needed, for both offset voltage and CMRR, as shown previously for the difference amplifier (Figure 5.74). More often you're given only the REF pin, though. That's enough to trim the offset voltage, but note that any offset applied to the REF pin must have a source impedance of no more than  $\sim 10^{-6} R_f$  (i.e., a few milliohms) in order not to compromise the amplifier's 100 dB+ CMRR. This is best done with a precision op-amp, as in Figure 5.86.

Some INAs helpfully include output-stage offset trim terminals, to be used with an external trimpot. A few of

these (e.g., the INA110 and others) even provide separate trim pairs for the front-end and output stages.



**Figure 5.86.** Trimming the offset of an INA that doesn't provide a SENSE pin, offset pins, or a buffered REF pin.

### 5.16.6 Sensing at the load

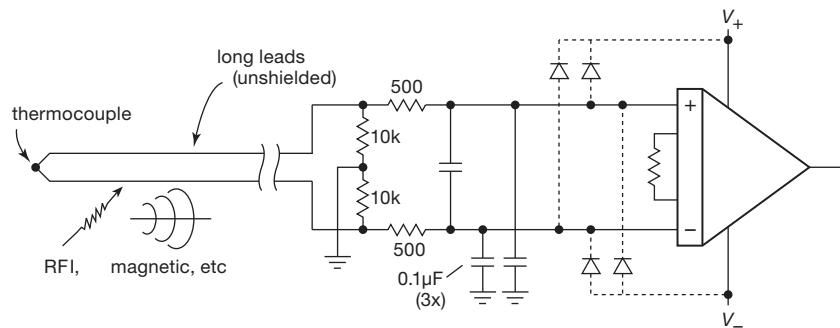
As with the difference amplifier, the REF and SENSE pins can be connected directly at the load, as in Figure 5.78(a), to eliminate errors from wiring resistance and unrelated ground currents.

### 5.16.7 Input bias path

Instrumentation amplifiers earn honest bragging rights from their very high input impedance, but, just as with op-amps, you've got to provide a dc return path. If you don't, the amplifier will saturate. This occurs naturally in circuits like the strain gauge of Figure 5.64A, but not in something like a thermocouple (Figure 5.85). For the latter you can use a resistor from one of the inputs to ground (or to mid-supply, for a single-supply amplifier), or you can use a bias resistor from each input to ground to preserve symmetry.

### 5.16.8 Output voltage range

If an INA is operated at low gain and with its common-mode input near the rails (but legally within the specified operating range), the internal amplifier may saturate, causing the INA's output to go to an incorrect voltage. For example, look at the "Maximum Output Voltage vs. Common-Mode Input" graph in the datasheet for the AD623 (this amplifier is like type A in Figure 5.88, except with *pnp*-input emitter followers). If the common-mode input is at 0 V with  $G=10$  (which is legitimate), the maximum output capability is only 1.0 V! You have been warned!



**Figure 5.85.** Electromagnetic interference couples nicely into long unshielded cables, which you can think of as antennas. Use lowpass filtering at high-gain inputs, with optional diode clamping to the rails. Note the  $10\text{k}\Omega$  resistor pair, to set the dc level of the floating differential input.

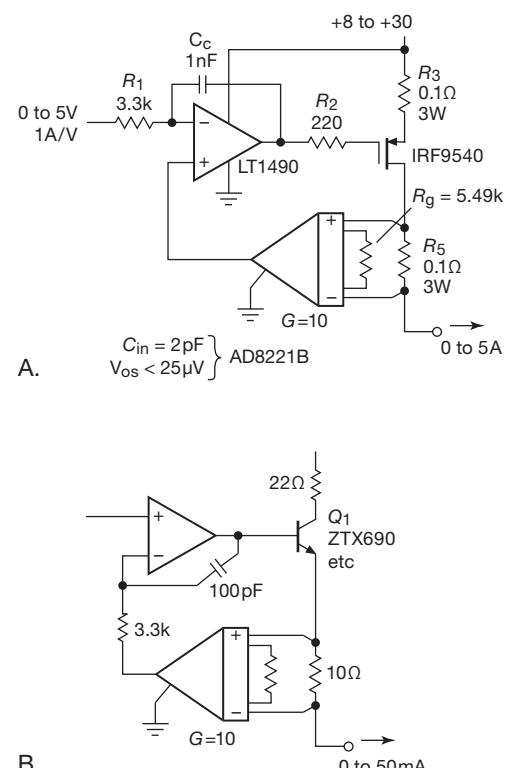
### 5.16.9 Application example: current source

The excellent CMRR of instrumentation amplifiers, combined with very low input capacitance (typically  $\sim 2\text{ pF}$ ), lets you design an active current source in which the current sense resistor rides on the high side of the output, with a “flying” INA converting its voltage drop to a ground-referenced output. Such a circuit is shown in Figure 5.87A, exploiting the  $>80\text{ dB}$  CMRR of the AD8221B out to  $50\text{ kHz}$ , where the stabilizing compensation network  $R_1C_C$  rolls off. The MOSFET, configured as a common-source transconductance stage, has inherently high output impedance, which makes it possible to maintain good current-source behavior at higher frequencies. This improves upon a current-source circuit like that in Figure 5.69 (§5.14.2D), in which the performance degrades with frequency owing to the op-amp’s compensation and limited slew rate. In this circuit the source-degeneration resistor  $R_3$  acts to reduce the MOSFET transconductance, to enhance stability (some breadboard testing, with anticipated load impedances, would not be inadvisable). See Table 3.4 for selected *p*-channel MOSFETs.

For this circuit the op-amp output must be able to swing to the positive rail (the LT1490 is RRIO), and the AD8221B needs a few volts of negative supply voltage ( $-3\text{ V}$  is adequate) because its common-mode input voltage does not extend to the negative rail.<sup>84</sup>

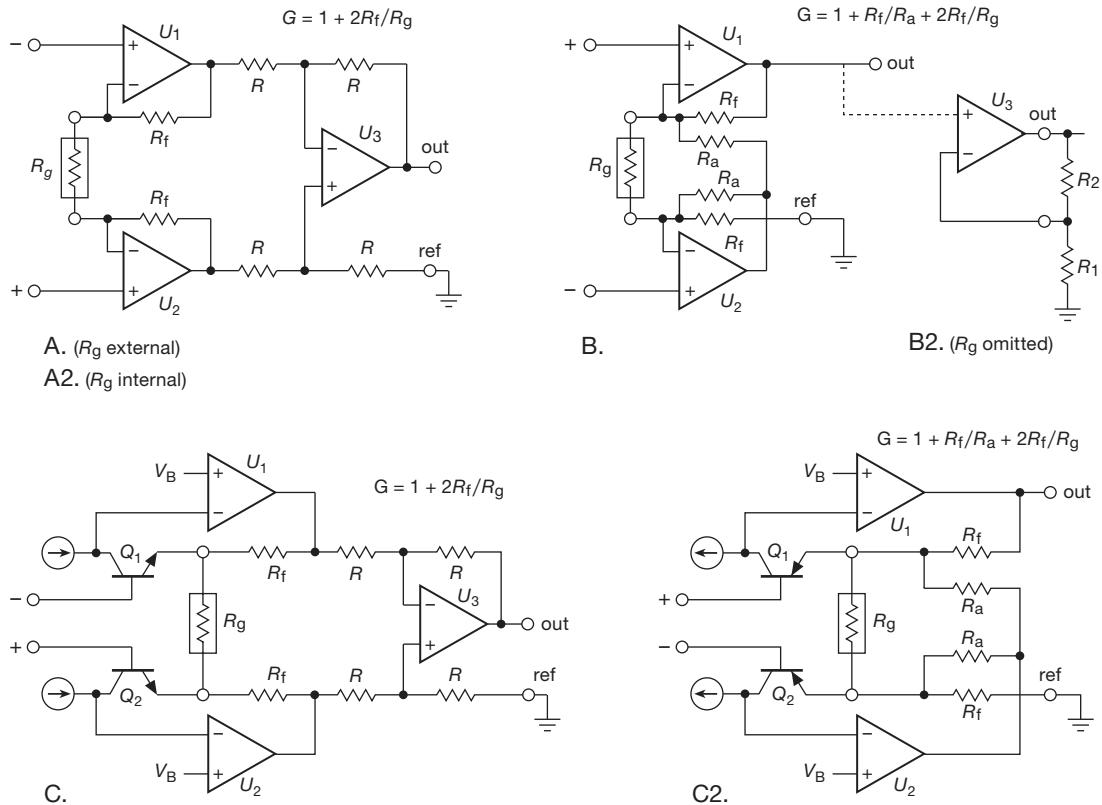
The AD8221B’s low offset voltage ( $25\text{ }\mu\text{V}$  max) provides a large dynamic range, corresponding to an output error of just  $0.25\text{ mA}$  out of the full-scale range of  $5\text{ A}$  ( $20,000:1$ ).

For lower output currents you can substitute a smaller bipolar transistor, as in Figure 5.87B; its lower capacitance permits greater loop bandwidth (we’ve configured it here as a follower). The AD8221B’s low input



**Figure 5.87.** Precision current source with flying instrumentation amplifier. A. Power MOSFET for  $5\text{ A}$  full-scale current, B. a small high-gain ( $\beta \sim 500$  over  $0.1\text{--}50\text{ mA}$ ) BJT provides greater bandwidth.

<sup>84</sup> The AD8227 variant allows  $V_{CM}$  to the negative rail, so you could run it single supply; but you pay a price in larger  $V_{OS}$  and  $I_B$ , and its CMRR degrades at a lower frequency.



**Figure 5.88.** Instrumentation amplifier configurations A–C, as listed in Table 5.8 (page 363). For these and other differencing circuits the “ref” pin need not be grounded.

current (0.4 nA max) means that you can scale down the full-scale output current range, say to 100  $\mu$ A full scale.

### 5.16.10 Other configurations

The classic 3-op-amp circuit of Figure 5.77 is widely used, notably in the INAs offered by Burr-Brown/TI (recognizable by their part numbering “INA<sup>nnn</sup>”); but you’ll see other circuit configurations (if you get far enough into the datasheets) representing different tradeoffs among the various performance parameters and cost. Although you can get along fine without “looking under the hood” (most of what you need to know comes from the tabulated data), some of these configurations have unusual quirks that can catch you off-guard. For example, amplifiers with configuration E (Table 5.8 on page 363, and Figures 5.88 and 5.89) can be damaged by input voltage differentials greater than  $\pm 0.5$  V (!),<sup>85</sup> and the amplifiers with configuration F do not

operate with the REF tied to ground (even though intended for low-voltage single-supply operation). Quite apart from these kinds of worries, curiosity impels a brief look at these circuits.

There’s a general principle in play in nearly all of these circuits (D and E excepted): (a) the voltage across a gain-setting resistor  $R_g$  is precisely the same as the input difference voltage, creating a current  $I_g = \Delta V_{in}/R_g$ ; and (b) that current is used to generate an accurately proportional output voltage  $V_{out} \propto I_g$ . Classic configuration A puts this in clear view: the input op-amps (or equivalent – they need not be fully featured op-amps) enforce a matched  $\Delta V_{in}$  across  $R_g$ , with the resultant current flowing through the two  $R_f$ ’s; thus a differential output  $\Delta V_{out} = (\Delta V_{in}/R_g)(R_g + 2R_f)$ . The

<sup>85</sup> Some of these have back-to-back clamp diodes across the inputs (true also of some op-amps and comparators), for which the damage is

caused by excessive input current; others tolerate greater differential inputs (as much as a few volts), though most of the E-configuration parts are considerably more restrictive than those with the other circuit topologies. What matters more, from the user’s point of view, is the maximum differential input without performance degradation.

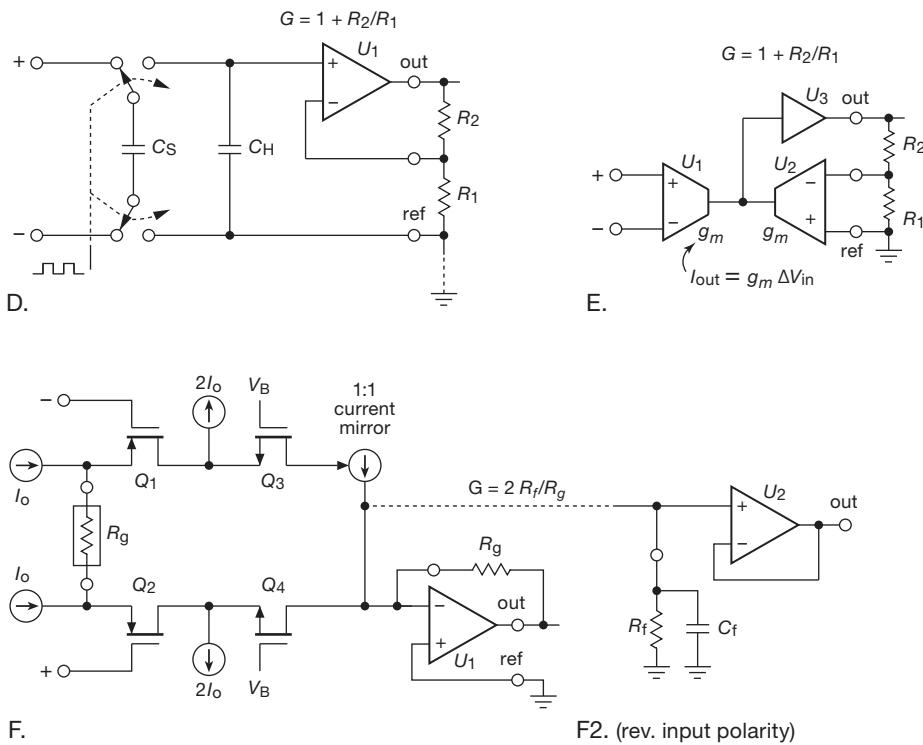


Figure 5.89. Instrumentation amplifier configurations D–F, as listed in Table 5.8 on page 363.

unity-gain difference amplifier converts this to a single-ended output, with gain  $G = 1 + 2R_f/R_g$ .

Configuration C works similarly, but here the matched emitter followers  $Q_1Q_2$  create the  $\Delta V_{in}$  replica across  $R_g$ , with the op-amps serving to ensure equal emitter currents (and thus no contribution to the differential output).<sup>86</sup> In this circuit and following,  $V_B$  is a reference “bias” voltage.

One nice thing about the C and C2 configurations is that a small RFI-defeating capacitor can be placed across base-to-emitter of the input transistors, because they’re exposed pins. Keep those capacitors small – 100 pF or less – so the amplifier’s bandwidth and stability are not degraded.<sup>87</sup>

The clever configuration B is different: it is more economical, requiring only two op-amps and fewer trimmed

resistors, but its performance suffers, with poorer CMRR (particularly at higher frequencies). (The reader also suffers, trying to figure out how this tangle of a circuit works.) Configuration C2 is the discrete differential-pair analog of B (just as C was for A), with similarly underwhelming specifications.

Configuration F continues the theme of replicating  $\Delta V_{in}$  across  $R_g$ , with the resultant unbalanced current at  $U_1$ ’s summing junction being converted to a single-ended output. In this circuit  $Q_3$  and  $Q_4$  form a “folded cascode,” pinning the drains of  $Q_1$  and  $Q_2$  while passing their currents through (offset by twice the quiescent current, thus sinking). This circuit requires accurate matching of the current sources and of the current mirror (stated more precisely, it requires constancy of the current sources and mirror over common-mode variations). Evidently that can be accomplished with good design (and the assistance of circuit tricks like the cascode), given the impressive 140 dB (typ) CMRR specification.<sup>88</sup>

<sup>86</sup> To achieve low input currents with the C configuration, LTC uses superbeta BJTs with base-current cancellation in some of the listed parts ( $I_B \approx 50$  pA); Analog Devices does even better using JFETs, but with greater offset and noise. Some of the TI/Burr-Brown INAs listed as A types may in fact use the C configuration; their datasheets are silent on the circuit details.

<sup>87</sup> BJT-input amplifiers are prone to RFI upset, because their inputs are forward-biased base-emitter (diode) junctions. And RFI is a real problem in these low-level circuits with inputs from remote sensors. Better to use a JFET-input amplifier if you are bedeviled by RFI.

<sup>88</sup> Because they contain only a few MOSFETs, current sources, and current mirrors, devices of configuration F can be quite inexpensive. For example, the AD8293 (an AZ with fixed  $G=80$  or 160) sells for only \$0.97 (qty 100).

Table 5.9 Selected Programmable-gain Instrumentation Amplifiers

Part #	Input device	circuit	Supply Voltage		I <sub>s</sub> min max <sup>e</sup>	shutdown <sup>f</sup>	Input Current		Offset Voltage		ΔV <sub>os</sub> typ <sup>g</sup>	Gain Error <sup>d</sup> typ	CMRR <sup>x</sup> G=100 <sup>d</sup>		Noise <sup>t</sup> curve <sup>o</sup>		
			(V)	(V)			(mA)	(nA)	(μV)	(μV)			typ	min			
<b>AD8250</b>	BJT	A2	10	34	4.1	-	5	90	260	1.7	0.04	5G	110	98	4	1	18
<b>AD8251</b>	BJT	A2	10	34	4.1	-	5	95	275	1.8	0.04	5G	110	98	4	1.2	18
<b>AD8253</b>	BJT	A2	10	34	4.6	-	5	-	160	1.2	0.04	4G	120	100	26	0.5	10
<b>PGA204B</b>	BJT	A2	9	36	5.2	-	0.5	10	50	0.1	0.024	10G	123	110	8a	0.4	13
<b>PGA202B</b>	JFET	C	12	36	6.5	-	0.01	500	1000	12	0.15	10G	120	92	3	1.7	12
<b>PGA207</b>	JFET	C	9	36	12	-	0.002	500	1500	2	0.05	10T	100	95	3	1	18
<b>PGA280</b>	BJT	A2	10	40 <sup>f</sup>	0.75	-	0.3	50 <sup>g</sup>	250	0.2	0.15	1G	140	120	5	0.42	22
<i>low-voltage</i>																	
<b>LTC6915</b>	switch	D	2.7	11	0.9	x	5	3	10	0.05	0.5	cap <sup>k</sup>	125	105	1	2.5	225
<b>AD8231</b>	BJT	A2	3	6	4	x	0.25	4	15	0.01	0.8	high	-	110	3	0.7	39 <sup>h</sup>
<b>LMP8358</b>	CMOS	E	2.7	6	1.9	x	0.006	1	10	0.05	0.1	50M	139	110	2	0.6	27 <sup>h</sup>
<b>PGA309</b>	CMOS	A2	2.7	7	1.2	x	0.1	3	50	0.2	1	30G	105	-	-	4	210

Notes: (a) at 25°C. (b) at 100Hz. (c) 0.01–10Hz or 0.1–10Hz, at G=100 if available. (d) at G=100 or max, not 1000; no source imbalance. (e) abs max. (f) separate LV output rails. (g) at G=max. (h) plus HF auto-zero noise. (k) caution: high bias currents. (m) plus filter stage and output clamp. (n) to within 50mV–200mV of rail. (o) see Fig. 5.82. (p) rail-to-rail input, provided AGND away from rails. (q) 250kHz chopper. (r) V<sub>out</sub> to V<sub>EE</sub>, but V<sub>ref</sub>=0.8V min. (s) 2V/μs at G=1000.

Parts using configurations D and E are unlike the others. In D a flying capacitor  $C_S$  periodically samples and conveys to the “hold” capacitor  $C_H$  the differential input voltage, thus creating a single-ended (ground-referenced) replica. This sounds good in principle; but the resulting noise is high, and the slow commutating rate (3–6 kHz, for the two exemplars in Table 5.8 on page 363) limits the bandwidth and extends the settling time. This technique is also vulnerable to aliasing at half the chopping frequency. However, these parts are inexpensive, and they may be well suited for some dc applications.

Finally, in E the output currents from a pair of differential-input transconductance amplifiers are combined and forced to equality: one amplifier sees the input signal pair, and the other sees a divided fraction of the output, producing a single-ended output voltage, as shown. This low-cost configuration (no laser-trimmed resistor pairs, etc.) is limited to small differential input signals (thus high gain), and generally exhibits relatively poor gain accuracy.<sup>89</sup> Interestingly, the fastest amplifier in Table 5.8, by far, uses this configuration.

### 5.16.11 Chopper and auto-zero instrumentation amplifiers

The same auto-zero techniques that are used in “zero-offset” CMOS op-amps (§5.11) are used in some CMOS low-voltage instrumentation amplifiers. These are recognizable by their very low offset-voltage specifications, down in the tens of microvolts, where non-AZ CMOSs do not venture (see Table 5.8 on page 363). These amplifiers attain excellent CMRR also, but they pay the price in broadband noise, switching-frequency noise,<sup>90</sup> and (sometimes) input bias and noise currents. The amplifiers are useful particularly in low-frequency applications, for example as input stages for integrating ADCs (see Figures 13.67), or combined with lowpass filtering.

### 5.16.12 Programmable gain instrumentation amplifiers

You set the voltage gain of a simple op-amp circuit with external resistors. In contrast, *difference* amplifiers (Table 5.7 on page 353) ordinarily are configured with fixed gain, set by a precision internal matched resistor network. And the gain of *instrumentation* amplifiers (Table 5.8 on page 363) is ordinarily set by a single gain-setting resistor  $R_g$ . Note, though, that some of the amplifiers in Table 5.8 include several internal  $R_g$  gain-setting resistors, allowing accurate gain selection with only an external jumper wire. Taking this one step further, you can get *programmable-*

<sup>89</sup> Some E types (e.g., the AD8130, a variant of the AD8129 in Table 5.8) are specified and characterized only for G=1. These are especially useful as differential line receivers, etc., but they are generally limited in swing, typically in the range of 3 to 4 V<sub>pp</sub> (the AD8237, with its flying switched capacitors, is an exception). See also §12.10 for a discussion of differential signaling in the *digital* context.

<sup>90</sup> Which may not be evident from the datasheets, which sometimes omit spectral noise plots.

**Table 5.9 Selected Programmable-gain Instrumentation Amplifiers (cont'd)**

Part #	Gain choices	BW <sup>d</sup>	Slew	Settled <sup>d</sup>	Swing to Supplies?				filter? out sense	Interface	Comments	Cost (\$ US)	
		-3dB (MHz)	-3dB Rate (V/μs)	0.01% (μs)	IN +	OUT <sup>v</sup> -	mux?						
<b>AD8250</b>	1,2,5,10	3	25	0.65	-	-	-	-	-	pins	fast settle to 10ppm	7.78	
<b>AD8251</b>	1,2,4,8	2.5	25	0.68	-	-	-	-	-	pins	fast settle to 10ppm	7.78	
<b>AD8253</b>	1,10,100,1000	0.55	20 s	1.5	-	-	-	-	-	pins	fast settle to 10ppm	7.68	
<b>PGA204B</b>	1,10,100,1000	0.01	0.7	100	-	-	-	-	•	pins	PGA205 for 1,2,4,8	17.09	
<b>PGA202B</b>	1,10,100,1000	1	20	2	-	-	-	-	• •	pins	PGA203 for 1,2,4,8	14.34	
<b>PGA207</b>	1,2,5,10	0.6	25	3.5	-	-	-	-	•	pins	PGA206 for 1,2,4,8	18.36	
<b>PGA280</b>	1/8–128, by x2 <i>low-voltage</i>	0.05	2	40	-	-	f	f	2	-	SPI	output supply 2.7–5V	6.46
<b>LTC6915</b>	1,2,4...4096	slow	0.2	5ms	x	x	x	x	-	• •	SPI,pins	chopper, high-gain	4.58
<b>AD8231</b>	1,2,4...128	0.07	1.1	slow	n	n	x	x	-	-	pins	auto-zero	3.08
<b>LMP8358</b>	10–1k, by 1–2–5	0.68	6.5	4	-	x	x	x	-	• -	SPI,pins	auto-zero, prog filter	5.27
<b>PGA309</b>	4–128, 2.7–1152	0.4	0.02	slow	-	n	n	x	•	-	SPI	sensor interface <sup>U</sup>	7.05

(t) typical. (u) 148-page user manual; includes ADC for sensor span, offset tempco cal, etc.; parameters stored in external SOT23 EEPROM. (v) caution: the RR output parts often don't allow the REF pin to be close to  $V_{EE}$ ; be sure to check the datasheet! (w) part with large gain error assume you'll perform gain calibration. (x) the CMRR spec is typically at 60Hz; check the plots if you care about performance at higher frequencies.

gain amplifiers (PGAs) in which that selection is made by a digital input code (either applied as a parallel logic-level code to a set of pins, or as a multibit serial code through a serial port like SPI or I<sup>2</sup>C; see Chapters 14 and 15). These are, in essence, integrated versions of the discrete digitally programmable instrumentation amplifier of Figure 5.80.

Some examples of stand-alone instrumentation PGAs are the PGA204/5, LMP8358, and PGA280, with more listed in Table 5.9. The PGA202/3 (JFET) and PGA204/5 (BJT) are traditional “high-voltage” (to  $\pm 18$  V) parts that accept a 2-bit parallel code (on two pins) to select gains of 1, 10, 100, or 1000 (PGA202/4) or 1, 2, 4, or 8 (PGA203/5). The more recent LMP8358 is a low-voltage single-supply part (2.7–5.5 V) with auto-zero and gains from 10 to 1000 in a 1–2–5 sequence (i.e., 10, 20, 50, 100, 200, 500, 1000), and with ambidextrous programming – the three pins can be used either as a 3-bit gain-setting parallel port or a 3-wire SPI serial port that programs both the gain and some additional parameters such as input polarity reversal, fault detection, and bandwidth. It is fast (8 MHz) and accurate ( $V_{OS} = 10 \mu\text{V}$ , max).

Finally, the very elegant PGA280 addresses the need for a PGA whose input signals can range over  $\pm 10$  V or more, but with a separately powered output stage matched to contemporary low-voltage single-supply ADCs and microcontrollers. The inputs can range over  $\pm 15.5$  V (with  $\pm 18$  V supplies), with the output supplied from the same +2.7 V

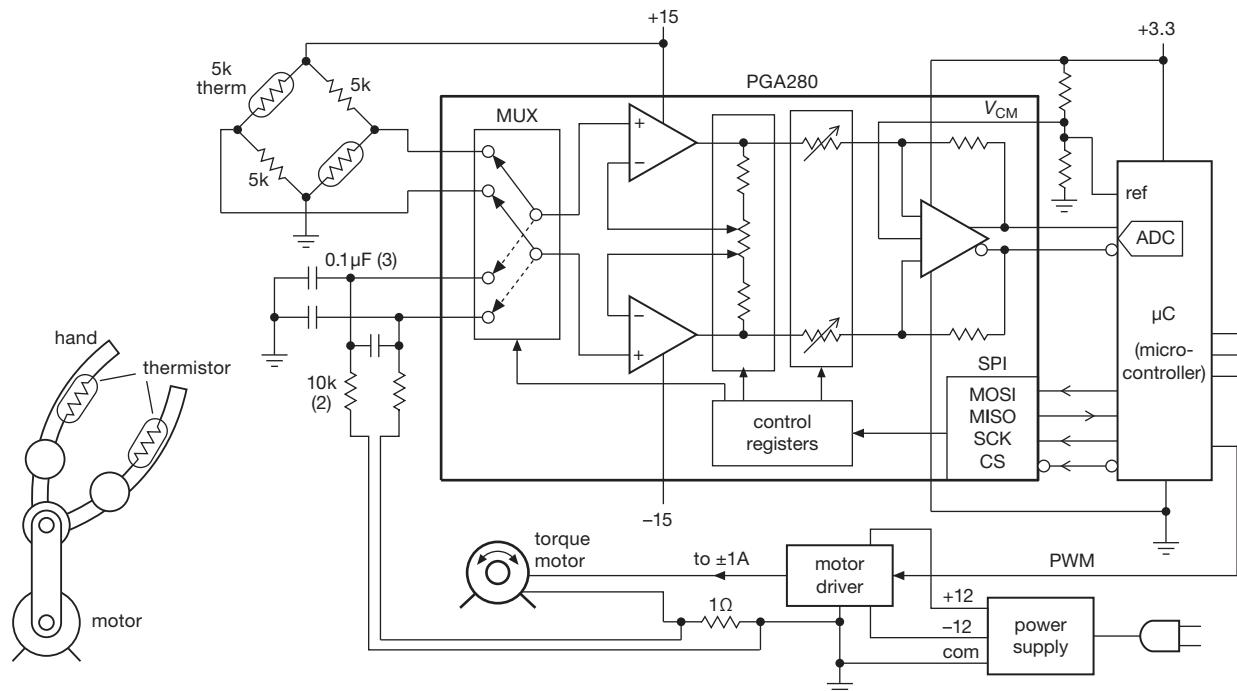
to +5 V supply that powers the ADC or  $\mu\text{C}$ .<sup>91</sup> This solves the messy problem of protecting the inputs of a low-voltage part when it is driven from an IC running from higher voltages. There's a REF-like pin that sets the output mid-span voltage; the output is actually *differential*, with a pair of complementary outputs, but you can ignore one and treat it as single-ended (with a slight loss of accuracy).

This amplifier has excellent performance: it programs through a digital serial port, with selectable gains from 1/8 to 128 by factors of two. It combines low offset voltage (auto-zeroing:  $\pm 15 \mu\text{V}$  max, at  $G = 128$ ), high input impedance ( $> 1 \text{ G}\Omega$  typ), low gain drift ( $\pm 3 \text{ ppm}/^\circ\text{C}$  max), and excellent CMRR (gain-dependent: 130–140 dB typ). Among its other tricks are an on-chip 2-input multiplexer (two differential-input pairs), an uncommitted byte-wide bidirectional digital port, and various signal conditioning and fault-detection options.

Figure 5.90 shows the sort of application for which this part would be ideal: our colleagues have developed an experimental robotic grasping hand<sup>92</sup> that is driven by a single torque motor through a set of passive linkages and couplings. For control you'd like to know the applied

<sup>91</sup> Or you can power the output from a split supply, staying within that total supply range.

<sup>92</sup> See, for example, Dollar and Howe, “The Highly Adaptive SDM Hand: Design and Performance Evaluation”, *International Journal of Robotics Research* **29**, (5), 585–597 (2010), available at the web page of The Harvard BioRobotics Laboratory: [biorobotics.harvard.edu](http://biorobotics.harvard.edu).



**Figure 5.90.** Two-channel level-shifting PGA application: reading out the applied torque and thermal response of a grasping robotic hand. This chopper-stabilized (250 kHz) instrumentation PGA works from  $\pm 15$  V rails, but delivers its output to a 3.3 V ADC – that’s being helpful! Recommended.

torque (from the motor current), and something about the contact pressure (if any) with the object being grasped. The PGA280’s wide input compliance and range of gains, combined with 2-channel differential inputs, make this an easy task: we imagine a pair of thermistors, self-heated a few degrees above ambient, that sense the temperature at the grasping points, and a low-side sense resistor in the motor return line. The low-voltage output stage connects to the on-chip ADC of a microcontroller, which controls the channel switching and gain of the PGA. We’ve rigged the thermistors so that first contact makes an upward step in differential voltage, followed by a second step in the same direction when the other “finger” makes contact. (The step size provides further information about the object’s material.) Note the differential “4-wire” connection at the motor-current-sense resistor, to eliminate errors from ground connection resistance; we’ve filtered that signal, because the use of pulse-width modulation tends to create high-frequency noise.

When looking around for a good instrumentation amplifier, be sure to consider specialized PGA variants that are intended to serve as front ends for low-level sensors and the like, for example the PGA309<sup>93</sup> or PGA2310.<sup>94</sup>

PGAs are quite popular as portions of more complex ICs, for example within ADCs (Figure 13.67 again, and Figures 13.70 and 13.71) and microcontrollers (Figure 15.10).

### 5.16.13 Generating a differential output

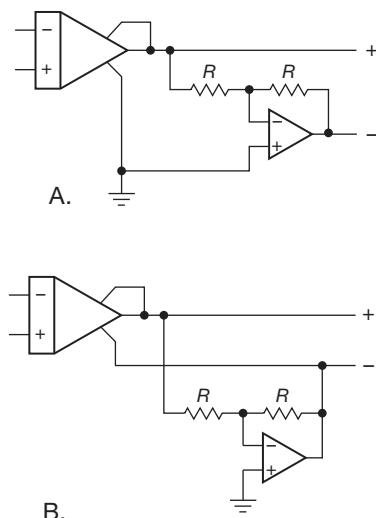
Both instrumentation amplifiers and difference amplifiers are used to convert a differential input signal to a single-ended output. That’s fine, if that’s what you want (and it usually is). There are some situations, though, in which you need a differential output signal, for example when

<sup>93</sup> Whose 148-page User’s Guide describes it thus: “The PGA309 is a smart programmable analog signal conditioner designed for resistive bridge sensor applications. It is a complete signal conditioner with bridge excitation, initial span and offset adjustment, temperature adjustment of span and offset, internal/external temperature measurement capability, output over-scale and under-scale limiting, fault detection, and digital calibration.”

<sup>94</sup> “The PGA2310 is a high-performance, stereo audio volume control designed for professional and high-end consumer audio systems.”

driving several varieties of analog-to-digital converter (ADC, a major subject of Chapter 13). Most simply, this can be done by adding a unity-gain inverter to the single-ended output, as in Figure 5.91A.<sup>95</sup> This works, for sure, but the gain accuracy will be degraded unless the resistor pair is matched to at least the precision and stability of the driving amplifier. This pitfall is circumvented in circuit B, where the unity-gain inverter forces the SENSE output pin to a symmetrical voltage relative to ground (or some other reference voltage). With this circuit the gain accuracy is retained; the effect of any resistor mismatch is simply to offset the symmetry of the outputs about ground (or reference voltage), which is generally of less importance because of the differential-input nature of the device being driven.

Both circuits, however, share the drawback of introducing a time delay (or phase shift), owing to the finite bandwidth of the inverting stage. One solution is to use a pair of matched amplifiers, as in Figure 5.66F. But a better way, particularly when plenty of bandwidth and rapid settling is needed (as with fast ADCs), is to use a *differential amplifier* (or *fully differential amplifier*, to emphasize the distinction), a term that has come to mean an amplifier with differential inputs *and* outputs; see §5.17. The PGA280 of the previous section is such an amplifier (though its designers officially call it an instrumentation amplifier).



**Figure 5.91.** Generating a differential output from an instrumentation amplifier or difference amplifier. Method B maintains gain accuracy.

<sup>95</sup> See also the parts listed under “single-ended to differential” in Table 5.10 (page 375).

## 5.17 Fully differential amplifiers

The term “fully differential amplifier” (or sometimes “differential in/out amplifier,” or just “differential amplifier”) is used to describe an amplifier with differential input *and* differential output, along with an additional input pin (“*V<sub>OCM</sub>*”) that sets the common-mode voltage of the output pair. We favor the “fully” term, to distinguish clearly from difference amplifiers and instrumentation amplifiers, both of which have single-ended outputs.

For some important applications you need to create a balanced differential *output*, from either a differential or a single-ended input signal. This is often the case with ADCs that have complementary inputs; see Figure 13.65 (a “charge-redistribution” ADC), Figure 5.102 (pipeline flash), Figure 13.28 (flash), Figure 13.37 (SAR), and Figure 13.68 (delta-sigma). For that application the important performance parameters are likely to be settling time, gain accuracy and stability, the ability to set the common-mode output voltage, and the ability to drive a rail-to-rail swing into a low-voltage ADC.

Other applications where differential signals are widely used include analog signaling over twisted pairs (e.g., through existing Cat-5-style network cable); telecom applications such as ADSL and HDSL links; oscilloscope input stages; and RF communications subcircuits such as IF and baseband blocks.

You can, of course, create a differential-output signal pair by using single-ended amplifiers (op-amps, difference amplifiers, and instrumentation amplifiers), as illustrated in Figures 5.66F, 5.70, 5.91, and 13.37. But you do better, particularly in terms of speed and noise, with an integrated differential amplifier, which also lets you set the common-mode output voltage (i.e., the midpoint of the output swing); this capability is particularly useful when driving differential-input ADCs powered from a single supply, because of their fussy insistence on common-mode input voltages.

Table 5.10 on page 375 includes a good selection of currently available differential amplifiers, keyed to the circuit diagrams in Figures 5.94–5.98. These illustrate the unflagging creativity of the human species, and some nice circuit tricks as well.

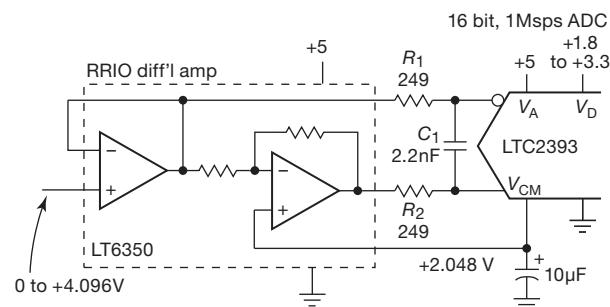
Let’s take a quick tour. Circuit A, intended for single-ended inputs, is simply a differential amplifier “kit,” with an input op-amp whose gain you set with external resistors. You can wire it as a noninverting amplifier (therefore high input impedance), or you can configure it as an inverting amplifier (for example, to handle a large input swing by setting the gain to less than unity). The noninverting

input of  $A_2$ , conveniently of high impedance, lets you set the output common-mode output voltage. The LT6350 is a low-noise, low-distortion amplifier of this configuration, with the added feature of rail-to-rail outputs.<sup>96</sup> Figure 5.92 shows how you would use it to drive an ADC, in this case the LTC2393 that cooperates by providing a mid-scale dc reference output ( $V_{CM}$ ).<sup>97</sup> The amplifier runs from the same +5 V and ground supply voltages, which eliminates the frequent worry about driving the ADC's input clamp diodes into conduction. The input lowpass filter to the ADC ( $R_1 R_2 C_1$ ) serves two functions: (a) it is an anti-alias filter, limiting the input bandwidth to  $\sim 150$  kHz; and (b) it provides the recommended shunt input capacitance to suppress the effects of the ADC's internal switching transients (which afflict many ADCs, including “charge redistribution SAR” converters, and delta-sigma converters – see Chapter 13).

Circuit B is a symmetrical balanced configuration optimized for pro-audio: well-balanced high-level drive ( $>15$  Vrms) with low distortion into balanced pairs, and stability into the load capacitances you get with long cables (to 10,000 pF or more). Figure 5.93 shows a typical application, in this case generating the high-level low-distortion balanced output needed for pro-audio cable driving, first discussed in §5.14.2E. Of note here is the very high *common-mode* output impedance, which preserves signal balance by allowing the receive end to override the driver's default (symmetrical about ground at the driver end, set weakly by the 10k resistors). In fact, it's even OK to ground one side at the receive end, in order to generate a single-ended signal there. We'll have more to say about this in §5.17.1.

Circuit C is the configuration of Figure 5.66F, with high-impedance (buffered) common-mode voltage-setting inputs ( $V_{OCM}$ ). The table entries with this configuration are low-noise wideband amplifiers, good for video and communications applications.

Circuit D is a popular configuration, sometimes with internal gain-setting resistors, sometimes external. The dif-



**Figure 5.92.** ADC driver circuit using a configuration A differential amplifier. The ADC provides a mid-span  $V_{CM}$  output reference, here used to set the amplifier's common-mode level. See also the better-performing AD4922.

ferential amplifier within the feedback loop consists of a symmetrical pair of transconductance amplifiers (voltage-to-current) generating a voltage across a resistive load, with voltage followers to generate the low-impedance output pair. The  $V_{OCM}$  input lets you assert the common-mode output voltage, which otherwise defaults to mid-supply (in which case it's good to attach a bypass capacitor). The  $V_{OCM}$  input bandwidth typically is comparable to that of the amplifier.

Circuit E continues the theme of differential transconductance amplifiers, but here they are configured in a feedback arrangement, the differential-output version of the analogous circuit E in Figure 5.89. Circuit F, used in the fastest amplifiers in the table, blends the output configuration of D with an input configuration like that of the classic instrumentation amplifier (Figure 5.65), again with differential transconductance amplifiers as the gain element.

Finally, circuit G is a completely different animal, a pair of offset-cancelled followers biased from the  $V_{OCM}$  input pin. This configuration, used in the table's 2 GHz entry, is intended for ac-coupled (or transformer-coupled) inputs.

### 5.17.1 Differential amplifiers: basic concepts

**A. Gain** The differential voltage gain is unity for most configurations in which the gain is set by matched pairs of feedback resistors  $R_f = R_g$ . The gain range column in Table 5.10 (page 375) identifies parts with fixed gains, minimum gains, or a set of gain selections. In some cases the exact gain is affected by source-impedance and termination-matching issues; see §5.17.4 and the formulas in Figure 5.104.

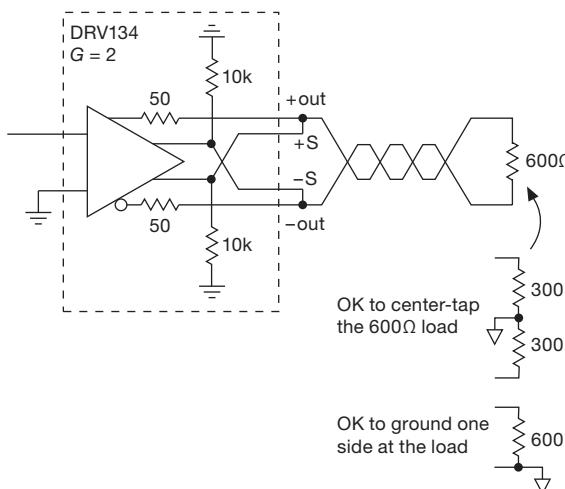
<sup>96</sup> The ADA4922-1 is a faster, 0.05% version, with a fixed unity-gain input stage.

<sup>97</sup> The ADC has an internal reference of reasonable accuracy (0.5%), but it lets you attach an external reference of better performance (e.g., the LT1790-4.096, with accuracy and drift of  $\pm 0.05\%$  and  $10 \text{ ppm}^{\circ}\text{C}$ , max). Sounds good – but it would not accomplish much here, because the gain accuracy of the system is limited by that of the LT6350 amplifier ( $\pm 0.6\%$ , max). You could use the ADA4922-1 instead.

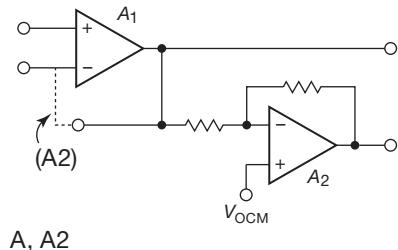
Table 5.10 Selected Differential Amplifiers

Part #	Circuit	Gain <sup>a</sup>		Bandwidth		Speed		$V_{out}$ diff <sup>b</sup>	Offset Voltage	Input bias	$Z_{in}^h$ diff	Noise			Supply	$I_s$ typ	RRO	Cost qty 25	
		Range	Set by	$\Delta G$ (%)	-3dB (MHz)	0.1dB (MHz)	Slew Rate	Settle 0.05% (ns)	$V_{pp}^b$ (mV)	typ	max	( $\mu A$ )	e <sub>n</sub> , typ @1MHz (nV/ $\sqrt{Hz}$ )	Range (V)	Comments				
<i>high voltage</i>																			
PGA280	q 1/8-128	PGA	0.15	1.5	-	-	1	40 $\mu s$	10	0.05	0.25	0.3nA	1G	22	10-40	0.8	• 6.46	INA, SPI, low $V_{out}$	
THAT1606	D3	2.0	fixed	2	10	-	-	15	-	67	4	15	-	5k	25	8-40	4.9	- 3.91	audio <sup>t</sup> , 7ppm dist
AD8270	C	0.5,1,2	fixed	0.08	15	-	-	30	700	55	0.2	0.75	0.5	20k	26 <sup>d</sup>	5-36	2.3	- 3.56	precision resistors
LME49724	D	1.0 up	Rf/Rg	-	50	-	-	18	200	52	0.2	1	0.06	$R_g^h$	2.1	5-36	10	- 3.53	audio, 0.3ppm dist
OPA1632	D	1-10	Rf/Rg	-	180	40	50	200	52	0.5	3	2	$R_g^h$	1.3	5-33	14	- 4.39	pro audio	
<i>low voltage</i>																			
LMP7312	D3	0.1-2	PGA	0.04	0.53	-	-	1.4	-	RR	-	0.1	c	160k	7.5	±5	2	y 6.43	2-ch-input, SPI
LTC1992	D	1-10	Rf/Rg	-	3.2	-	-	1.5	-	RR	0.25	2.5	2pA	30k	45 <sup>d</sup>	2.7-12	0.7 w 2.80	$R_f = 30k$ to 50k	
LTC1992-x	D3	1-10	p/n	0.3	3.2	-	-	1.5	-	RR	0.25	2.5	2pA	30k	45 <sup>d</sup>	2.7-12	0.7 w 5.23	-x for G=1,2,5,10	
AD8137	D	1.0 up	Rf/Rg	-	76	10	450	100	RR	0.7	2.6	0.5	$R_g^h$	8 <sup>d</sup>	2.7-12	2.6	• 1.98	$R_f=1k$ -10k, cheap	
THS4521	D	1-10	Rf/Rg	-	145	20	490	13 <sup>k</sup>	RR	0.2	2	0.65	$R_g^h$	4.6 <sup>d</sup>	2.5-5.5	1.1 w 2.81	$R_f=1k$ ; 4522 dual		
EL5170	E3	2.0	2.0	1	100	12	1100	20	6.0	6	25	6	300k	28 <sup>d</sup>	5-12	7.4	- 1.44	cheap, '5370 triple	
LTC6605-14	D3	1, 4, 5	1,4,5	0.3	14 <sup>e</sup>	f	f	f	RR	0.25	1	12	400/G	2.6	2.7-5.5	16 w 9.81	dual, 16 bit ready		
LT6600-x	D2	1-8	402/Rg	7	10 <sup>g</sup>	f	f	f	4.8	5	25	30	$R_g^h$	10 <sup>d</sup>	2.7-11	36	- 3.90	4th-order filters	
LTC6601	D3	1-7	1 to 7	3	6-27 <sup>i</sup>	f	f	f	RR	0.25	1.5	12	400/G	2.2	2.7-5.5	16 w 5.58	pin-select filter		
AD8138	D	1.0 up	Rf/Rg	-	320	30	1150	16	15.6	1	2.5	3.5	$R_g^h$	5	3-10	20	- 4.12	$R_f=500\Omega$	
LMH6551	D	1.0 up	Rf/Rg	-	370	50	2400	18	15.6	0.5	4	4	$R_g^h$	6 <sup>d</sup>	3.3-10	13	- 3.50	$R_f=365\Omega$	
EL5173	E3	2.0	2.0	0.5	450	60	900	10	6.0	3	30	11	150k	25 <sup>d</sup>	4.8-12	12	- 2.80	EL5373 triple	
EL5177	E	1.0 up	Rf,Rg	1.5	550	120	1100	10	6.0	1.4	25	14	150k	21	4.8-12	12	- 1.85	$G=1+R_f/R_g$ , cheap	
AD8139	D	1.0 up	Rf/Rg	-	410	45	800	45	RR	0.15	0.5	2.3	$R_g^h$	2.2	4.5-12	25	- 6.00	$R_f=200\Omega$	
AD8132	D	1.0 up	Rf/Rg	-	350	90	1200	15	14.4	1	3.5	3	$R_g^h$	8 <sup>d</sup>	2.7-11	12	- 3.07	$R_f=348\Omega$	
LT6402-20	C	10	fixed	10	300	30	400	8 <sup>k</sup>	7.0	1	6.5	c	100	1.9	4.0-5.5	30	- 3.90	opt 75MHz filter	
LTC6404-4	D	4 up	Rf/Rg	-	600	450	1200	13	RR	0.5	2	23	$R_g^h$	1.5	2.7-5.5	30 w 4.91	$R_f=402$ ; ver -1,-2 <sup>j</sup>		
THS4520	D	1-10	Rf/Rg	-	620	30	570	7	RR	0.25	2.5	6.5	$R_g^h$	2	3-6	14	• 4.07	$R_f=499\Omega$	
PGA870	s	0.3-10	PGA	4	650	100	2900	5	4.8	5 <sup>z</sup>	30 <sup>z</sup>	c	150	6	4.8-6	143	- 9.70	0.5dB atten, pins	
ADA4932-1	D	1.0 up	Rf/Rg	-	560	300	2800	9	15.0	0.5	2.2	2.5	$R_g^h$	3.6 <sup>d</sup>	3-10	10	- 4.72	$R_f=499\Omega$ ; -2=dual	
LT1993-10	C	10	fixed	12	700	50	1100	4	7	1	6.5	c	100	1.7	5	100	- 4.20	-2,-4 for G=2,4	
ADA4950-1	D	1, 2, 3	pins	1.7	750	210	2900	9	15	0.2	2.5	c	500/G	5.5 <sup>d</sup>	3-11	9.5	- 4.78	-2=dual	
LMH6553	D	1.0 up	Rf/Rg	-	900	50	2300	10	15.4	-	-	50	$R_g^h$	1.2	5-12	29	- 6.43	CFB, 274Ω; clamp	
ADA4938-1	D	1-10	Rf/Rg	-	1000	150	4700	6.5	15.2	1	4	13	$R_g^h$	2.6	4.5-11	37	- 6.06	$R_f=200\Omega$	
LMH6552	D	1.0 up	Rf/Rg	-	1500	450	3800	10	15.4	-	-	80	$R_g^h$	1.1	5-12	23	- 8.14	CFB, 357Ω	
THS4513	D	1.0 up	Rf/Rg	-	1600	700	5100	16	5.6	1	4	8	$R_g^h$	2.2	3-6	35	- 10.34	$R_f=348\Omega$	
LMH6555	D3	4.84	4.84	3	1200	180	1300	2.2 <sup>k</sup>	1.0	15	50	c	78	4 <sup>d</sup>	3.3	120	- 10.55	assumes $R_S=50\Omega$	
THS4511	D	1-10	Rf/Rg	-	1600	620 <sup>r</sup>	4900	3.3 <sup>k</sup>	5.2	1	4	8	$R_g^h$	2	3.5-5.5	39	v 9.13	$R_f=349\Omega$	
ADA4937-1	D	1-5	Rf/Rg	-	1900	200	6000	7	6	0.5	2.5	21	$R_g^h$	2.2 <sup>d</sup>	3-10	40	- 6.06	$R_f=200\Omega$ ; -2=dual	
THS4508	D	2-10	Rf/Rg	-	2000	400	6600	2 <sup>k</sup>	5.2	1	4	8	$R_g^h$	2.3	3.8-5.5	39	v 9.13	$R_f=349\Omega$	
THS4509	D	2-10	Rf/Rg	-	2000	300	6600	10	5.6	1	4	8	$R_g^h$	1.9	3-6	38	- 9.13	$R_f=349\Omega$	
LTC6416	G	0.98	1.0	2	2000	300	3400	1.8 <sup>k</sup>	4.2	0.5	5	5	12k	1.8	2.7-3.9	42	- 8.27	emitter follower	
THS770006	D	2	fixed	3	2400	350	3100	2.2	4.9	0.5	5	c	100	1.7	5	100	- 10.45	$Z_{out}=270$ at 1GHz	
AD8352	F	1-8	Rg	12	2500	190	8000	2 <sup>k</sup>	6	-	6	5	3k	2.7	3.5-5	37	- 5.65	specs for G=10dB	
ADL5561	D3	2,4,6	pins	0.5	2900	600	9800	2 <sup>k</sup>	4.3	0.25	-	c	800/G	1.7	3-3.6	40	- 5.89	good for IF-strip	
ADL5562	D3	2,4,6	pins	0.7	3300	270	9800	2 <sup>k</sup>	4.3	0.25	-	c	800/G	1.6	3-3.6	80	- 5.88	good for IF-strip	
ADA4960-1	F	1-10	Rf/Rg	6	5000	300	8700	1.4	3.5	-	36	20	10k	5.4	5	60	- 11.47	specs for G=6dB	
<i>single-ended to differential<sup>m</sup></i>																			
DRV134	B	2.0	fixed	2	1.5	-	15	2500	60	1	10	-	10k	35	9-40	5.2	- 4.21	audio <sup>t</sup> , 5ppm dist	
ADA4922-1	A2	2.0	fixed	0.05	38	-	260	580	43	0.18	0.55	3.5	11M	6	9-26	5.4	n 5.81	18-bit ADC driver	
<i>low voltage</i>																			
LT6350	A	2.0 up	opamp	0.6	33	-	48	240	RRIO	0.1	0.7	1.2	4M	4.1 <sup>d</sup>	2.7-12	4.8	• 3.42	precision	
ADA4941-1	A	2.0 up	opamp	1	30	-	22	300	RR	0.2	0.8	3	24M	5.1 <sup>d</sup>	2.7-12	2.5	w 4.45	odd $f_{3dB}$ specs	

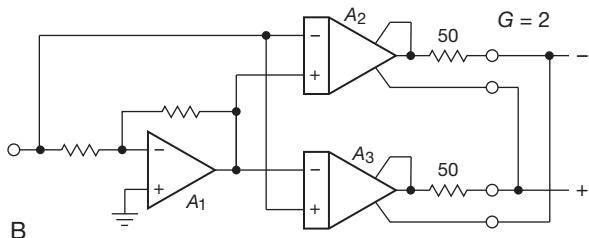
**Notes:** (a) fixed gain, programmable gain (PGA), or gain set by input resistor  $R_g$ . (b) RR output means  $V_{out}$  diff =  $2 \times V_{supply}$  max. (c) included in  $V_{os}$  spec. (d) includes feedback resistor noise. (e) matched 2nd-order antialias filters; 7, 10MHz avail. (f) set by filter. (g) 6600-x specifies 2.5 to 20MHz 4th-order filters. (h) nominally  $Z_{in} = R_g$ , where  $G = R_f/R_g$ ; but for "D" circuit types it's greater than  $R_g$ , see later section on Differential Amplifier Input Impedance. (i) filter, 6 to 27MHz strappable. (j) -4 version compensated for  $G < 4$ . (k) settle to 1%. (m) most differential-to-differential amplifiers can convert single-ended input to differential output. (n) near. (o) instrumentation amplifier figures. (p) inst-amp input, differential-amp output. (r)  $G=2$ . (s) R-2R ladder input, differential-amp output. (t) high common-mode  $Z_{out}$ , like an isolation transformer. (v) inputs to  $-V_{EE}$ . (w) RRO and inputs to  $-V_{EE}$  or within 0.2V of  $-V_{EE}$ . (y) RRO and input beyond rails, to  $\pm 15V$ . (z) RTO.



**Figure 5.93.** Balanced audio driver with a high common-mode output impedance, so that the receiver (load end) sets the common-mode voltage.  $V_{CM(out)}$  defaults to 0 V if the load is left floating.

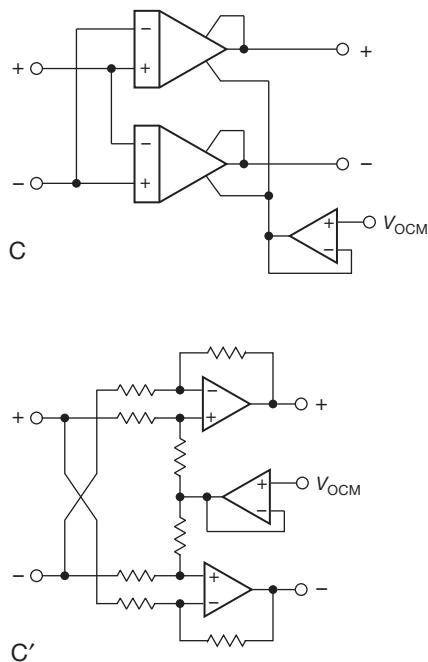


A, A'



**Figure 5.94.** Differential-output amplifier configurations A and B, as listed in Table 5.10 on page 375. The uncommitted input op-amp can be configured as a noninverting amplifier (or follower), or an inverting amplifier.

**B. Input impedance** The input impedance of amplifiers with configuration D is equal to  $R_g$ , making them unsuitable for high gains because the input impedance becomes unmanageably low: the signal source is heavily loaded, the effective  $R_g$  is increased by the source impedance  $R_S$ , and the CMRR is degraded by source-



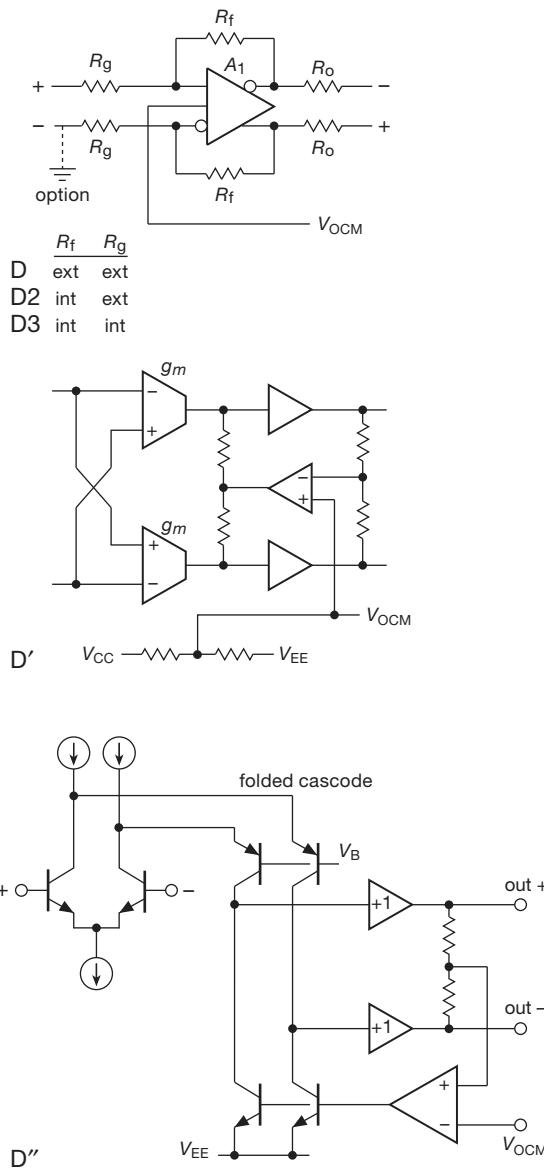
**Figure 5.95.** Fully differential amplifier configuration C, as listed in Table 5.10 on page 375. The symmetry is evident in the redrawn version, C', where the gain is  $G = 2R_f/R_g$ .

impedance imbalance. The exact value of  $R_g$  (and therefore the input impedance) will be affected by source termination and loading considerations; see §5.17.4 and the formulas in Figure 5.104.

**C. Single-ended input** Most fully differential amplifiers work fine with single-ended inputs, i.e., with the “-” input grounded. But you may wish to use  $G = 2$  or higher to achieve full peak-to-peak drive into a differential ADC.<sup>98</sup>

**D. Common-mode rejection** With differential pairs at both input and output, there are *two* measures of common-mode rejection: *differential*  $V_{out}$  versus common-mode  $V_{in}$ , which is usually quite good (e.g., 80 dB up to 1 MHz); and *common-mode*  $V_{out}$  versus common-mode  $V_{in}$ , which can be significantly poorer (e.g., 50 dB up to 1 MHz, degrading above that). But the latter is not terribly worrisome if the receiving device (e.g., an ADC) has good common-mode rejection itself. Resistor matching is important for configurations

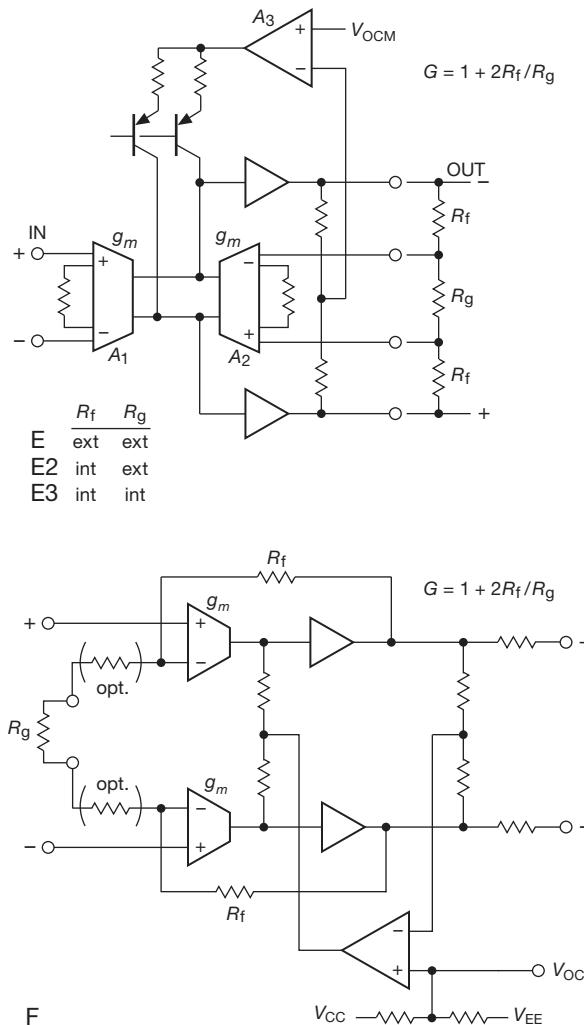
<sup>98</sup> This configuration is also available with a voltage gain less than unity. In a burst of cuteness, Analog Devices calls their AD8475 (with gains of 0.4 and 0.8) a *funnel amplifier* (get it?). With it you can reduce a 20 Vpp signal to a differential pair of 4 Vpp signals, for input to a low-voltage ADC.



**Figure 5.96.** Fully differential amplifier configuration D, as listed in Table 5.10 on page 375; the gain is  $G=R_f/R_g$ . A typical configuration for output amplifier  $A_1$  is shown in D', and TI's version for their THS45xx series is D''; the THS4508/11/21, for example, use polarity complements (*pnp* input pair, etc.), permitting operation down to  $V_{in}=V_{EE}-0.2\text{ V}$ .

with external  $R_f$  and  $R_g$  gain-setting resistors (configuration D); see §5.17.5 for more details.

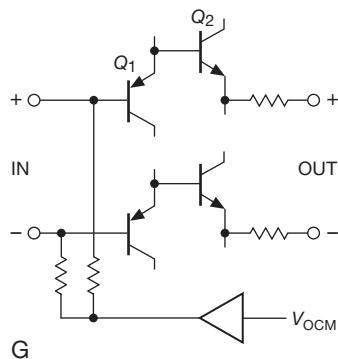
**E. Single-ended output** The datasheets of some differential amplifiers describe operation with a single-ended



**Figure 5.97.** Fully differential amplifier configurations E and F, as listed in Table 5.10 on page 375.

output.<sup>99</sup> When it's operated in single-ended output mode, however, you care about the *output* offset voltage  $\Delta V_{OCM}$  (i.e., the output error with respect to the  $V_{OCM}$  reference), which translates back to an input-referred error of  $\Delta V_{OCM}/G$ . For the LMP7312 the output offset is  $\pm 20\text{ mV}$ , far larger than the maximum input offset of  $\pm 0.1\text{ mV}$ . This is a low-gain amplifier ( $G=0.1$  to 2), so this output offset looks like a corresponding worst-case

<sup>99</sup> For example, the datasheet for the “Precision SPI-Programmable AFE with Differential/Single-Ended Input/Output” LMP7312 says that “the output can be configured in both single-ended and differential modes with the output common mode voltage set by the user.”



**Figure 5.98.** Fully differential amplifier configuration G, as listed in Table 5.10 on page 375. This type is intended for ac- or transformer-coupled inputs.

input error of  $\pm 200 \text{ mV}$  to  $\pm 10 \text{ mV}$ ! This is hardly what one would call “precision.”

**F.  $V_{\text{OCM}}$  bias pin** You set the output common-mode voltage by asserting a dc bias at this pin. Some devices buffer this input to get high  $R_{\text{in}}$ , but many parts present you with an input impedance in the tens of kilohms. Generally the operating  $V_{\text{OCM}}$  range does not extend to the negative rail. If this pin is left unconnected, most parts default to mid-supply. It’s always a good idea to bypass this pin, because the fast signals associated with these wide-bandwidth amplifiers couple to the  $V_{\text{OCM}}$  node.

**G. Input common-mode range** The input common-mode voltage range of most fully differential amplifiers does *not* extend to the negative rail, which can seriously constrain a circuit that’s running from a single positive supply. This isn’t as bad as it sounds, though: the outputs, sitting up there around a positive common-mode output voltage (set by the dc you apply to the  $V_{\text{OCM}}$  input pin), bring the input terminals up by the voltage dividers formed by  $R_f$  and  $R_g$ . This effect is largest when operating at low gains; when operating at higher gains it’s best to check that the input common-mode range is not violated. Assuming that there’s plenty of loop gain (i.e., that  $G_{\text{OL}} \gg G$ ) the (equal) voltage at the amplifier’s inverting and noninverting pins is

$$V_{(+,-)} = \frac{V_{\text{OCM}} + GV_{\text{in(CM)}}}{G+1},$$

where the differential gain  $G=R_f/R_g$ , and  $V_{\text{in(CM)}}$  is the common-mode voltage of the (differential) input signal source. If the input is single ended (with the other differential input grounded) then (substituting  $V_{\text{in}}/2$  for

$V_{\text{in(CM)}}$ ) you get

$$V_{(+,-)} = \frac{V_{\text{OCM}} + GV_{\text{in}}/2}{G+1}.$$

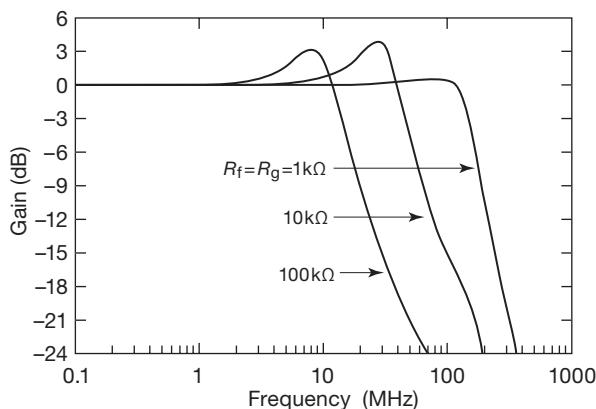
Note that with a balanced input signal source (i.e., fixed  $V_{\text{in(CM)}}$ ), the voltages at the amplifier’s (+) and (-) pins don’t vary as the input differential signal does its thing. This is in contrast with the single-ended arrangement, in which the input signal amplitude causes a variation in input common-mode voltage. For the latter be sure to check for input common-mode violations at the extremes of the input signal.

Of course, you can do an end run around the problem by choosing an amplifier whose input common-mode range includes the negative rail, for example the THS4521 illustrated in §5.17.3.

**H. Voltage feedback versus current feedback** All of the “ $R_f/R_g$ ” amplifiers in Table 5.10 (page 375) use conventional voltage-feedback amplifiers, with the exception of the LMH6552/3, which uses current feedback (CFB). Because they are VFB amplifiers, they work well with bandwidth-limiting capacitors across the feedback resistors when operated at higher gains (helpful for taming the integrated voltage noise  $v_n=e_n\sqrt{\text{GBW}}$ , which can be excessive owing to the large bandwidth of many of these amplifiers). To a decent approximation, current-feedback amplifiers have an  $f_{3\text{dB}}$  bandwidth independent of gain, whereas voltage-feedback amplifiers have a bandwidth inversely proportional to the closed-loop gain ( $f_{3\text{dB}} = \text{GBW}/G_{\text{CL}}$ ). See the discussion in Chapter 4x.

**I. Gain-setting resistors** Large  $R_f$  and  $R_g$  values can cause problems that are due to parasitic circuit-board capacitance. For example,  $R_f$  above 1k for the modest-capability 145 MHz THS4521 creates peaking (Figure 5.99). Dual and quad package options may also suffer from unavoidable peaking problems caused by lead-frame issues, so it’s generally better to choose fixed-gain types in multiple-amplifier packages. Large values of  $R_f$  and  $R_g$  also create (a) loss of speed, (b) increased input offset error from the relatively large bias currents characteristic of fast bipolar amplifiers, and (c) increased input-referred voltage noise, both from resistor Johnson noise and from the noise voltage developed across  $R_f$  by the amplifier’s input noise current.

To put some flesh on these latter bones, consider the THS4521 again, with our (somewhat extreme) straw-man  $R_f = R_g = 100\text{k}$ . From Figure 5.99 you can see the  $\times 10$  reduced bandwidth and peaking. This peaking (which occurs with VFB amplifiers but not CFB types)



**Figure 5.99.** Large gain-setting resistor values create peaking in the frequency response, as shown in this datasheet plot for the THS4521 configured for unity gain ( $R_f = R_g$ ).

can be tamed by putting a small capacitor across each feedback resistor, but you'll lose a bit more bandwidth in the process.<sup>100</sup> Turning this vice into a virtue, we note that you may wish to add feedback capacitors anyway, to reduce the bandwidth intentionally.

As for offset voltage, this part has an input bias current of 650 nA (typ), which would create a drop of 65 mV across 100k. But the bias currents are reasonably well matched, with an offset current spec of  $\Delta I_B = \pm 50$  nA (typ), thus creating an input offset of 5 mV. That's much better, but it does seriously degrade the amplifier's typical  $V_{OS} = \pm 0.2$  mV ( $\pm 2$  mV max). You need to keep  $R_f$  and  $R_g$  less than 10k to preserve the amplifier's accuracy.

Finally, noise. There are two contributions, the resistors' Johnson noise ( $e_n = \sqrt{4kTR} = 0.13\sqrt{R}$  nV/ $\sqrt{\text{Hz}}$ ), and the noise voltage developed by the amplifier's current noise ( $e_n = i_n R_f$ ). For  $R_f = 100k$  the Johnson noise voltage is 40 nV/ $\sqrt{\text{Hz}}$ , and the noise produced by the amplifier's  $i_n = 0.6$  pA/ $\sqrt{\text{Hz}}$  is 65 nV/ $\sqrt{\text{Hz}}$ . These disastrously degrade the amplifier's typical 4.6 nV/ $\sqrt{\text{Hz}}$  (taking the usual root-sum-of-squares, the total added noise voltage is 76 nV/ $\sqrt{\text{Hz}}$ ). The table below summarizes these figures, and also those corresponding to  $R_f R_g = 10k$  and  $R_f = R_g = 1k$ .<sup>101</sup>

<sup>100</sup> You can look at this another way: the manufacturer's recommended gain-setting resistor values are chosen such that a small amount of peaking is exploited to extend the amplifier's natural bandwidth.

<sup>101</sup> Many parts (the D and E configurations) let you add feedback capacitors to reduce the bandwidth. With some parts this may introduce instability at low gains, but with others it may improve the stability, especially when larger resistor values are used.

Bottom line: compared with the nominal 1k value, the use of 100k gain-setting resistors reduces your bandwidth  $\times 10$ , increases the typical offset voltage  $\times 25$ , and increases the typical input-referred noise voltage  $\times 16$ . You wouldn't want to do this. But you could reasonably use something like 2.49k, 4.99k, or perhaps 10k, buying increased input impedance at the expense of modest degradation of bandwidth, noise, and accuracy.

$R_f, R_g$	-3 dB Bandwidth (MHz)	Offset voltage* (mV, typ)	Input-referred noise		
			$\sqrt{4kTR}$ (nV/ $\sqrt{\text{Hz}}$ )	$i_n R_f$ (nV/ $\sqrt{\text{Hz}}$ )	Total* (nV/ $\sqrt{\text{Hz}}$ )
1k	150	$\pm 0.2$	4	0.7	4.6
10k	45	$\pm 0.5$	13	6.5	15
100k	15	$\pm 5$	40	65	76

\* Includes  $V_{OS}$  and  $e_n$  of the amplifier.

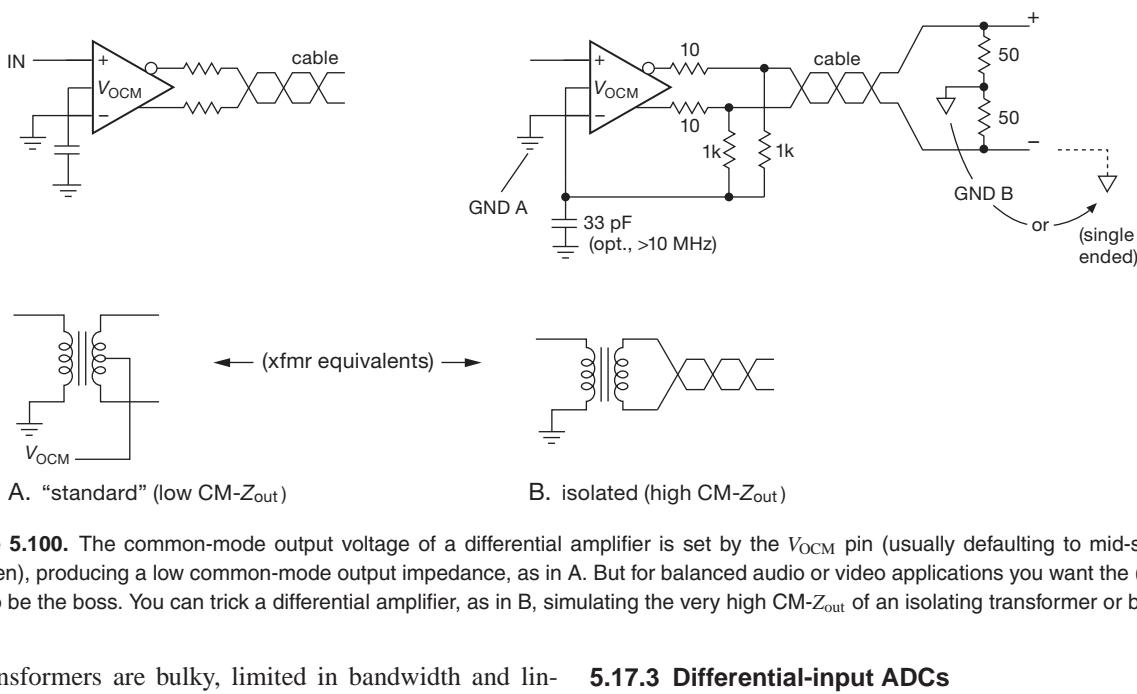
**J. Common-mode output impedance** The voltage asserted onto the  $V_{OCM}$  pin sets the common-mode output voltage. Put another way, differential amplifiers have a low common-mode output impedance. That's usually what you want; after all, that's the reason there is a  $V_{OCM}$  pin. But this can create difficulties if the output is sent to a remote load that needs to establish its preferred common-mode level. That is the case in balanced audio (or video), sent substantial distances over balanced twisted-pair cable.

Take a look at Figure 5.100B. By driving the  $V_{OCM}$  pin with the average output voltage, you create an amplifier that cooperates by letting the load lead the dance. In fact, the load can even unbalance the signal intentionally (by grounding one side), in which case the other output swings symmetrically around ground with the full desired differential output voltage.<sup>102</sup> There are a few differential amplifiers that are designed specifically for this kind of application, with an internal configuration that creates a high common-mode output impedance. We saw one example (the DRV134, similar to the SSM2142) in Figure 5.93. Another excellent one is the THAT1606, from the curiously named THAT Corporation.<sup>103</sup>

The traditional solution has been to use an isolating transformer, which also can do the job of converting between single-ended and balanced signals (this is known as a "balun," for *balanced-unbalanced*) as shown. But

<sup>102</sup> Omit the small bypass capacitor shown, if this mode of operation is anticipated.

<sup>103</sup> A member of the hard-to-Google corporate name club, which includes AND Displays, and ON Semiconductor. (Try it: you'll get more than ten billion Google hits for "AND" or "ON.")



**Figure 5.100.** The common-mode output voltage of a differential amplifier is set by the  $V_{OCM}$  pin (usually defaulting to mid-supply if undriven), producing a low common-mode output impedance, as in A. But for balanced audio or video applications you want the (distant) *load* to be the boss. You can trick a differential amplifier, as in B, simulating the very high CM- $Z_{out}$  of an isolating transformer or balun.

transformers are bulky, limited in bandwidth and linearity, and not inexpensive. High CM- $Z_{out}$  differential amplifiers can be an attractive alternative.

### 5.17.2 Differential amplifier application example: wideband analog link

We conclude the discussion of differential amplifiers with several application examples: a wideband analog link over twisted-pair cable, and a riff on driving differential-input ADCs.

In §5.14.2F we illustrated the use of a difference amplifier as the receiving end of an analog link over differential twisted-pair cable. In that circuit  $R_4C_1$  creates a rising response at high frequencies ("equalization"), to compensate for the increasing cable attenuation. To complete the link, of course, we need a differential driver.

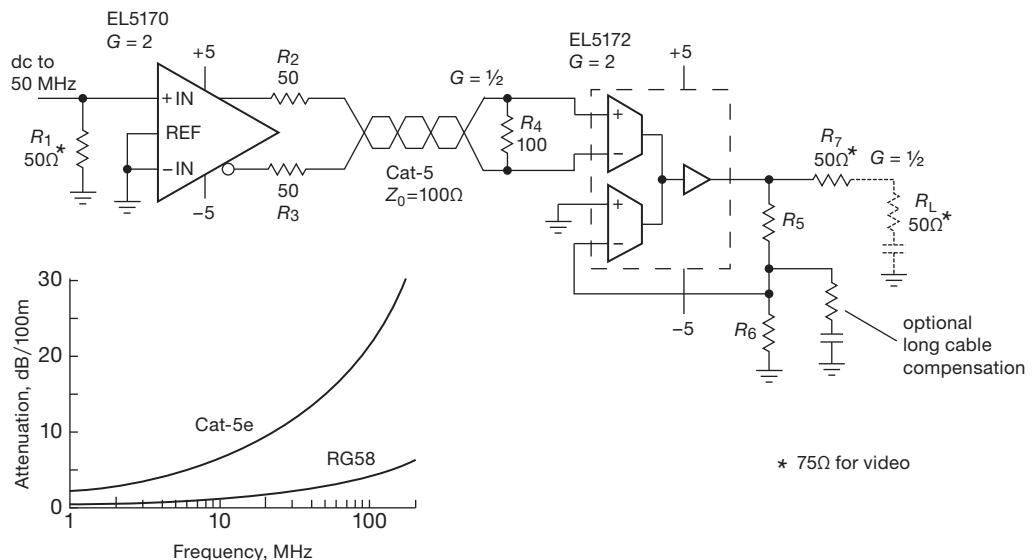
Figure 5.101 shows the whole enchilada, here implemented with Intersil's EL5170/72 differential line driver/receiver pair. They make triple units (EL5370/72) also, with application to color video. Bandwidths to tens of megahertz are easily achievable over tens of meters of Cat-5 cable, with modest equalization at the receiver end. Coaxial cable is considerably better, and two  $50\Omega$  coax lines can replace the  $100\Omega$  differential pair. As always, the balanced signal combines with excellent receiver CMRR (here 95 dB typ) to provide high rejection of interference.

### 5.17.3 Differential-input ADCs

Many analog-to-digital converters require differential signal inputs. This is almost universally true of high-speed converters (e.g., pipelined flash converters), as well as the varieties known as "charge-redistribution SAR" and "delta-sigma" (ADCs, in all their glory, are the major subject of Chapter 13). And in many cases the input is hardly benign – the internal switched capacitors introduce charge transients at the input terminals, mandating some external shunt capacitance. An additional annoyance is the requirement that the driver must be able to swing the inputs over the full conversion range (which may include ground), but without driving them beyond the ADC's power rails (with the risk of damage from input clamp conduction and possible SCR latchup).

#### A. First iteration: single-supply ADC driver with $V_{OCM}$ offset

Figure 5.102 shows two iterations of an input stage for fast single-supply ADCs with differential inputs. Our first design was based around the AD9225 12-bit 25 Msps pipelined flash converter, which runs from a single +5 V analog supply, and has a separate digital supply pin for interfacing to microcontrollers running from +3 V to +5 V. Its input span is programmable, either 0–2 V or 0–4 V, and it provides a mid-span dc output (+1 V or +2 V) that can be used to set the differential amplifier's common-mode output (via the  $V_{OCM}$  pin).



**Figure 5.101.** Wideband analog link over Cat-5 network cable. The EL5370/72 package three similarly performing drivers–receivers in a single IC, convenient for sending analog video (RGB, S-Video, or YPbPr component video) over a single cable (which has four twisted pairs). See also Figure 5.71.

By running the differential amplifier from the same +5 V and ground, we are assured that the ADC’s inputs cannot be driven beyond the rails. We chose the AD8139 differential amplifier for its low noise ( $2.2 \text{ nV}/\sqrt{\text{Hz}}$ ), ample bandwidth ( $\sim 15 \text{ MHz}$  at  $G = 20$ ), and ability to swing its output rail-to-rail (to drive the ADC over its full-input span). We used the recommended pair of series resistors to suppress amplifier ringing from charge transients at the ADC input; and we added a shunt capacitor to both reduce these transients and also provide a second stage of anti-alias filtering.<sup>104</sup>

That’s all good news. The bad news is that this amplifier, in common with most differential amplifiers, does not include ground in its input common-mode operating range: you’ve got to stay a volt away from either rail. So you cannot simply tie one input to ground and drive the other with a small signal around ground.<sup>105</sup> The amplifier *does* let you run from split supplies (e.g.,  $\pm 5 \text{ V}$ ), which solves the input signal-level problem; but then you have to worry about

power-supply sequencing and the risk of driving negative current into the ADC’s clamp diodes.

### B. Second iteration: single-supply ADC driver with $V_{in(CM)}$ to ground

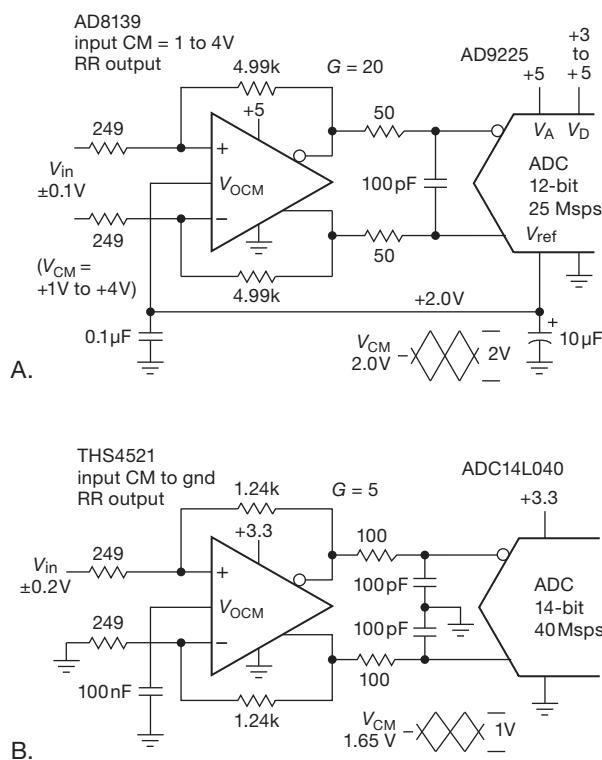
What to do? Find a single-supply amplifier that operates with inputs to the negative rail! That’s what we did in the second circuit, whose THS4521’s common-mode input range includes ground (“NRI” – negative rail input), and in fact guarantees proper operation with inputs to  $-0.1 \text{ V}$ .<sup>106</sup> It also has the needed rail-to-rail output, but it is somewhat noisier and slower than the AD8139 ( $4.6 \text{ nV}/\sqrt{\text{Hz}}$ , and a gain of only  $\times 5$ , to maintain 18 MHz bandwidth).

We teamed it with the ADC14L040, a more accurate and faster ADC (14-bit, 40 Msps) that runs on a single +3.3 V supply and uses less power (235 mW versus 335 mW). The ADC’s span is  $\pm 0.5 \text{ V}$ , centered on an allowable midspan voltage of +0.5 V to +2.0 V. We could have used the ADC’s reference-derived +1.5 V output to drive the amplifier’s  $V_{OCM}$  pin, as before; but when that pin is not driven the amplifier defaults to mid-supply

<sup>104</sup> Repeating some important advice: when designing with high-speed ICs, it’s particularly important to pay attention to special instructions in the datasheet; the AD9255, for example, devotes nearly a full page to a discussion of input  $R$ ’s and  $C$ ’s.

<sup>105</sup> Except when operating at low gain: here that would require  $G \leq 1$ , so that the AD8139’s input terminals are brought up to 1 V or more by the  $R_f R_g$  divider. See the discussion in §5.17.1.

<sup>106</sup> Other differential amplifiers whose feedback inputs can operate at or near ground are indicated with **w** or **v** in Table 5.10 (page 375), and include the LTC1992, LTC6605, LTC6601, LTC6404, THS4508, and THS4511. These parts span the bandwidth range from 3 MHz to 2000 MHz.



**Figure 5.102.** The care and feeding of single-supply differential-input ADCs. A. The AD9225 provides a midscale  $V_{ref}$  output to set the amplifier's common-mode output; but the AD8139's inputs cannot go closer than 1 V from either rail. B. The THS4521 is unusual in allowing input voltages to the negative rail (here ground).

(+1.65 V), which is fine. As before we added the recommended decoupling filter.

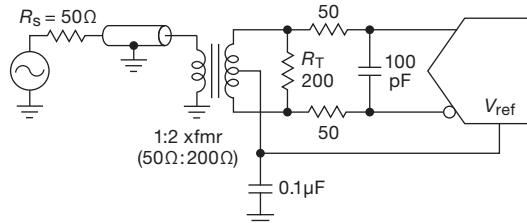
Given the ADC's higher resolution, it's worth checking to see how the noise voltage contributed by the amplifier and resistors compares with the converter's step size. Taking into account the input gain, the step size is  $400 \text{ mV}/2^{14}$ , or  $25 \mu\text{V}$ . The amplifier's noise density ( $4.6 \text{ nV}/\sqrt{\text{Hz}}$ ) combined with the (uncorrelated) resistor noise ( $2.7 \text{ nV}/\sqrt{\text{Hz}}$ ) is about  $5.3 \text{ nV}/\sqrt{\text{Hz}}$ , or about  $18 \mu\text{V}_{\text{rms}}$  in the  $\sim 12 \text{ MHz}$  effective bandwidth of the amplifier and  $RC$  filter. In other words, the noise voltage is comparable to the converter's LSB step size. This is OK, though it would be nice to have it somewhat lower.<sup>107</sup> Perhaps a way to think about it is that the circuit's virtues of speed and resolution have made

the relatively small noise contribution look bad. You can always throw away bandwidth (if you don't need it), or just look at the top 12 bits, if that makes you feel better.

### C. Third Iteration: transformer coupling

If you don't need dc coupling, an easy way to drive differential-input converters is with a wideband transformer. They're widely used in radiofrequency applications, and you can get them in small surface-mount packages. Figure 5.103 shows how to do it. Use the ADC's mid-span reference output (suitably bypassed) to set the common-mode voltage, and use a resistive termination that matches the transformed impedance of the driving source. Here we used a 1:2 turns ratio transformer, which transforms impedances by that ratio squared, thus  $50 \Omega:200 \Omega$ . This nicely fineses the problems of amplifier input- and output-voltage ranges, noise, and so on. But note that there's no intrinsic protection against ADC overdrive.

T4-1: 0.2–350MHz  
T4-6T: 0.02–250MHz  
LA0511: 0.001–15MHz



**Figure 5.103.** Wideband transformers can drive differential-input ADCs. They have excellent CMRR, and are available with frequency spans of 10,000:1.

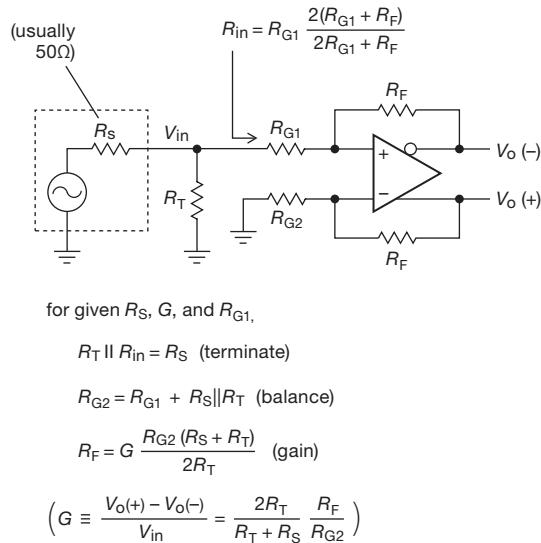
### 5.17.4 Impedance matching

Differential amplifiers are commonly used for widebandwidth applications, in which the (single-ended) input must be properly terminated to match the signal's source impedance (usually  $50 \Omega$ ). This is particularly important when the signal arrives through a length of transmission line, in order to prevent reflections (see Appendix H). This is not difficult, as long as you keep your wits about you.<sup>108</sup>

Figure 5.104 shows the situation when differential amplifiers of the D type (Figure 5.96) are used. The extra resistor  $R_T$  is chosen so that the signal source sees an input impedance equal to  $R_S$  (that is,  $R_T \parallel R_{in} = R_S$ ). Note

<sup>107</sup> But sometimes a little bit of noise can be a *good* thing, as it can improve ADC linearity and dynamic range by way of "dithering." See, for example, *The Art of Digital Audio* by John Watkinson (3rd ed., 2001); or Vanderkooy and Lipshitz "Dither in digital audio," *J. Audio Eng. Soc.*, **35**, (12), 966–975, (1987).

<sup>108</sup> And perhaps read the helpful Analog Devices MT-076 "Differential Driver Analysis."



**Figure 5.104.** Terminated single-ended to differential amplifier: design equations.

especially that the amplifier's noninverting input is not a virtual ground, so  $R_{in}$  is somewhat larger than  $R_{G1}$  alone, according to the equation in the figure. The differential amplifier has the usual equal-value feedback resistors  $R_F$ , but the gain-setting resistors  $R_G$  are unequal, to take account of the finite impedance at the drive point (marked  $V_{in}$ ). That is,  $R_{G2}$  must be larger by the parallel resistance  $R_S \parallel R_T$ . Finally, the feedback resistors  $R_F$  must be adjusted upward to bring the gain back to the desired value.

Note that the gain is defined in terms of  $V_{in}$ , i.e., with respect to the *loaded* input signal amplitude (*not* the open-circuit source amplitude). This makes good sense, because signal amplitudes (from signal generators, etc.) are normally specified as their properly terminated amplitudes.

As an example, for a  $50\Omega$  source,  $G = 2$ , and  $R_{G1} = 200\Omega$ , you would find (choosing nearest 1% standard resistor values)  $R_T = 60.4\Omega$ ,  $R_{G2} = 226\Omega$ , and  $R_F = 412\Omega$ .

Unlike the situation at high frequencies, it is not necessary to terminate a signal source when operating at low frequencies (e.g., audio). In that case  $R_T$  is omitted and  $R_{G2}$  is simply  $R_{G1} + R_S$ . The gain, defined now in terms of the *open-circuit* source signal amplitude, is just  $G = R_F / (R_{G2})$ .

### 5.17.5 Differential amplifier selection criteria

All differential amplifiers are not created equal. There are plenty of subtleties lurking, associated with tradeoffs of bandwidth, accuracy, output drive capability, supply voltage, and the like. Here's a summary collection of such con-

siderations, tied closely to the differential amplifiers listed in Table 5.10 on page 375.

#### A. Supply voltage, RR output capability

The first group in Table 5.10 lists the high-voltage differential amplifiers, with  $\pm 12$  V to  $\pm 15$  V supplies (though some can work down to  $\pm 5$  V). These are usually used with bipolar (split) supplies, but most have the  $V_{OCM}$  common-mode output capability to drive single-supply ADCs. This common-mode capability distinguishes the parts on this table from other types. Most parts have internal rail-splitting divider resistors (which require a bypass capacitor) to establish the common-mode output voltage, but this can be overridden by a dc mid-span output provided from the ADC (see Figure 13.28 in §13.6.2). Be sure to check both datasheets – sometimes an op-amp will be required (in the manner of Figure 5.86).

Keep in mind that " $V_{out}$  diff'1 max (Vpp)" means  $(V_{a+} - V_{b-}) + (V_{b+} - V_{a-})$ ; that is, twice the peak-to-peak output swing of any one output.

These parts have high differential-output capability,  $>50$  Vpp (each output going  $\pm 12.5$  V), and higher still for  $\pm 18$  V supplies, thus well suited for line-driving applications. The differential THAT1606, OP1632, and LME49724, and the single-ended-input DRV134 are all intended for pro-audio (see §13.9.11D). As described before, the differential types can also be used with single-ended inputs. For lowest distortion, all four of these parts should be driven with a low-impedance source, such as an op-amp output.

Next in Table 5.10 (page 375) are the low supply-voltage amplifiers. Most high-frequency, low supply-voltage differential amplifiers are limited to a maximum of  $\pm 5$  V supplies, or even less. Many cannot be used with total supply voltages greater than 5 V, or even 3.3 V in some cases. Some can run from a single supply as low as +2.7 V or +3.3 V, others need at least +5 V.

Many of the low- to mid-frequency low-voltage parts have rail-to-rail (RRO) outputs, convenient for single-supply ADCs, which don't allow signals beyond their supply rails: simply power the amplifier from the same supply rail as the ADC. But note especially that high-frequency RRO types may suffer degraded high-frequency capability when used near their supply rails. For example the LTC6404, with a featured 600 MHz bandwidth, reveals even at 10 MHz distortion that climbs dramatically when the output approaches within 400 mV of the rails.

An alternative to the use of low-voltage RR outputs to protect ADC inputs is to use an amplifier with output-voltage clamping. That's a nice feature of the LMH6553.

This part is also a CFB amplifier, good for wide bandwidth at high gains, but bad for noise (see the next subsection).

### B. Common-mode input range, and the negative rail

Most of the parts have summing junctions that must be operated at least a volt or more above the negative supply rail (the low-power THS4521 is an exception). However, as explained in §§5.17.1G and 5.17.3), that does not necessarily prevent the input *signals* from going down to ground, most particularly when the amplifier is operated in a fully differential configuration at low gains ( $G = 1$  or  $G = 2$ , for example).

Nine of the parts (marked **w** or **v** in Table 5.10's "RRO" column – see page 375) allow use of their summing junctions down to  $-V_{EE}$ . The datasheets will declare something like "Common-Mode Input Range Includes the Negative Rail" or "NRI" on the front page. In most cases performance is not degraded under this condition, in contrast to the loss of performance when RRO-capable amplifiers are run to the extremes of output swing.<sup>109</sup> This is especially useful when fully differential amplifiers are used as single-ended to differential converters, with the  $(-)$  input grounded, as in Figure 5.102B. But be careful – if either summing-junction input is taken further than the specified  $-0.2$  V below  $-V_{EE}$ , polarity reversal at the output may occur, similar to the situation with the legacy (but still wildly popular) LM324/358 single-supply op-amps.<sup>110</sup>

### C. Low $Z_{in}$

Most of these amplifiers present rather low input impedances to their signal sources, especially when configured for high gains, because the specified  $R_f$  gain-setting resistor values are low, and  $Z_{in}$  is roughly  $R_f/G$  (exceptions: LTC6416, and the EL5170 family). Most parts with higher input impedances are noisier, mostly because of resistor Johnson noise (exception: the AD8352, which uses the F rather than the D configuration).

Signal impedance matching is often a concern, especially at high frequencies, say 30–100 MHz and above, even with short PCB traces. An amplifier's low  $Z_{in}$  complicates the problem of matching the impedance of the signal source, and it also affects the amplifier's gain. See §5.17.4 for helpful formulas.

<sup>109</sup> Some parts (for example the THS4008 and THS4511) even specify " $V_{S-} = 0$ " and "input referenced to ground" as the operating condition for their specifications.

<sup>110</sup> For the latter you can substitute the improved LT1013/1014, which eschew that nasty habit and provide better performance overall; but no such solution is available for differential amplifiers.

### D. Offset voltage, CMRR

Many of the amplifiers in Table 5.10 (see page 375) have poor offset voltage and other dc specifications. Most of these are fixed-gain parts with internal resistors (e.g., with the D3 configuration). They appear to be suffering from offsets arising from the high  $V_{OCM}$  common-mode output voltage that's natural with differential amplifiers, combined with modest internal resistor mismatches; for example, a 1% resistor mismatch operating on a  $V_{OCM}$  of 1.5 V would produce an effective input offset of 15 mV. By contrast, most "bare" parts (e.g., the D configuration) in the table have attractive low offset-voltage specs. But they will surely develop high offsets if you were to use 1% gain-setting resistors to complete the amplifier. Note also that dc accuracy is degraded by mismatched external resistance when driving D-configuration differential amplifiers.

If we had room for a CMRR column in the table, we'd see a similar bare- versus fixed-gain dichotomy, and for the same reason. Taking two examples from their respective datasheets, the bare ADA4932 has a 100 dB typical CMRR, compared with 64 dB for the similar fixed-gain ADA4950. Likewise for the bare LTC1992 (90 dB) and the fixed-gain LTC1992-10 (60 dB).

In many fully differential applications CMRR isn't important. But if it matters in your design, use 0.1% resistors, or matched resistor arrays. And be sure to take special care with circuit-board wiring capacitances, which really matter at high frequencies: for example, to achieve  $-80$  dB matching at 1 MHz with  $500\ \Omega$  gain-setting resistors, you've got to match capacitances to a difficult 0.03 pF! And no good deed goes unpunished – the CMRR will degrade by 20 dB for each decade increase in frequency.

A better solution would be to use an E- or F-configuration part. Intersil's EL5170-series parts have good CMRR, e.g., 80 dB at 1 MHz (but poor 25 mV offset), and Analog Device's AD8352 boasts 60 dB at 100 MHz and 6 mV offset. Both are much less affected by unbalanced input resistances than D-configuration types.

### E. Fixed gain, external resistor-settable gain

One good reason to select a fixed-gain amplifier is that some of them have better gain accuracy than can be easily and inexpensively achieved with discrete resistors – for example, attractive worst-case gain errors of  $\pm 0.04\%$  for the NSC LMP3712,  $\pm 0.15\%$  for the TI (Burr-Brown) PGA280 (both programmable-gain parts), and  $\pm 0.08\%$  for the Analog Devices pin-programmed AD8270.

The simplicity of fixed-gain types may seem appealing, but some external-resistor types have their attractive aspects, e.g., TI's THS4520 and Analog's ADA4932

draw much lower supply currents than their competitors. The THS4520 can be used to make  $G=10$  amplifiers with 120 MHz bandwidth.

Fixed-gain types are easier to use at high frequencies, because they avoid painful wiring and pin-capacitance problems. But most have poor absolute gain accuracy,  $\Delta G$  in Table 5.10 (exception: the ADL5561). Most don't allow bandwidth-limiting filter capacitors to be added, and most restrict you to low gain values (exceptions: the PGA870 and LT1993-10, but note their high power consumption).

#### F. VFB, CFB, $f_{-3\text{dB}}$ , GBW, and filters

Few of the fixed-gain amplifiers offer special filtering capability. Three LTC offerings are an exception, especially the LT6600-x, with five fixed 4th-order filter frequencies (alliteration!) available from 2.5 MHz to 20 MHz.

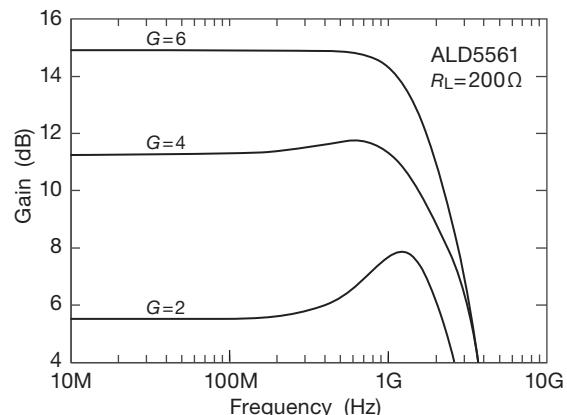
All of the adjustable-gain types employ voltage-feedback-mode op-amps, with two exceptions (the LMH6552 and LMH6553). For gains  $G \geq 4$  they follow the GBW rule, namely  $f_{-3\text{dB}} = \text{GBW}/G$ . But be aware that the "Bandwidth" value in the table is generally considerably higher ( $1.5\times$  or more) than the part's GBW, because it is determined at unity gain, where the amplifier benefits from response peaking that extends its  $-3\text{ dB}$  rolloff frequency. (The manufacturer may want to show off its best-looking values, but the peaking for  $G = 1$  may be so severe with some parts that you may not want to use them that way.) You may have to study datasheet response plots, etc., to determine the actual true (and very useful) GBW value. Because these are VFB op-amp types, they are stable with bandwidth-limiting filter capacitors added across the feedback resistors. You can increase the  $R_f$  value (which increases the input impedance) and add a small parallel  $C_f$  capacitor to bring peaking under control or a larger one to provide a bandwidth filter for your signal.<sup>111</sup>

#### G. Response peaking, GBW, and 0.1 dB bandwidth

Gain peaking is a primary killer of good "bandwidth-to-0.1 dB" ratings. The 0.1 dB bandwidth numbers may be

much improved for higher gains, where low-gain peaking is eliminated. Take, for example, the ALD5561, which boasts a  $-3\text{ dB}$  bandwidth of 2900 MHz at its minimum gain ( $G = 2$ ), but its  $-0.1\text{ dB}$  bandwidth is a disappointing 200 MHz (i.e., just 7% of its  $-3\text{ dB}$  bandwidth). However, at its maximum gain ( $G = 6$ ), where the  $-3\text{ dB}$  is reduced somewhat (to 1800 MHz), its  $-0.1\text{ dB}$  bandwidth improves to 600 MHz (i.e., 33% of its  $-3\text{ dB}$  bandwidth). This behavior is displayed nicely in the datasheet's graphs (see Figure 5.105). It is possible that the settling time is also improved (from a lack of ringing), although this isn't specified.

Note that some of the parts in the table are available in dual configurations (noted in the Comments column in Table 5.10), which can be helpful in providing matched time-delay responses, important for many applications.



**Figure 5.105.** Amplifier peaking at low gain settings produces an extended " $-3\text{ dB}$  bandwidth," at the expense of flatness of response.

#### H. Slew rate, settling time, large-signal bandwidth

Just as with some high-speed op-amps, the datasheets indicate a much wider bandwidth for small ( $\sim 100\text{ mV}$ ) signals than for large ( $\sim 2\text{ V}$ ) signals. This is a slew-rate issue: amplifier output-swing capabilities are reduced as their slew-rate limits are approached. For example, Analog Devices' low-power ADA4932 has a  $2800\text{ V}/\mu\text{s}$  specified slew rate, which implies that 1 V amplitude sinewave outputs are possible to  $f = S/2\pi A = 445\text{ MHz}$ . Indeed the part's datasheet shows a  $-3\text{ dB}$  response to 560 MHz (or even 1 GHz with smaller  $R_f$ ) for 100 mV output, but only to 360 MHz for 2 Vpp. Higher slew-rate parts are available,

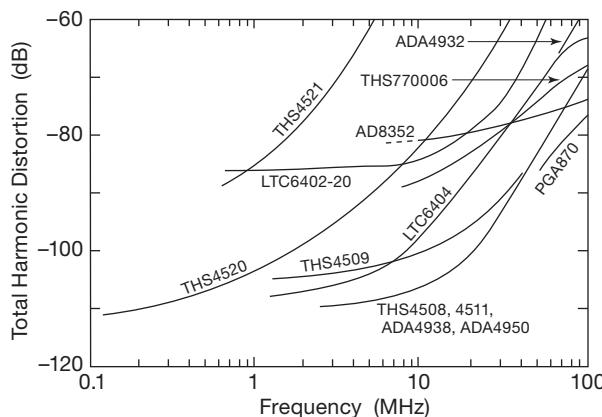
<sup>111</sup> Taking the example of the LMH6553 and LMH6552 CFB amplifiers, these are specified with  $R_f = 274\Omega$  and  $357\Omega$ , at which the respective bandwidths are 900 and 1500 MHz, and the slew rates are 2300 and 3800  $\text{V}/\mu\text{s}$ . These specs are for  $G=1$ , but with CFB op-amps you can dramatically increase their gain without losing too much bandwidth. For example the 1500 MHz LMH6552 claims to have 800 MHz of bandwidth still at  $G=4$ . For higher gain you may prefer not to reduce  $R_i$  much, but rather to increase  $R_f$ . With CFB amplifiers a primary effect of increasing  $R_f$  is to reduce the slew rate proportionally; but, hey, you had plenty to begin with! Increasing  $R_f$  with CFB does cause an increase in noise.

even to  $10\text{ kV}/\mu\text{s}$  (the ALD5561), implying a  $2\text{ Vpp}$  capability to  $1.5\text{ GHz}$ .<sup>112</sup>

### I. Distortion

Two of the high-voltage parts (the OP1632 and LME49724), often used for professional music applications, have their distortion performance plotted in Figure 5.43. We would hope that a fully differential circuit, with differential input signals, might have lower distortion than competing single-ended circuits, at least for the symmetrical 2nd-harmonic. And indeed the differential LME49724 does very well, in the  $-140\text{ dB}$  territory. However, the single-ended LME49990 and OPA134 op-amps do better on the graph.

Figure 5.106 plots distortion (from manufacturers' datasheets) for a few of the differential amplifiers in Table 5.10 (page 375) for frequencies to  $100\text{ MHz}$ . As we cautioned earlier, the conditions for distortion data are not standardized, complicating direct comparisons. Consequently, many of these parts provide multiple plots, taken with different combinations of gain, load resistance, signal amplitude, and supply voltage, and in which 2nd- and 3rd-harmonic distortion curves are shown separately.



**Figure 5.106.** Total harmonic distortion (THD) versus frequency, taken from the respective datasheets, for a selection of the differential amplifiers in Table 5.10 on page 375.

Be careful when evaluating distortion (a parameter that we've not listed in the table). For example, the ADA4932 has  $560\text{ MHz}$  (or  $360\text{ MHz}$ ) bandwidths, as discussed earlier, but the luster of its front-page claim to being a low-distortion amplifier ( $-90\text{ dB}$  at  $20\text{ MHz}$ ) is tarnished some-

what when you discover that it deteriorates by a factor of  $10\times$  by  $50\text{ MHz}$ , well below its  $360\text{ MHz}$  bandwidth.

As we saw with the op-amp distortion plots in Figures 5.43 and 5.44, there's a strong correlation between speed (high GBW, fast slew rate) and low distortion at high frequencies. This is especially clear in Figure 5.106 above  $1\text{ MHz}$ . For example, the  $145\text{ MHz}$  THS4521 that we've often mentioned performs poorly even in the under- $5\text{ MHz}$ -region, compared with its TI relative, the  $1.6\text{ GHz}$  THS4511 (note that both have NRI front ends, i.e., operating common-mode input range to the negative rail).<sup>113</sup> Four of the best parts in this class maintain better than  $-100\text{ dB}$  distortion out to  $20\text{ MHz}$ , compared with just  $7\text{ MHz}$  for the best-in-class AD8045 op-amp in Figure 5.44. In other words, at high frequencies (say above  $10\text{ MHz}$ ) fully differential amplifiers excel in low distortion compared with single-ended op-amps.

As we will see in Chapter 13, 16-bit ADCs are available with sampling rates to  $250\text{ Msps}$  (e.g., the AD9467, see Table 13.4), justifying a need for better than  $0.01\%$  linearity ( $-80\text{ dB}$  distortion) at frequencies approaching  $100\text{ MHz}$ . Happily, manufacturers of differential amplifiers are rising to this challenge.<sup>114</sup>

### J. Noise, high $1/f$ noise corners

We conclude with a few comments about noise. We've tabulated amplifier *voltage* noise, but not *current* noise. But we do list the input bias current  $I_B$ , which is roughly predictive of current noise, which must equal or exceed the shot-noise contribution of  $i_n = \sqrt{2qI_B}$ . Note that CFB amplifiers, with their exceptionally high input currents, have much greater input current noise, generally as much as  $10\times$  greater than VFB amplifiers.

Looking at Table 5.10 (page 375), we see many amplifiers with noise densities in the range of  $25\text{--}45\text{ nV}/\sqrt{\text{Hz}}$ . Assuming no increase above say  $10\text{ MHz}$  (check the datasheet plots) this corresponds to a wideband input-voltage noise of  $V_n = e_n \sqrt{BW}$ , which evaluates to  $175\text{--}700\text{ }\mu\text{VRms}$  for bandwidths of  $50\text{--}250\text{ MHz}$ . This is quite a bit larger than the  $30\text{ }\mu\text{V}$  LSB step size of a 16-bit ADC with a  $2\text{ Vpp}$  full-scale differential input. While some

<sup>113</sup> In fairness, the THS4521 gets by on a miserly  $1.1\text{ mA}$  of supply current and has RR outputs, whereas the THS4511 takes  $39\text{ mA}$  and lacks RROs.

<sup>114</sup> But are they rising to the challenge posed by NSC's ADC12D1800, a 12-bit 3600 Msps converter? Or the even greater challenge posed by the even-speedier ADCs that surely will be available by the time this book's ink has dried? Such ADCs will likely have to be driven with transformers.

<sup>112</sup> The AD8351 (not in our table, but similar to the AD8352), offers  $3\text{ GHz}$  bandwidth with  $13\text{ V/ns}$  slew rate and  $2\text{ Vpp}$  capability to nearly  $2\text{ GHz}$ .

dither is a good thing, it's clear that even at  $G = 1$  these amplifiers are too noisy for some applications.

There are many other amplifiers in the table with  $e_n$  specs down in the  $1.1\text{--}5 \text{ nV}/\sqrt{\text{Hz}}$  range. But these are the bare amplifier summing-junction noise specs, without the necessary feedback resistors taken into account. Amplifier gains of 5 or 10 are generally assumed, not the least to overcome the amplifier's own output-stage noise. Many of the amplifiers specify  $R_f$  values of  $350\text{--}500 \Omega$ . For  $G = 1$  the input resistor  $R_g$  would have the same value, and its Johnson noise of  $2.4\text{--}2.8 \text{ nV}/\sqrt{\text{Hz}}$  would dominate the intrinsic noise of the quieter amplifiers. However, for  $G = 10$  the  $35\text{--}50 \Omega$  resistor noise is under  $1 \text{ nV}/\sqrt{\text{Hz}}$ , which would not badly degrade the completed amplifier's noise.

Finally, many amplifiers have good looking noise specs, but we must warn you to examine the noise-versus-frequency curves on their datasheets; many have very high  $1/f$  noise corners. This is especially true for current noise, with some  $1/f$  corners way up at 1 MHz or beyond. An example of a troublesome spec might be the THS4508, with its *pnp* input transistors (for operation to GND) and  $4.7 \text{ pA}/\sqrt{\text{Hz}}$  of current noise at 1 MHz. This creates  $1.6 \text{ nV}/\sqrt{\text{Hz}}$  across a  $349 \Omega$  resistor, which is OK compared with the part's  $e_n = 2.3 \text{ nV}/\sqrt{\text{Hz}}$ . But if you were to use 1k resistors, the corresponding current-induced noise voltage would be  $4.7 \text{ nV}/\sqrt{\text{Hz}}$ , dominating the amplifier's  $e_n$ . Depending on the application, this might be of concern. Or it might not.

## Review of Chapter 5

An A-to-M summary of what we have learned in Chapter 5. This summary reviews basic principles and facts in Chapter 5, but it does not cover application circuit diagrams and practical engineering advice presented there.

### ¶A. Precision and Dynamic Range.

A *precision* circuit is one that exhibits (through careful design and choice of components) both initial *accuracy*, and also *stability* (i.e., maintenance of accuracy over time and temperature). A precision circuit may (but need not) exhibit wide *dynamic range* (the ratio of signal amplitudes over which it operates); see §5.1.

### ¶B. Error Budget.

When designing a precision circuit, you need to keep track of numerous error contributions (from voltage offsets, current-induced offsets, component tolerances, and the like); this is best tallied in an *error budget*, which promotes design discipline, and which assists in helping you spot the dominant error sources; see the example in §5.5.

### ¶C. Strict versus Pragmatic.

Strict worst-case design mandates that all components be operated within their datasheet specifications, and that the effects of all their worst-case errors be added (as unsigned magnitudes) to determine the circuit's performance. The benefit of such conservatism is a circuit guaranteed to perform within specifications (assuming proper design); the downside is that it may be impossible to achieve design goals with available and/or affordable components, taking their worst-case specifications in a worst-case arithmetical combination (and noting that some critical performance parameters may be unspecified, or show only "typical" values). We favor a pragmatic approach (§5.3), taking with a grain of salt some of the published worst-case parameters (for example, input leakage current, where testing limits encourage highly conservative worst-case specs), or adopting reasonable estimates of unspecified parameters. It may be sufficient to establish that the circuit effects of the unspecified parameter are completely insignificant; or, if it's a closer call, you may have to set up a testing regime of incoming components to ensure that you meet specs. If have a situation in which there are many components contributing to an overfull error budget, you may have to validate the performance of the overall circuit, subassembly, or complete instrument at final test.

### ¶D. Component Errors – Resistors.

Taking a simple example, resistor accuracy (including the effects of tempco) sets a limit on the accuracy of an amplifier's gain. But it's not that simple, because amplifier gain usually depends on a resistance *ratio*; so the situation is greatly improved if you use a resistor pair with accurate ratio and matched tempco, and it's the *ratio* mismatch and the tempco of the *ratio* that are limiting. Likewise, the CMRR limit of a difference amplifier depends on resistor-pair matching,  $\text{CMRR}(\text{dB}) \approx 20\log(100/p)$ , where  $p$  is the ratio mismatch in percent. To give scale to these statements, typical metal-film resistors come in accuracy ranges of 0.05%–1%, with a typical tempco spec of 25–100 ppm/ $^{\circ}\text{C}$ ; resistor arrays intended for high precision have accuracies in the range of 0.01%–0.05%, tempcos down to 1 ppm/ $^{\circ}\text{C}$  or better, and matching accuracies and tempcos down to 0.01%–0.001% and 1–0.1 ppm/ $^{\circ}\text{C}$ , respectively. Real resistors depart from the ideal also in *linearity*, i.e., they exhibit some resistance change with applied voltage. See §5.6.1; see also §1x.2 for further discussion of resistors and their vicissitudes (e.g., long-term drift, and parasitic inductance and capacitance).

### ¶E. Component Errors – Capacitors and Switches.

Capacitors have several amusing traits that affect the accuracy of integrators and of sample-and-hold circuits; these include resistive leakage (thus exponential decay), and more seriously *dielectric absorption* (memory effect, see §5.6.2, §1x.3, and the dielectric-absorption plots in that section.) These application circuits include analog switches (for integrator reset, and sample-to-hold switching), which introduce errors through leakage and charge injection (§§3.4.2E and 5.6.3).

### ¶F. Amplifier Input Errors.

This is where most of your troubles are located. The primary qualification for membership in the category of *precision op-amps* is a small offset voltage  $V_{\text{os}}$ , and a correspondingly low tempco TCV $_{\text{OS}}$  (sometimes called  $\Delta V_{\text{OS}}$ ). The offset voltage operates at the input, so the RTI (referred to the input) error is simply  $V_{\text{os}}$ ; at the output of an amplifier of voltage gain  $G_V$  the  $V_{\text{os}}$ -induced error is  $\Delta V_{\text{out}} = G_V V_{\text{os}}$ . In an integrator circuit with input resistor  $R_{\text{in}}$ , the input offset voltage is equivalent to an input current error of  $\Delta I_{\text{in}} = V_{\text{os}}/R_{\text{in}}$ . To give an idea of scale, a typical precision op-amp has an offset voltage of 10–50  $\mu\text{V}$ , and a tempco of 0.1–0.5  $\mu\text{V}/^{\circ}\text{C}$ . *Auto-zero* op-amps (see ¶I) improve on these figures by about a factor of ten; see Table 5.5 on pages 320–321 and Table 5.6 on 335. See also §5.10.5.

But there's more. Input bias current  $I_B$  flowing through

the source resistance  $R_S$  seen at the amplifier's input produces an RTI voltage error of  $\Delta V_{in} = I_B R_S$ . For precision *bipolar* op-amps, whose bias currents are of order 10 nA, this becomes serious for source resistances greater than a few kilohms (where  $I_B R_S$  amounts to a few tens of microvolts, in the same ballpark as a precision op-amp's  $V_{os}$ ). High- $R_S$  situations thus mandate low-bias op-amps, usually those with JFET or MOSFET inputs, or (for moderate source resistance) a bipolar precision op-amp with bias-cancellation (where  $I_B$  is of order 1 nA). A *warning*: the very low bias current of FET-input op-amps rises dramatically at higher temperatures (see Figure 5.6), where its input current may even exceed that of a bias-cancelled bipolar op-amp. See also §5.10.7.

Looking deeper, additional error sources at op-amp inputs include  $I_B$  variation with common-mode input voltage  $V_{CM}$  (§5.7.2, Figure 5.7),  $V_{os}$  variation with  $V_{CM}$  (i.e., CMRR, §5.7.3, §5.10.9, and Figures 5.10, 5.29, and 5.30), PSRR, and input noise ( $e_n$ ,  $i_n$ , §§5.10.6 and 5.10.8).

### ¶G. Amplifier Output Errors.

While much of precision analog design concerns dc and low frequencies, some applications require accuracy at higher speeds: audio and video, communications, scientific measurement, and so on. With falling op-amp loop gain, input errors are rising, output impedance is rising, and slew-rate limitations may come into play, along with reduced suppression of output-stage crossover distortion, gain nonlinearity, and phase error. And overshoot, ringing, and settling time are critical in dynamic applications.

These can be made quantitative. The gain error  $\delta_G \equiv (G_{ideal} - G_{actual})/G_{ideal} = 1/(1 + AB)$ , where  $B$  is the feedback fraction around open-loop gain  $A$ ; see §5.8.5. The closed-loop bandwidth is  $f_{3dB} = f_T/G_{CL}$ , corresponding to a time constant of approximately  $\tau \approx G_{CL}/2\pi f_T$ , which (if well compensated, with good phase margin) produces a settling time of order  $5-10\tau$ ; see §§5.8.2 and 5.10.10. An op-amp's *actual* settling time can be considerably longer, and there's no substitute for real data; see Table 5.5 on pages 320–321. A well-compensated op-amp exhibits a closed-loop phase error of  $\phi \approx f/f_C$ , where  $f_C = f_T/G_{CL}$  is the frequency at which the closed-loop gain has fallen to unity.<sup>115</sup> Op-amp distortion depends strongly on the output-stage circuit; see the extensive plots in Figures 5.43 and 5.44.

### ¶H. Rail-to-rail Op-amps.

It's tempting to choose op-amps with rail-to-rail common-mode input range (RRI), rail-to-rail output (RRO), or both (RRIO), especially for operation at low supply voltages.

**RRI op-amps.** But there are drawbacks associated with the duplicate complementary input stages that can seriously compromise precision, notably an abrupt change of both  $I_B$  and  $V_{os}$  at the input crossover voltage (§5.9.1). Some RRIO op-amps (e.g., the OPA360-series) circumvent this problem by using an on-chip charge pump. If you don't need full rail-to-rail input, choose an op-amp whose input common-mode range extends only to one rail (usually the negative rail).

**RRO op-amps.** Op-amps that feature rail-to-rail output have their own issues, stemming from the use of a common-source (or common-emitter) output-stage topology, rather than the conventional source follower (or emitter follower). These include increased distortion, and higher output impedance (thus gain and phase shift are more susceptible to load impedance). However, many RRO op-amps mitigate this latter problem by using internal capacitive feedback to lower the open-loop output impedance at high frequencies, as seen for example in Figure 5.34.

### ¶I. Auto-zero Op-amps.

This subclass of op-amps includes on-chip offset-nulling circuitry that operates cyclically (with an on-chip oscillator) to trim the input offset voltage (§5.11). As the entries in Table 5.6 on page 335 demonstrate, this produces typical  $V_{os}$  values around a microvolt, with tempcos around  $10\text{nV}/^\circ\text{C}$ , roughly an order of magnitude better than the best conventional precision op-amps. Notable, too, is the absence of a  $1/f$  low-frequency rising noise voltage (see for example Figures 5.53 and 5.54).

That's the good news. The bad news is that the switching action produces noise-spectra peaks at the switching frequency and its harmonics (Figure 5.51), superposed on a low-frequency noise voltage floor that is already considerably higher than that of conventional low-noise op-amps (compare Table 5.6 on page 335 with Table 5.5 on pages 320–321). The input switching circuitry also results in relatively high input noise *current*, compared with low-noise precision JFET-input op-amps.

Auto-zero op-amps are built with CMOS, and (except for some legacy parts) are generally limited to low supply voltages ( $\leq 7\text{ V}$  total supply). A shining exception is the recent LTC2057HV (4.75 V to 60 V total supply!). Another caution: because of their internal voltage-storage

<sup>115</sup> Throughout the book we use  $f_T$  as shorthand for the proper term *gain-bandwidth product* (GBP, GBW, or GBWP), which is the extrapolated unity-gain crossover frequency.

correction capacitors, auto-zero op-amps can exhibit slow recovery from saturation, as long as several milliseconds.

#### ¶J. Difference, Differential, and Instrumentation

##### Amplifiers: Taxonomy.

These share the property of accurate and stable differential gain, with high common-mode rejection:  $V_{\text{out}} = G_V \Delta V_{\text{in}} = G_V (V_{\text{in+}} - V_{\text{in-}})$ . In common usage, the terms are distinguished as follows.

##### Difference amplifier.

Differential-in, single-ended out; op-amp plus two matched resistor pairs (Figure 5.65, §5.14); CMRR 90–100 dB; accurate but low gain ( $G_V = 0.1$ –10); input impedance 25–100 k $\Omega$ , intended to be driven by a low impedance; inputs typically can go beyond rails.

##### Instrumentation amplifier.

Differential-in, single-ended out; very high input impedance (10 M $\Omega$  to 10 G $\Omega$ ), wide gain range ( $G_V = 1$ –1000), and very high CMRR at higher gains (110–140 dB at  $G_V = 100$ ); see §5.15 and Figure 5.77.

##### Fully differential amplifier.

Differential or single-ended in, differential-out; most are low voltage, fast-settling,

and wideband; ideal for twisted-pair cable drivers and fast differential-input ADCs; see §5.17.

#### ¶K. Difference Amplifiers.

The classic *difference amplifier*, typified by the original INA105 (see Table 5.7 on page 353, §5.14), consists of an op-amp plus a pair of matched resistor dividers, with SENSE and REF inputs (at the bottom of the divider strings) in addition to the  $V_{\text{in+}}$  and  $V_{\text{in-}}$  signal inputs. With REF grounded and SENSE tied to the output, the gain is  $G_{\text{diff}} \equiv V_{\text{out}} / \Delta V_{\text{in}} = R_f / R_i$ , where  $R_f$  and  $R_i$  are the feedback and input resistors, respectively. Figure 5.66 shows circuit variations that exploit alternative connections of the SENSE and REF pins.

Difference amplifiers are simple, and good enough for many applications. Some permit common-mode inputs well beyond the rails (e.g.,  $\pm 200$  V for the INA117). But difference amplifiers suffer from relatively low input impedance (tens of k $\Omega$ ), limited gain (typically in the range of  $G=1$ –10), relatively high noise voltage, unimpressive CMRR (typically  $\lesssim 90$  dB), and degradation of both gain and CMRR when driven with signals of non-zero source impedance.

#### ¶L. Instrumentation Amplifiers.

These drawbacks are nicely remedied in the *instrumentation amplifier* configuration of three op-amps (Figure 5.77, §5.15). The input impedance is high (10 M $\Omega$ –1 T $\Omega$ ), and the gain (which can be as large as  $\times 1000$ ) is set with a single external resistor (or pin-selectable choice of internal resistors). When configured for high gain, most instrumentation amplifiers deliver a typical CMRR around 120 dB, and an input noise voltage  $e_n$  less than 10 nV/ $\sqrt{\text{Hz}}$ ; they do not permit common-mode outputs beyond the rails. Instrumentation amplifiers are available in BJT, JFET, MOSFET, auto-zero, and programmable-gain types; see Tables 5.8 on page 363 and 5.9 on pages 370–371.

#### ¶M. Fully Differential Amplifiers.

Unlike difference amplifiers and instrumentation amplifiers, fully differential amplifiers (§5.17) generate a balanced *differential output* centered on a settable common-mode voltage. This is useful for driving fast ADCs with complementary inputs, balanced transmission lines, and RF communication circuits. Befitting these applications, they tend to be wideband (up to 1 GHz or more), high slew-rate (1000 V/ $\mu$ s or more), fast settling (a few nanoseconds), and quiet (5 nV/ $\sqrt{\text{Hz}}$  or less). There are many internal circuit topologies (labeled A–G in §5.17); see Table 5.10 for a representative selection.

# MATH REVIEW

## APPENDIX A

Some knowledge of algebra and trigonometry is essential for a full understanding of this book. In addition, a limited ability to deal with complex numbers and derivatives (a part of calculus) is helpful, although not entirely essential. This appendix is meant as the briefest of summaries of complex numbers and differentiation, preceded by a collection of useful formulas from trigonometry, exponentials, and logarithms. It is not meant as a textbook substitute. For a highly readable self-help book on calculus, we recommend *Quick Calculus*, by D. Kleppner and N. Ramsey, Wiley, 2nd ed., 1985.

### A.1 Trigonometry, exponentials, and logarithms

Here is a collection of useful formulas:

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

is the solution of the quadratic equation

$$ax^2 + bx + c = 0.$$

$$\sin(x \pm y) = \sin x \cos y \pm \cos x \sin y,$$

$$\cos(x \pm y) = \cos x \cos y \mp \sin x \sin y,$$

$$\sin 2x = 2 \sin x \cos x,$$

$$\cos x \cos y = \frac{1}{2} [\cos(x+y) + \cos(x-y)],$$

$$\cos x \sin y = \frac{1}{2} [\sin(x+y) - \sin(x-y)],$$

$$\sin x \sin y = \frac{1}{2} [\cos(x-y) - \cos(x+y)]$$

$$e^{x+y} = e^x e^y,$$

$$e^{x-y} = e^x / e^y,$$

$$x^{a/b} = \sqrt[b]{x^a},$$

$$e^{\log_e x} = x,$$

$$\log_e(xy) = \log_e x + \log_e y,$$

$$\log_e(x/y) = \log_e x - \log_e y,$$

$$\log_e x^n = n \log_e x,$$

$$\log_e e^x = x,$$

$$\log_e x = \log_e 10 \log_{10} x \approx 2.3 \log_{10} x,$$

$$a^x = e^{x \log_e a}.$$

### A.2 Complex numbers

A complex number is an object of the form

$$\mathbf{N} = a + ib,$$

where  $a$  and  $b$  are real numbers and  $i$  is the square root of  $-1$ ;  $a$  is called the real part, and  $b$  is called the imaginary part.<sup>1</sup> Boldface letters or squiggly underlines are sometimes used to denote complex numbers. At other times you're just supposed to know!

Complex numbers can be added, subtracted, multiplied, etc., just as real numbers:

$$(a + ib) + (c + id) = (a + c) + i(b + d),$$

$$(a + ib) - (c + id) = (a - c) + i(b - d),$$

$$(a + ib)(c + id) = (ac - bd) + i(bc + ad),$$

$$\frac{a + ib}{c + id} = \frac{(a + ib)(c - id)}{(c + id)(c - id)} = \frac{ac + bd}{c^2 + d^2} + \frac{bc - ad}{c^2 + d^2} i.$$

<sup>1</sup> Electrical engineers depart from the universal convention of  $i \equiv \sqrt{-1}$ , using instead the symbol  $j$  in order to avoid duplicating the use of the symbol  $i$  (which designates small-signal current). We follow the EEs in this book, but not in this Math appendix. Were we to do so, we would likely be disowned by our math colleagues.

All these operations are natural, in the sense that you just treat  $i$  as something that multiplies the imaginary part, and go ahead with ordinary arithmetic. Note that  $i^2 = -1$  (used in the multiplication example) and that division is simplified by multiplying top and bottom by the *complex conjugate*, the number you get by changing the sign of the imaginary part. The complex conjugate is sometimes indicated with an asterisk. If

$$\mathbf{N} = a + ib,$$

then

$$\mathbf{N}^* = a - ib.$$

The magnitude (or *modulus*) of a complex number is a real number with no imaginary part:

$$|\mathbf{N}| = |a + ib| = \sqrt{(a + ib)(a - ib)} = \sqrt{a^2 + b^2},$$

i.e.,

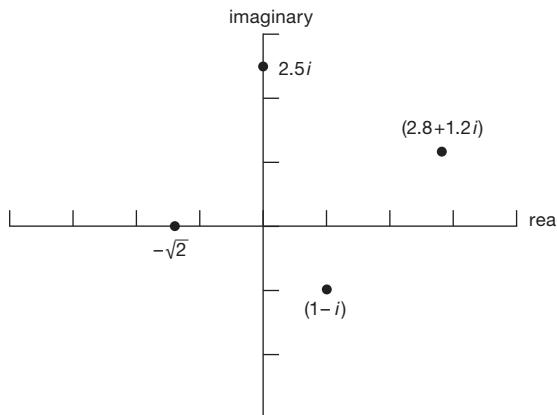
$$|\mathbf{N}| = \sqrt{\mathbf{N}\mathbf{N}^*},$$

simply obtained by multiplying by the complex conjugate and taking the square root. The magnitude of the product (or quotient) of two complex numbers is simply the product (or quotient) of their magnitudes.

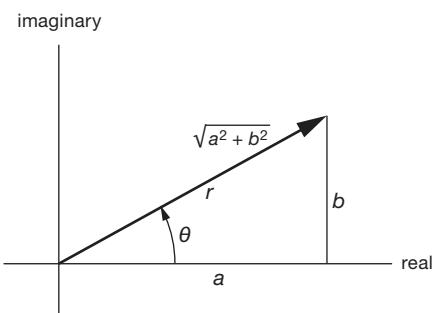
The real (or imaginary) part of a complex number is sometimes written as

$$\begin{aligned} \text{real part of } \mathbf{N} &= \mathcal{R}e(\mathbf{N}), \\ \text{imaginary part of } \mathbf{N} &= \mathcal{I}m(\mathbf{N}). \end{aligned}$$

You get them by writing out the number in the form  $a + ib$ , then taking either  $a$  or  $b$ . This may involve some multiplication or division, since the complex number may be a real mess.



**Figure A.1.** Complex numbers in the “complex plane.”



**Figure A.2.** Complex numbers, as magnitude and angle.

Complex numbers are sometimes represented on the complex plane. It looks just like an ordinary  $x,y$  graph, except that a complex number is represented by plotting its real part as  $x$  and its imaginary part as  $y$ , as shown in Figure A.1. In keeping with this analogy, you sometimes see complex numbers written just like  $x,y$  coordinates:

$$a + ib \leftrightarrow (a, b).$$

Just as with ordinary  $x,y$  pairs, complex numbers can be represented in polar coordinates; that’s known as “magnitude, angle” representation. For example, the number  $a + ib$  can also be written as (Figure A.2)

$$a + ib = r\angle\theta,$$

where<sup>2</sup>  $r = \sqrt{a^2 + b^2}$  and  $\theta = \tan^{-1}(b/a)$ . This is usually written in a different way, using the astonishing fact that

$$e^{i\theta} = \cos\theta + i\sin\theta.$$

(You can derive the preceding result, known as Euler’s<sup>3</sup> formula, by expanding the exponential in a Taylor series.) Thus we have the following equivalents:

$$\begin{aligned} \mathbf{N} &= a + ib = re^{i\theta}, \\ r &= |\mathbf{N}| = \sqrt{\mathbf{N}\mathbf{N}^*} = \sqrt{a^2 + b^2}, \\ \theta &= \tan^{-1}(b/a), \end{aligned}$$

i.e., the modulus  $r$  and angle  $\theta$  are simply the polar coordinates of the point that represents the number in the complex plane. Polar form is handy when complex numbers have to be multiplied; you just multiply their magnitudes and add their angles (or, to divide, you divide their magnitudes and subtract their angles):

$$(r_1 e^{i\theta_1})(r_2 e^{i\theta_2}) = r_1 r_2 e^{i(\theta_1 + \theta_2)}.$$

<sup>2</sup> Caution: the formula for  $\theta$  returns values only between  $-\pi/2$  and  $+\pi/2$ ; the signs of both  $a$  and  $b$ , and not merely their quotient, are required for a correct value of  $\theta$  in all four quadrants.

<sup>3</sup> Leonhard Euler, pronounced like “oiler.”

Finally, to convert from polar to rectangular form, just use Euler's formula:

$$re^{i\theta} = r \cos \theta + i r \sin \theta,$$

i.e.,

$$\Re(re^{i\theta}) = r \cos \theta,$$

$$\Im(re^{i\theta}) = r \sin \theta.$$

(These can be used to easily derive the sum and difference of trigonometric functions, so you never have to remember those pesky formulas. Just work out  $e^{i(x \pm y)}$ .)

If you have a complex number multiplying a complex exponential, just do the necessary multiplications. If

$$\mathbf{N} = a + ib,$$

$$\mathbf{N}e^{i\theta} = (a+ib)(\cos \theta + i \sin \theta),$$

$$= (a \cos \theta - b \sin \theta),$$

$$+ i(b \cos \theta + a \sin \theta).$$

When dealing with circuits and signals, the angular argument  $\theta$  often takes the form of an evolving wave:  $\theta = \omega t = 2\pi ft$ ; thus, for example,  $V(t) = \Re(V_0 e^{i\omega t}) = V_0 \cos \omega t$ , etc.

### A.3 Differentiation (Calculus)

We start with the concept of a *function*  $f(x)$ , i.e., a formula that gives a value  $y = f(x)$  for each  $x$ . The function  $f(x)$  should be *single valued* i.e., it should give a single value of  $y$  for each  $x$ . You can think of  $y = f(x)$  as a graph, as in Figure A.3. The derivative of  $y$  with respect to  $x$ , written  $dy/dx$  ("dee  $y$  dee  $x$ "), is the *slope* of the graph of  $y$  versus  $x$ . If you draw a tangent to the curve at some point, its slope is  $dy/dx$  *at that point*; i.e., the derivative is itself a function, since it has a value at each point. In Figure A.3 the slope at the point  $(1, 1)$  happens to be 2, whereas the slope at the origin is zero (we'll see shortly how to compute the derivative).

In mathematical terms, the derivative is the limiting value of the ratio of the change in  $y$  ( $\Delta y$ ) to the change in  $x$  ( $\Delta x$ ), as  $\Delta x$  goes to zero. To quote a song once sung in the hallowed halls of Harvard (by Tom Lehrer and Lewis Branscomb),

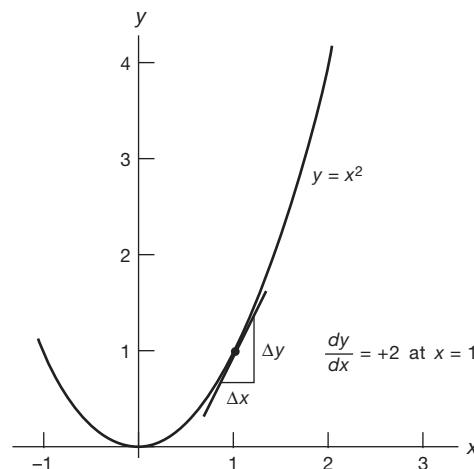
You take a function of  $x$ , and you call it  $y$

Take any  $x$ -nought that you care to try

Make a little change and call it delta  $x$

The corresponding change in  $y$  is what you find  
nex'

And then you take the quotient, and now, carefully



**Figure A.3.** A single-valued function:  $f(x) = x^2$ .

Send delta  $x$  to zero, and I think you'll see  
That what the limit gives us (if our work all checks)  
Is what you call  $dy/dx$ ...  
It's just  $dy/dx$ .

(*The Derivative Song*, sung to the tune of *There'll Be Some Changes Made*, W. Benton Overstreet).

Differentiation is a straightforward art, and the derivatives of many common functions are tabulated in standard tables and automatically calculated in programs like Mathematica®. Here are some rules ( $u$  and  $v$  are arbitrary functions of  $x$ , and  $a$  represents a constant).

#### A.3.1 Derivatives of some common functions

$$\frac{d}{dx} a = 0$$

$$\frac{d}{dx} ax = a$$

$$\frac{d}{dx} ax^n = anx^{n-1},$$

$$\frac{d}{dx} \sin ax = a \cos ax,$$

$$\frac{d}{dx} \cos ax = -a \sin ax,$$

$$\frac{d}{dx} e^{ax} = ae^{ax},$$

$$\frac{d}{dx} \log_e x = 1/x.$$

### A.3.2 Some rules for combining derivatives

Here  $u(x)$  and  $v(x)$  represent generic functions of  $x$ :

$$\frac{d}{dx} au(x) = a \frac{d}{dx} u(x),$$

$$\frac{d}{dx}(u+v) = \frac{du}{dx} + \frac{dv}{dx},$$

$$\frac{d}{dx} uv = u \frac{d}{dx} v + v \frac{d}{dx} u,$$

$$\frac{d}{dx} \left( \frac{u}{v} \right) = \frac{v \frac{du}{dx} - u \frac{dv}{dx}}{v^2},$$

$$\frac{d}{dx} \log_e u = \frac{1}{u} \frac{du}{dx}$$

$$\frac{d}{dx} \{u[v(x)]\} = \frac{du}{dv} \frac{dv}{dx}.$$

The last one is very useful and is called the chain rule.

### A.3.3 Some examples of differentiation

$$\frac{d}{dx} x^2 = 2x,$$

$$\frac{d}{dx} (1/x^{\frac{1}{2}}) = -\frac{1}{2} x^{-\frac{3}{2}},$$

$$\frac{d}{dx} xe^x = xe^x + e^x \quad (\text{product rule}),$$

$$\frac{d}{dx} e^{-x^2} = -2xe^{-x^2} \quad (\text{chain rule}),$$

$$\frac{d}{dx} a^x = \frac{d}{dx} (e^{x \log_e a}) = a^x \log_e a \quad (\text{chain rule}).$$

Once you have differentiated a function, you often want to evaluate the value of the derivative at some point. Other times you may want to find a minimum or maximum of the function; that's the same thing as having a zero derivative, so you can just set the derivative equal to zero and solve for  $x$ . For example, you can easily determine that the slope of the function plotted in Figure A.3 equals 2 at  $x=1$ , and that its minimum occurs at  $x=0$  (where its slope is zero).

# HOW TO DRAW SCHEMATIC DIAGRAMS

## APPENDIX

# B

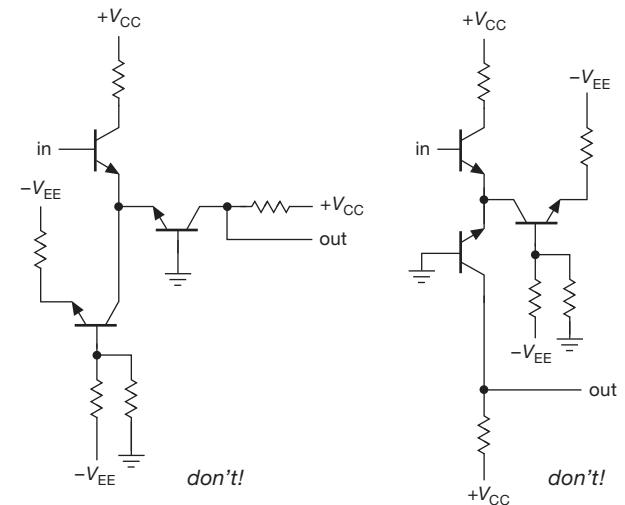
A well-drawn schematic makes it easy to understand how a circuit works, and it aids greatly in troubleshooting. A poor schematic only creates confusion. By keeping a few rules and suggestions in mind, you can draw a good schematic in no more time than it takes to draw a poor one. In this appendix we dispense advice of three varieties: general principles, rules, and hints. We have also drawn some real knee-slappers to illustrate habits to avoid.

### B.1 General principles

- Schematics should be unambiguous. Therefore pin numbers, parts values, reference designators, polarities, etc., should be clearly labeled to avoid confusion.
- A good schematic makes circuit functions clear. Therefore keep functional areas distinct; don't be afraid to leave blank areas on the page, and don't try to fill the page. There are conventional ways to draw functional subunits; for instance, don't draw a differential amplifier as in Figure B.1, because the function won't be easily recognized. Likewise, flip-flops are usually drawn with clock and inputs on the left, set and clear on top and bottom, and outputs on the right.

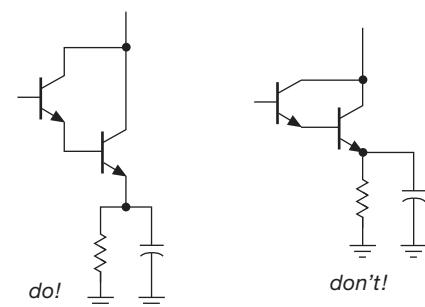
### B.2 Rules

- Wires connecting are indicated by heavy black dots; wires crossing, but not connecting, have no dot (don't use a little half-circular "jog"; it went out in the 1950s).
- Four wires must not connect at a point; i.e., wires must not cross *and* connect. You sometimes see this rule violated, but it's poor practice (because a missing or undersized dot is a different circuit).
- Always use the same symbol for the same device; e.g., don't draw flip-flops in two different ways (exception: assertion-level logic symbols show each gate in two possible ways).
- Wires and components are aligned horizontally or vertically, unless there's a good reason to do otherwise.



**Figure B.1.** Arrange components so that the function (here a differential amplifier) is clear. Don't corrupt the presentation to save space.

- Label pin numbers on the outside of a symbol, signal names on the inside.
- All parts should have values or types indicated; it's best to give all parts a label ("refdes"), too, e.g.,  $R_7$  or  $U_3$ .



**Figure B.2.** Bring leads away from component symbols before connecting or jogging.

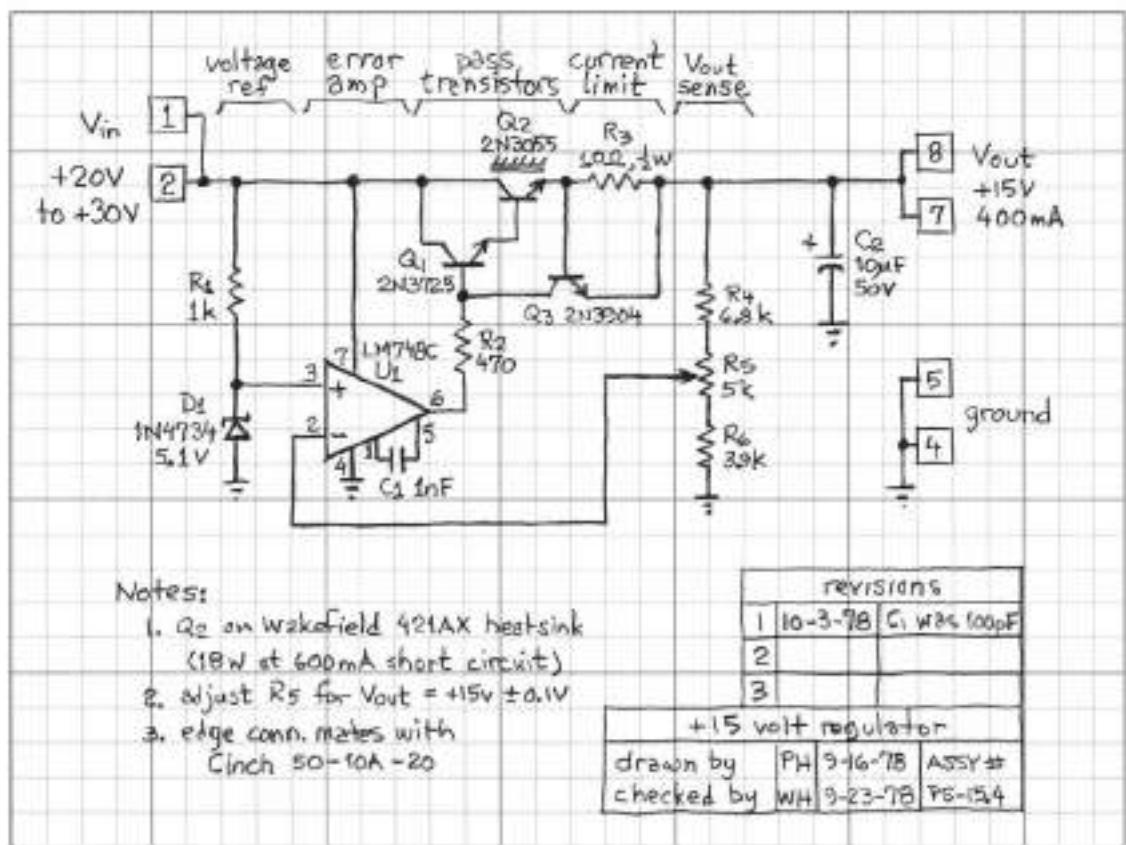
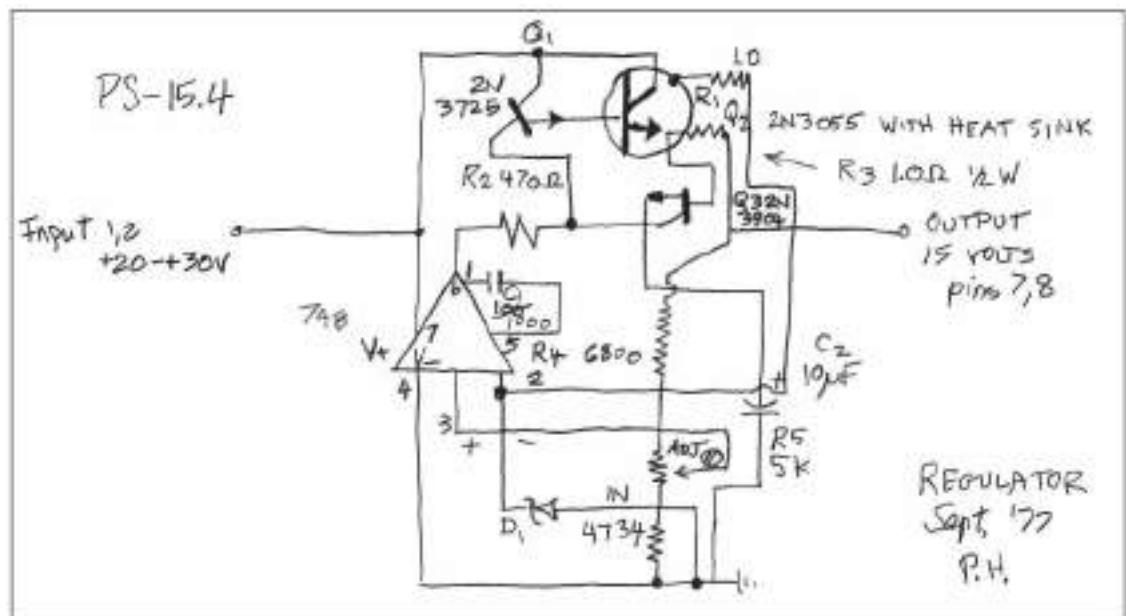


Figure B.3. Good drawing (perhaps too good) on “engineering paper,” and very bad. Guess which is which.

### B.3 Hints

- Identify parts immediately adjacent to the symbol, forming a distinct group giving symbol, label, and type or value.
- In general, signals go from left to right; don't be dogmatic about this, though, if clarity is sacrificed.
- Put positive supply voltages at the top of the page, negative at the bottom. Thus *npn* transistors will usually have their emitter at the bottom, whereas *pnp*s will have their emitter topmost.
- Don't attempt to bring all wires around to the supply rails, or to a common ground wire. Instead, use the ground symbol(s) and labels like  $+V_{CC}$  to indicate those voltages where needed.
- It is helpful to label signals and functional blocks and show waveforms; in logic diagrams it is especially important to label signal lines, e.g., RESET' or CLK.
- It is helpful to bring leads<sup>1</sup> away from components a short distance before making connections or jogs. For example, draw transistors as in Figure B.2.
- Leave some space around circuit symbols; e.g., don't draw components or wires too close to an op-amp symbol. This keeps the drawing uncluttered and leaves room for labels, pin numbers, etc.
- Label all boxes that aren't obvious: comparator versus op-amp, shift register versus counter, etc. Don't be afraid to invent a new symbol.
- Use small rectangles, ovals, or circles to indicate card-edge connections, connector pins, etc. Be consistent.

<sup>1</sup> Leads? Yeah, sure... I'll just check with the boys down at the crime lab. They got four more detectives working on the case. They got us working in *shifts*. Hahahaha... LEADS!

- The signal path through switches should be clear. Don't force the reader to follow wires all over the page to find out how a signal is switched.
- Power-supply connections are normally assumed for op-amps and logic devices. However, show any unusual connections (e.g., an op-amp run from a single supply, where  $V_-$  = ground), and the disposition of unused inputs.
- It is very helpful to include a small table of integrated circuit (IC) numbers, types, and power-supply connections (pin numbers for  $V_{CC}$  and ground, for instance).
- Include a title area near the bottom of the page, with name of circuit, name of instrument, by whom drawn, by whom designed or checked, date, and assembly number. Also include a revision area, with columns for revision number, date, and subject.
- We recommend drawing schematics freehand on coarse graph paper (pale gridlines, five per inch, for example National® Brand "Engineer's Computation Pad" in "Eye-Ease"® green), or on plain paper on top of graph paper. This is fast, and it gives very pleasing results. Use dark pencil (we like HB hardness, 0.5 mm diameter) or ink; avoid ballpoint or felt-tip pen.

### B.4 A humble example

As an illustration, we've drawn a humble example (Figure B.3) showing "awful" and "good" schematics of the same circuit; the former violates nearly every rule and is almost impossible to understand. See how many bad habits you can find illustrated. We've seen all of them in professionally drawn schematics! (We drew the "bad" schematic in an airport while waiting for a flight. It was an occasion of great hilarity; we laughed ourselves silly.)

# RESISTOR TYPES

## APPENDIX

C

### C.1 Some history

For a half century people used “leaded” (pronounced lēd'-ed) resistors: if you look inside a very old radio (before ~1950), you’ll see colored cylindrical objects with some colored dots painted on them and with a wire wrapped around each end that comes off perpendicularly to the axis (“radial leads”). These carbon-composition resistors evolved into the standard carbon-comp “axial-lead” resistors (still cylindrical, but with colored stripes all the way around, and with the wires now sticking straight out each end) that were dominant through the last half of the 20th century (and that we recommended for noncritical applications in our previous book editions). Axial-lead resistors are still popular for some uses, such as easy breadboarding in the lab. They’re also used in applications that require very high resistance ( $\geq 100\text{ M}\Omega$ ), or high voltage or power ratings, or for resistors of very high precision.

However, contemporary electronics has embraced surface-mount packaging, because of its high density (SMT devices are *small*, and you don’t have to take up space with holes for the leads). Surface-mount resistors, in common with other two-terminal SMT components (capacitors, inductors), are available in a range of package sizes, characterized by a 4-digit code giving their length and width in units of 0.010”; for example, an “0603” package is  $0.06'' \times 0.03''$  (1.5 mm  $\times$  0.75 mm). We favor that size, or the larger 0805 package, for general prototyping of surface-mount circuits. The smaller packages (0402, 0201, and even “01005”) are a major pain – you basically have to work under a microscope (and don’t sneeze).

an analogous 5% resistor,<sup>1</sup> so you might as well use 1% resistors by default. They come in the E96 set of standard values (96 values per decade, spaced approximately 2% apart; thus 481 values from  $10\Omega$  through  $1\text{ M}\Omega$ , see below). Resistors of greater precision (e.g., 0.1%) are sometimes available in the E192 superset,<sup>2</sup> and in convenient round-number values (e.g., 250, 300, 400, or 500) that are not included in the EIA sequences.

Here is the E24 set of “5%” values (the E12 subset, used for 10% components, is shown in **bold**):

<b>10</b>	16	<b>27</b>	43	<b>68</b>
11	<b>18</b>	30	<b>47</b>	75
<b>12</b>	20	<b>33</b>	51	<b>82</b>
13	<b>22</b>	36	<b>56</b>	91
<b>15</b>	24	<b>39</b>	62	<b>100</b>

And here is the E96 set of “1%” values (the E48 set, used for 2% components, or for a reduced set of 1% parts, is in **bold**):

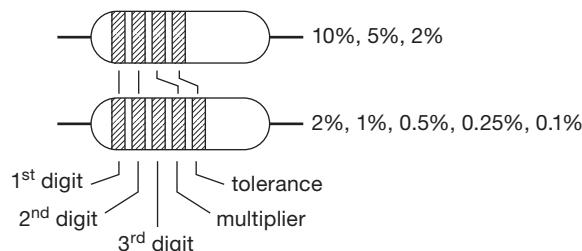
<b>100</b>	137	<b>187</b>	255	<b>348</b>	475	<b>649</b>	887
102	<b>140</b>	191	<b>261</b>	357	<b>487</b>	665	<b>909</b>
<b>105</b>	143	<b>196</b>	267	<b>365</b>	499	<b>681</b>	931
107	<b>147</b>	200	<b>274</b>	374	<b>511</b>	698	<b>953</b>
<b>110</b>	150	<b>205</b>	280	<b>383</b>	523	<b>715</b>	976
113	<b>154</b>	210	<b>287</b>	392	<b>536</b>	732	
<b>115</b>	158	<b>215</b>	294	<b>402</b>	549	<b>750</b>	
118	<b>162</b>	221	<b>301</b>	412	<b>562</b>	768	
<b>121</b>	165	<b>226</b>	309	<b>422</b>	576	<b>787</b>	
124	<b>169</b>	232	<b>316</b>	432	<b>590</b>	806	
<b>127</b>	174	<b>237</b>	324	<b>442</b>	604	<b>825</b>	
130	<b>178</b>	243	<b>332</b>	453	<b>619</b>	845	
<b>133</b>	182	<b>249</b>	340	<b>464</b>	634	<b>866</b>	

<sup>1</sup> For example, the Digi-Key catalog shows a full selection of Vishay/Dale CRCW-series surface-mount resistors, in sizes from 1210 down to 0201. For the convenient 0603 size, the current prices for 1% and 5% resistors are \$0.025 and \$0.023 apiece, respectively, in quantity 200. (You’ll pay about triple that, in quantity 10, and about a fifth as much, in a full reel of 5000 resistors.)

<sup>2</sup> The full E192 set, along with subsets, is nicely displayed at [http://www.logwell.com/tech/components/resistor\\_values.html](http://www.logwell.com/tech/components/resistor_values.html).

### C.3 Resistance marking

Leaded resistors are marked in one of two ways: (a) with a set of four or five color bands, indicating resistance and tolerance; or (b) with a 4-digit resistance code, followed by a letter that indicates the tolerance. Surface-mount resistors use either (a) a 3- or 4-digit resistance code, or, for the smallest package sizes, (b) no marking at all!



color	digit	multiplier	tolerance	(tol. suffix)
black	0	1	–	
brown	1	10	1%	F
red	2	100	2%	G
orange	3	1k	–	
yellow	4	10k	–	
green	5	100k	0.5%	D
blue	6	1M	0.25%	C
violet	7	10M	0.1%	B
gray	8	–	0.05%	A, W
white	9	–	–	
gold	–	0.1	5%	J
silver	–	0.01	10%	K
(none)	–	–	20%	M
		0.02%		N, Q, P
		0.01%		T, L
		0.005%		V
		0.0025%		X
		0.002%		U
		0.001%		S

**Figure C.1.** The resistor color code, used on some axial-lead resistors (notably carbon-film and carbon-composition types). The resistance is read as a 2- or 3-digit integer (depending on resistor precision) followed by a band indicating the power-of-10 multiplier. For example, yellow-violet-orange-gold is  $47\text{ k}\Omega \pm 5\%$ , and yellow-white-white-black-brown is  $499\Omega \pm 1\%$ . The alphabetic tolerance suffix is used on resistors with numerical printed resistance values.

Although it may seem diabolical to the beginner, the practice of color banding makes it easy to recognize resistor values in a circuit or parts bin, without having to search for a printed legend. Each color corresponds to a digit, in a sort of floating-point format (with the final digit indicating the power of ten); a last color band signifies the tolerance.

See Figure C.1. Resistors with numerical markings use the same system, but with the digits themselves printed along the body of the resistor (for leaded resistors), or on the top side of a surface-mount package; a final letter signifies the tolerance, as shown in the figure.

### C.4 Resistor types

The usual choices for general-purpose use are metal-film (axial-lead) or thick-film (surface-mount) parts. Thin-film surface-mount resistors offer improved characteristics (accuracy, stability, and ability to operate in cryogenic environments). For power applications you usually use wire-wound resistors, either in an air-cooled ceramic package or a conduction-cooled (“Dale-type”) metal package. High-value resistors ( $>10\text{ M}\Omega$ , say) are usually of metal-oxide construction (e.g., Ohmite “Mini-Mox” or “Super Mox,” or Vishay RNX-series). Film resistors are not tolerant of high peak power; for such applications use something like ceramic or carbon composition, or other styles specified for peak-power use. For the utmost in stability and low temperature coefficient (tempco), you can’t beat the excellent metal-foil types from Vishay. They exploit a clever design, in which the positive tempco of the resistive metal element (firmly attached to an insulating substrate) is cancelled by the negative strain-induced tempco caused by differential expansion of the substrate.<sup>3</sup> We’ve listed some comparative resistor properties in Table C.1; for much more detail see §1x.2.

General-purpose resistors are ridiculously inexpensive – thick-film surface-mount resistors cost a few cents apiece in small quantities, and just fractions of a cent apiece in full reel quantities (5000 pieces, for 0603 size). Distributors may be unwilling to sell fewer than 25 to 50 pieces of one value; thus an assortment box (e.g., from Yageo or Vishay/BC) may be a wise purchase. We particularly like the nice packaging and good pricing of the kits from SMT Zone ([www.smtzone.com](http://www.smtzone.com)).

### C.5 Confusion derby

Component markings should be clear and unambiguous. Sometimes it just ain’t so! See Figure 1.130 for some real head-scratchers, both resistive and otherwise.

<sup>3</sup> Check it out: Felix Zandman’s 1982 US patent #4,318,072, “Precision resistor with improved temperature characteristics.”

**Table C.1 Selected Resistor Types**

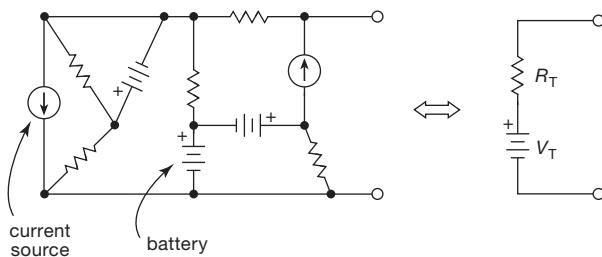
Parameter	Resistor Type					Units
	carbon comp axial (RC-07)	thick film SMT-0603 (Vishay CRCW)	thin film SMT-0603 (KOA Speer RN73)	metal film axial (RN-55D)	metal foil SMT (Vishay VSMP)	
Tolerances	5%, 10%	1%, 5%	0.05%-1%	0.1%-1%	0.01%-1%	$\Delta R/R$
Temp coef	~1000	100, 200	5, 10, 25, 50, 100	50, 100	0.05 (typ)	ppm/C
Load life	10%	2%	0.25%	0.5%	0.01%	$\Delta R/R$
Moisture	10%	2%	0.5%	0.5%	0.02%	$\Delta R/R$
Thermal cycle	2%	2%	0.25%	0.25%	0.01%	$\Delta R/R$
Low temp	3%	-	-	0.25%	0.01%	$\Delta R/R$
Overload	2%	0.5%	0.1%	0.25%	0.01%	$\Delta R/R$
Soldering	3%	0.5%	0.1%	0.25%	0.01%	$\Delta R/R$
Vibration	2%	-	-	0.25%	-	$\Delta R/R$
Voltage coef	-	-	-	5	0.1	ppm/V
Self-heating	-	-	-	-	5ppm	$\Delta R/R$
Price (approx) (for tol and TC)	\$0.35 (5%)	\$0.025 (1%, TC=200)	\$0.32 (0.1%, TC=25)	\$0.05 (1%, TC=100)	\$10 (0.01%, TC=0.05)	ea, qty 100

Properties of selected resistor types. The legendary axial-lead “carbon-composition” resistors have been superseded by inexpensive metal-film (or carbon-film) types, with greatly improved properties (except for peak-power endurance, see Chapter 1x). We like Vishay’s CMF-55 metal-film resistors (industrial version of MIL RN-55D). For most surface-mount applications the “thick-film” (a metal-ceramic composite) types are fine, though thin-film and metal-film resistors have somewhat better properties. The extraordinary Vishay “Z-foil” ultraprecision hermetically sealed resistor is listed to show the best that is currently available (but if you have to ask the price, you probably can’t afford it). It’s useful to note that a parameter like a voltage coefficient of 5 ppm/V corresponds to a change of 0.1% over a full 200 V operating range.

# THÉVENIN'S THEOREM

## APPENDIX D

In Chapter 1 we stated (but did not “prove”) Thévenin’s Theorem, namely that any two-terminal network whose internal circuitry consists solely of resistors, batteries, and current sources, interconnected in any manner whatsoever, is equivalent (and indistinguishable) from the two-terminal network consisting of a single battery  $V_{\text{TH}}$  in series with a single resistor  $R_{\text{TH}}$ ,<sup>1</sup> see Figure D.1. We did not prove it, because, in the spirit of this book, we don’t *prove* anything; we show you how to design circuits, instead. We make an exception here, because it’s nice to see *something* proved, right?



**Figure D.1.** Thévenin’s Theorem: a single resistor in series with a single battery can mimic any mess of a two-terminal network made from resistors, batteries, and current sources.

### D.1 The proof

For linear circuit elements (here resistors), the “nodal equations” (Kirchhoff’s voltage law, KVL, and Kirchhoff’s current law, KCL) are a set of linear equations. So we can find any circuit quantity (a voltage or a current), which depends on all the “independent sources” (batteries, current sources), by turning on each source in turn, and adding the partial contributions. (This is exactly analogous to using superposition to find, say, the electric field from a set of charges.) This technique is often useful in circuit analysis.

Here we wish to mimic the  $V$  versus  $I$  of the actual circuit with the (simpler) Thévenin equivalent of a single battery in series with a single resistor. Imagine we determine

<sup>1</sup> A related theorem is Norton’s, where the equivalent circuit consists of a resistor  $R_N$  in parallel with a current source  $I_N$ .

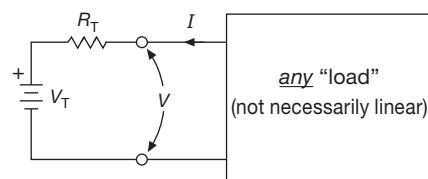
that  $V$  versus  $I$  function by applying an external current  $I_{\text{ext}}$  that flows through the two-terminal circuit, and observing the resultant  $V$  across those same two terminals.  $V$  depends on  $I_{\text{ext}}$  and on all the internal batteries ( $V_{\text{int}}$ ) and current sources ( $I_{\text{int}}$ ).

1. Set all  $V_{\text{int}} = 0$  and all  $I_{\text{int}} = 0$ ; that is, replace all internal batteries with short circuits and all current sources with open circuits. Now, with a given applied  $I_{\text{ext}}$ , observe  $V_1$ .
2. Define  $R_T = V_1/I_{\text{ext}}$ . (They must be proportional, by linearity.)
3. Now set  $I_{\text{ext}} = 0$ , and turn on the internal batteries and current sources. Observe  $V_2$ , which we will call  $V_T$ .
4. Finally, by superposition it must be the case that

$$V(\text{actual}) = V_1 + V_2 = I_{\text{ext}}R_T + V_T.$$

This is true for all  $I_{\text{ext}}$ , and is exactly what you get with the Thévenin equivalent circuit, when connected to *any* load (which need not be linear); see Figure D.2.

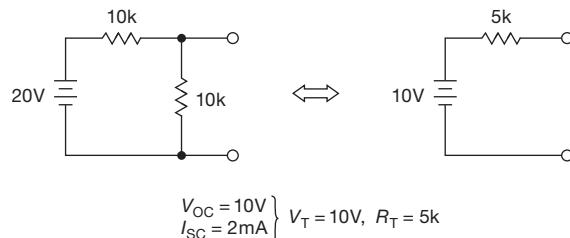
To summarize: (a) you determine  $R_T$  and  $V_T$  by first finding the open-circuit voltage, which equals  $V_T$ ; then (b) you find the short-circuit current,  $I_{\text{SC}}$ , which equals the ratio of  $V_T$  to  $R_T$ . In other words,  $V_T = V_{\text{OC}}$  and  $R_T = V_{\text{OC}}/I_{\text{SC}}$ . You do this by analysis, if you know the “black-box” circuit; or by measurement, if you don’t.



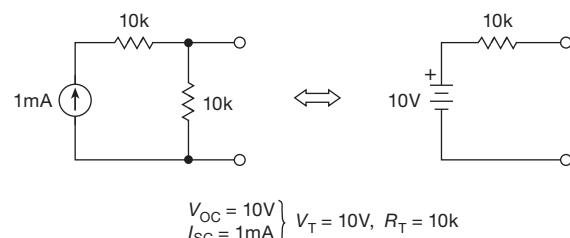
**Figure D.2.** The Thévenin equivalent circuit behaves exactly like the original network, regardless of the nature of the load.

#### D.1.1 Two examples – voltage dividers

Figures D.3 and D.4 show two simple examples, variations on the resistive divider. Interestingly, their Thévenin equivalent circuits are different, even though the resistor values and the open-circuit voltages are the same.



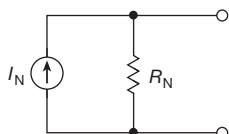
**Figure D.3.** Thévenin equivalent of a simple resistive divider. Note that  $R_T$  is the parallel resistance of the divider (as if the voltage source were replaced with a short circuit).



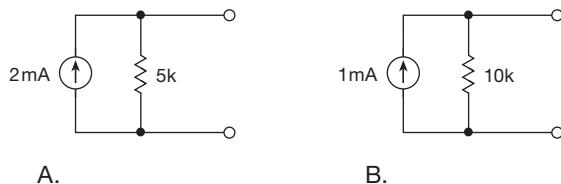
**Figure D.4.** Note that the Thévenin equivalent resistance is here *not* equal to the parallel resistance of the divider components. Instead it equals the value of the resistor across the output alone (as if the current source were replaced with an open circuit).

## D.2 Norton's theorem

You can replace a Thévenin circuit with a Norton circuit, which consists of a current source  $I_N$  in parallel with a resistor  $R_N$  (Figure D.5). It is easy to show that  $I_N = I_{SC}$  and  $R_N = R_T$  ( $= V_{OC}/I_{SC}$ ). So, for the two examples above, the Norton equivalents are as shown in Figure D.6.



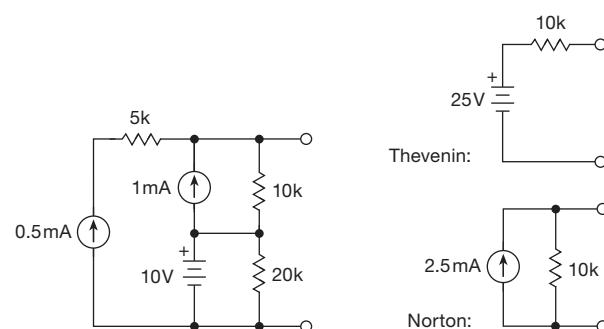
**Figure D.5.** Norton equivalent circuit: a current source in parallel with a resistor.



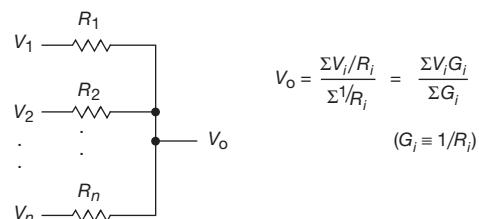
**Figure D.6.** Norton equivalents of the circuits of Figure D.3 (A) and Figure D.4 (B).

## D.3 Another example

Figure D.7 shows a complicated-looking circuit, for which it is pretty easy to see that  $V_{OC}=25\text{V}$  (the bottom of the 10k resistor sits at +10V, and 1.5 mA flows into the top) and that  $I_{SC}=2.5\text{mA}$  (10V across the 10k, plus the two current sources). From that you get the equivalent circuits shown.



**Figure D.7.** Thévenin and Norton equivalents of a complicated-looking circuit.



**Figure D.8.** Millman's theorem for parallel circuits.

## D.4 Millman's theorem

A related – and useful – tool is *Millman's Theorem* (also known as the parallel generator theorem), which is helpful when dealing with circuits with several parallel branches. It's shown in Figure D.8, where a set of input voltages  $V_i$  are combined via resistors  $R_i$ , producing an output voltage  $V_o$ . The latter is just  $V_o = (\sum V_i G_i) / \sum G_i$ , where the  $G_i$  are the conductances  $G_i \equiv 1 / R_i$ . The input voltages  $V_i$  can of course include ground, forming a voltage divider. Millman's theorem, which comes from the more general class of network theorems, can be generalized to include input currents  $I_k$ , whose sum is added to the numerator (but whose series resistances, if any, do not appear in the denominator).

# LC BUTTERWORTH FILTERS

## APPENDIX

## E

Active filters (see Chapter 6) are convenient at low frequencies, but they are impractical at radio frequencies because of the slew-rate and bandwidth requirements they impose on the operational amplifiers. At frequencies of 100 kHz and above (and often at lower frequencies), the best approach is to design a passive filter with inductors and capacitors. (At UHF and microwave frequencies these “lumped-component” filters are replaced by stripline and cavity filters.)

As with active filters, there are many methods and filter characteristics possible with *LC* filters. For example, you can design the classic Butterworth, Chebyshev, and Bessel filters, each in lowpass, bandpass, highpass, and band-reject varieties. It turns out that the Butterworth filter is particularly easy to design, and we can present in just a page or two all the essential design information for lowpass and highpass Butterworth *LC* filters, and even a few examples.

### E.1 Lowpass filter

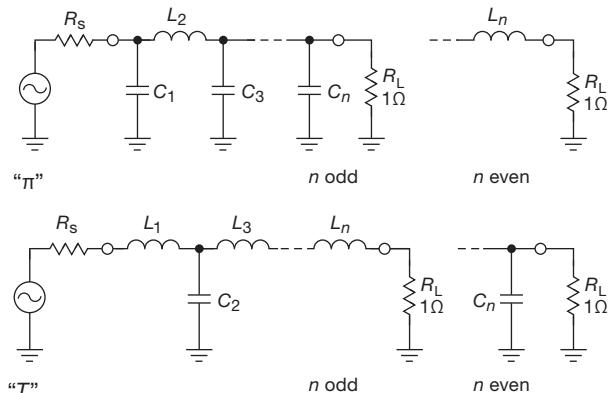
Table E.1 gives the values of normalized inductances and capacitances for low-pass filters of various orders, from which actual circuit values are obtained by the frequency and impedance scaling rules

$$L_n(\text{actual}) = \frac{R_L L_n(\text{table})}{\omega},$$

$$C_n(\text{actual}) = \frac{C_n(\text{table})}{\omega R_L},$$

where  $R_L$  is the load impedance and  $\omega$  is the angular frequency ( $\omega=2\pi f$ ).

Table E.1 gives normalized values for two-pole through eight-pole lowpass filters for the two most common cases, namely (a) equal source and load impedances and (b) either source or load impedance much larger than the other. To use the table, first decide how many poles you need, based on the Butterworth response (graphs are plotted in Figure 6.30). Then use the preceding equations to determine the filter configuration (*T* or  $\pi$ ; see Figure E.1) and component values. For equal source and load impedances, either



**Figure E.1.**  $\pi$  and *T* configurations. See Table E.1 and text.

configuration is OK; the  $\pi$  configuration may be preferable because it requires fewer inductors. For a load impedance much higher (lower) than the source impedance, use the *T* ( $\pi$ ) configuration.

### E.2 Highpass filter

To design a highpass filter, follow the same procedure to determine which filter configuration to use and how many poles are necessary. Then do the universal lowpass to highpass transformation shown in Figure E.2, which consists simply of replacing inductors by capacitors, and vice versa. The actual component values are determined from the normalized values in Table E.1 by the following frequency and impedance scaling rules:

$$L_n(\text{actual}) = \frac{R_L}{\omega C_n(\text{table})},$$

$$C_n(\text{actual}) = \frac{1}{R_L \omega L_n(\text{table})}.$$

### E.3 Filter examples

Here are a few examples showing how to design both lowpass and highpass filters.

**Example I.** Design a five-pole lowpass filter for source

**Table E.1 Butterworth Lowpass Filters<sup>a</sup>**

$\pi \longrightarrow R_S$	$C_1$	$L_2$	$C_3$	$L_4$	$C_5$	$L_6$	$C_7$	$L_8$
$T \longrightarrow 1/R_S$	$L_1$	$C_2$	$L_3$	$C_4$	$L_5$	$C_6$	$L_7$	$C_8$
$n = 2$	{ 1 ∞	1.4142 1.4142	1.4142 0.7071					
$n = 3$	{ 1 ∞	1.0000 1.5000	2.0000 1.3333	1.0000 0.5000				
$n = 4$	{ 1 ∞	0.7654 1.5307	1.8478 1.5772	1.8478 1.0824	0.7654 0.3827			
$n = 5$	{ 1 ∞	0.6180 1.5451	1.6180 1.6944	2.0000 1.3820	1.6180 0.8944	0.6180 0.3090		
$n = 6$	{ 1 ∞	0.5176 1.5529	1.4142 1.7593	1.9319 1.5529	1.9319 1.2016	1.4142 0.7579	0.5176 0.2588	
$n = 7$	{ 1 ∞	0.4450 1.5576	1.2470 1.7988	1.8019 1.6588	2.0000 1.3972	1.8019 1.0550	1.2470 0.6560	0.4450 0.2225
$n = 8$	{ 1 ∞	0.3902 1.5607	1.1111 1.8246	1.6629 1.7287	1.9616 1.5283	1.9616 1.2588	1.6629 0.9371	1.1111 0.5776
								0.3902 0.1951

Notes: (a) Values of  $L_n$ ,  $C_n$  for  $1\Omega$  load resistance, and cutoff frequency ( $-3\text{dB}$ ) of 1 rad/s.  
See text for scaling rules.

and load impedances of  $75\Omega$ , with a cutoff frequency ( $-3\text{ dB}$ ) of 1 MHz.

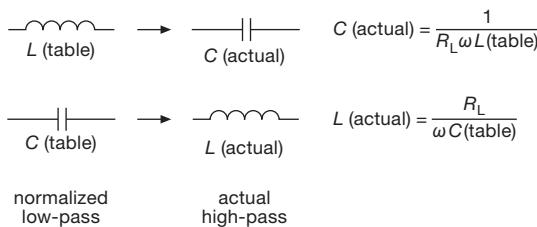
We use the  $\pi$  configuration to minimize the number of required inductors. The scaling rules give us

$$C_1 = C_5 = \frac{0.618}{2\pi \times 10^6 \times 75} = 1310\text{ pF},$$

$$L_2 = L_4 = \frac{75 \times 1.618}{2\pi \times 10^6} = 19.3\mu\text{H},$$

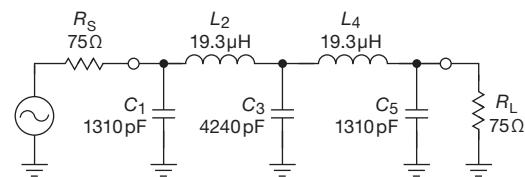
$$C_3 = \frac{2}{2\pi \times 10^6 \times 75} = 4240\text{ pF}.$$

The complete filter is shown in Figure E.3. Note that all filters with equal source and load impedances will be symmetrical.



**Figure E.2.** Lowpass to highpass transformation.

**Example II.** Design a three-pole lowpass filter for a



**Figure E.3.** Circuit for Example I. Five-pole 1 MHz lowpass with equal source and load impedances.

source impedance of  $50\Omega$  and a load impedance of  $10\text{k}$ , with a cutoff frequency of  $100\text{ kHz}$ .

We use the  $T$  configuration, because  $R_S \ll R_L$ . For  $R_L=10\text{k}$ , the scaling rules give

$$L_1 = \frac{10^4 \times 1.5}{2\pi \times 10^5} = 23.9\text{ mH},$$

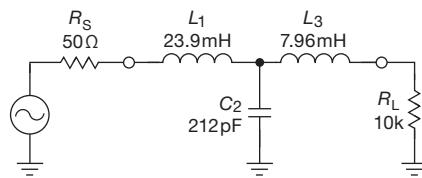
$$C_2 = \frac{1.3333}{2\pi \times 10^5 \times 10^4} = 212\text{ pF},$$

$$L_3 = \frac{10^4 \times 0.5}{2\pi \times 10^5} = 7.96\text{ mH}.$$

The complete filter is shown in Figure E.4.

**Example III.** Design a four-pole lowpass filter for a zero-impedance source (voltage source) and a  $75\Omega$  load, with a cutoff frequency of  $10\text{MHz}$ .

We use the  $T$  configuration, as in the previous example,



**Figure E.4.** Circuit for Example II. Three-pole 100 kHz lowpass with 50  $\Omega$  source and 10k load.

because  $R_S \ll R_L$ . The scaling rules give

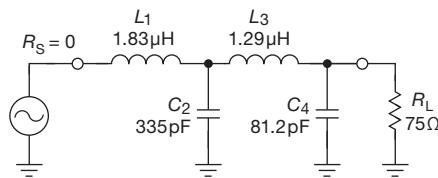
$$L_1 = \frac{75 \times 1.5307}{2\pi \times 10^7} = 1.83 \mu\text{H},$$

$$C_2 = \frac{1.5772}{2\pi \times 10^7 \times 75} = 335 \text{ pF},$$

$$L_3 = \frac{75 \times 1.0824}{2\pi \times 10^7} = 1.29 \mu\text{H},$$

$$C_4 = \frac{0.3827}{2\pi \times 10^7 \times 75} = 81.2 \text{ pF}.$$

The complete filter is shown in Figure E.5.



**Figure E.5.** Circuit for Example III. Four-pole 10 MHz lowpass with voltage source and 75  $\Omega$  load.

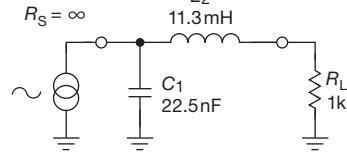
**Example IV.** Design a two-pole lowpass filter for current-source drive and 1k load impedance, with a cutoff frequency of 10 kHz.

We use the  $\pi$  configuration because  $R_S \gg R_L$ . The scaling rules give

$$C_1 = \frac{1.4142}{2\pi \times 10^4 \times 10^3} = 0.0225 \mu\text{F},$$

$$L_2 = \frac{10^3 \times 0.7071}{2\pi \times 10^4} = 11.3 \text{ mH}.$$

The complete filter is shown in Figure E.6.



**Figure E.6.** Circuit for Example IV. Two-pole 10 kHz lowpass with current source drive and 1k load.

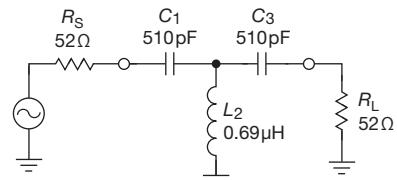
**Example V.** Design a three-pole highpass filter for 52  $\Omega$  source and load impedances, with a cutoff frequency of 6 MHz.

We begin with the  $T$  configuration, then transform inductors to capacitors, and vice versa, giving

$$C_1 = C_3 = \frac{1}{52 \times 2\pi \times 6 \times 10^6 \times 1.0} = 510 \text{ pF},$$

$$L_2 = \frac{52}{2\pi \times 6 \times 10^6 \times 2.0} = 0.690 \mu\text{H}.$$

The complete filter is shown in Figure E.7.



**Figure E.7.** Circuit for Example V. Three-pole 6 MHz highpass with equal source and load impedances.

We would like to emphasize that the field of passive filter design is rich and varied and that this simple table of Butterworth filters doesn't even begin to scratch the surface.

# LOAD LINES

## APPENDIX F

The graphic method of “load lines” usually makes an early appearance in electronics textbooks. We have avoided it because, well, it just isn’t useful in transistor design the way it was in vacuum-tube circuit design. However, it is of use in dealing with some nonlinear devices (tunnel diodes, for example), and in any case it is a useful conceptual tool.

### F.1 An example

Let’s start with an example. Suppose you want to know the voltage across the diode in Figure F.1. Assume that you know the voltage-versus-current ( $V$ - $I$ ) curve of the particular diode (of course, it would have a manufacturing “spread,” as well as a dependence on ambient temperature); it might look something like the curve drawn. How would you figure out the quiescent<sup>1</sup> point?

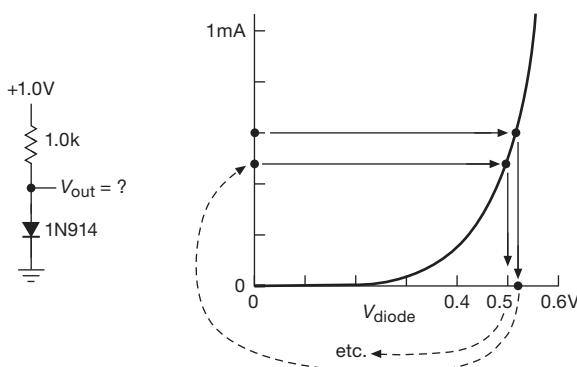


Figure F.1. Finding the operating point by iteration.

One method might be to guess a rough value of current, say 0.6 mA, then use the curve to get the drop across the resistor, from which you get a new guess for the current (in this case, 0.48 mA). This iterative method is suggested in

Figure F.1. After a few iterations, this method will get you an answer, but it leaves a lot to be desired.

The method of load lines gets you the answer to this sort of problem immediately. Imagine *any* device connected in place of the diode; the 1.0k resistor is still the load. Now plot, on a  $V$ - $I$  graph, the curve of resistor current versus device voltage. This turns out to be easy: at zero volts the current is just  $V_+/R$  (full drop across the resistor); at  $V_+$  volts the current is zero; points in between fall on a straight line between the two. Now, on the same graph, plot the  $V$ - $I$  curve of the device. The operating point lies on both curves, i.e., at the intersection, as shown in Figure F.2.

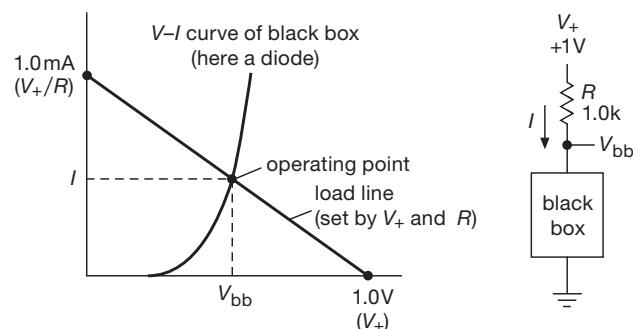


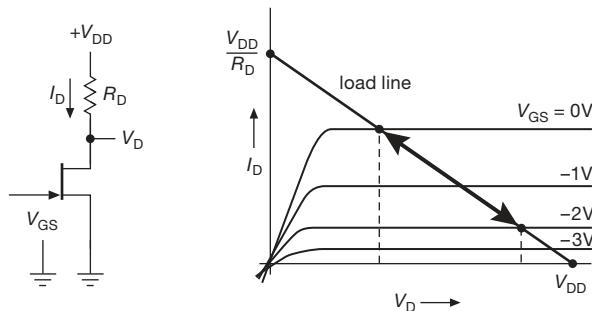
Figure F.2. A “load line” lets you find the operating point directly.

### F.2 Three-terminal devices

Load lines can be used with a three-terminal device (tube or transistor, for example) by plotting a family of curves for the device. Figure F.3 shows what such a thing would look like for a depletion-mode field-effect transistor (FET), with the curve family parameterized by the gate-source voltage. You can read off the output for a given input by sliding along the load line between appropriate curves corresponding to the input you’ve got, then projecting onto the voltage axis. In this example we’ve done this, showing the drain

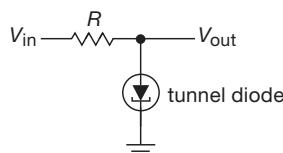
<sup>1</sup> The quiescent point, also known as the *operating point*, describes the various dc voltages and currents in a circuit with no ac signals applied.

voltage (output) for a gate swing (input) between ground and  $-2\text{ V}$ .



**Figure F.3.** Load-line solution for a three-terminal device.

As nice as this method seems, it has quite limited use for transistor or FET design, for a couple of reasons. For one thing, the curves published for semiconductor devices are “typical,” with manufacturing spread that can be as large as a factor of five. Imagine what would happen to those nice load-line solutions if all the curves shrank to one-fourth their height! Another reason is that for an inherently logarithmic device like a diode junction, a linear load-line graph can be used to give accurate results over only a narrow region. Finally, the nongraphic methods we’ve used in this book are entirely adequate for handling solid-state design. In particular, these methods emphasize the parameters you can count on ( $r_e$ ,  $I_C$  versus  $V_{BE}$  and  $T$ , etc.), rather than the ones that are highly variable ( $\beta$ ,  $V_{th}$ , etc.). If anything, the use of load lines on published curves for transistors only gives you a false sense of security, since the device spread isn’t also shown.

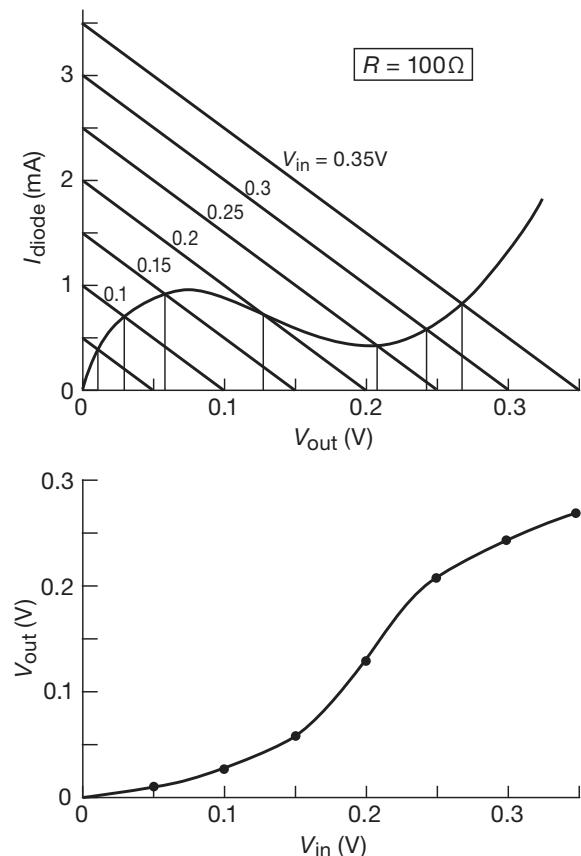


**Figure F.4.** The tunnel diode: a two-terminal nonlinear device with a region of negative resistance (see Figure F.5).

### F.3 Nonlinear devices

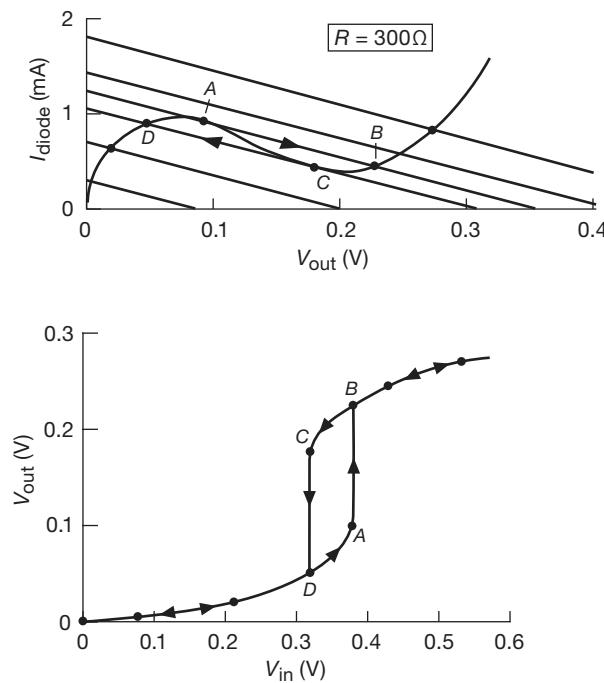
Load lines turn out to be useful in understanding the circuit behavior of highly nonlinear devices. The example of tunnel diodes illustrates a couple of interesting points. Let’s analyze the circuit in Figure F.4. Note that in this case,  $V_{in}$

takes the place of the supply voltage in the previous examples. So a signal swing will generate a family of parallel load lines intersecting with a single device  $V-I$  curve (Figure F.5). The values shown are for a  $100\Omega$  load resistor. As can be seen, the output varies most rapidly as the input swing takes the load line across the negative-resistance portion of the tunnel-diode curve. By reading off values of  $V_{out}$  (projection on the  $x$  axis) for various values of  $V_{in}$  (individual load lines), you get the “transfer” characteristics shown. This particular circuit has some voltage gain for input voltages near  $0.2\text{ V}$ .



**Figure F.5.** Load lines and transfer characteristic for the tunnel-diode circuit.

An interesting thing happens if the load lines become flatter than the middle section of the diode curve. That happens when the load resistance exceeds the magnitude of the diode’s negative resistance. It is then possible to have *two* intersection points, as in Figure F.6. A rising input signal carries the load lines up until the intersection point has nowhere to go and has to jump across to a higher  $V_{out}$  value. On returning, the load lines similarly carry the



**Figure F.6.** Having  $|R_{\text{load}}| > |R_{\text{neg}}|$  produces hysteretic switching behavior in the tunnel-diode circuit.

intersection point down until it must again jump back. The overall transfer characteristic has *hysteresis*, as shown. Tunnel diodes have been used in this manner as fast-switching devices (triggers).

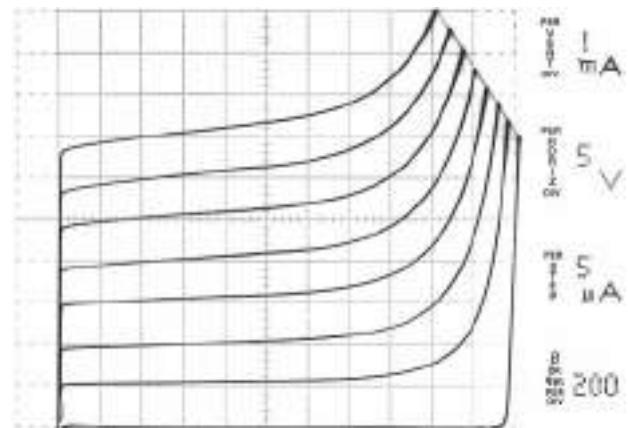
# THE CURVE TRACER

## APPENDIX G

A handy instrument for exploring transistor behavior of both BJTs (Chapter 2) and MOSFETs (Chapter 3) is the *curve tracer*. Most simply, it plots collector current versus collector voltage for a family of equally spaced base currents (or, if you want to be an Ebers–Mollian, base *voltages*), and with a selectable current-limiting collector resistor.<sup>1</sup> Figure G.1 shows what you get from a random 2N3904, driven with seven successive base current steps of  $5\ \mu\text{A}$  each while sweeping the collector voltage from 0 to 50 V. You can see clearly the rise of beta with collector voltage, and the onset of breakdown somewhat below 50 V (maximum  $V_{\text{CEO}}$  is specified as 40 V). This particular curve tracer obligingly displays the scale factors, including “ $\beta$  per div,” which is about 200 for this specimen (the datasheet specifies  $100 \leq \beta \leq 300$  at  $I_C=10\ \text{mA}$ ). A curve tracer makes it easy to select closely matched pairs.

Sadly, the traditional curve tracer has disappeared from the product lines of most T&M (test and measurement) manufacturers, including the venerable Tektronix. You can still find them used, for example on eBay, for a thousand dollars or so. Agilent offers some pretty fancy contemporary instruments that will do the job, though it’s best to be sitting down when you ask the prices; they go by names like Semiconductor Parameter Analyzer (model 4155C), or Power Device Analyzer/Curve Tracer (model B1505A).

<sup>1</sup> You can run it as common base, if you like; and it has many amusing knobs to play with.



**Figure G.1.** Tektronix 576 Curve Tracer display of a 2N3904.

A less expensive alternative is to use a “source-measure unit” (SMU), a delightful instrument that lets you source voltages and currents to selected terminals of a device (or subcircuit), simultaneously measuring and logging other voltages or currents. You can program the excitation as dc, or ramps, steps, or pulses, and you can display the logged results via software running on an attached laptop computer; you can also save the logged data as a spreadsheet, to be manipulated to your heart’s content. Take a look, for example, at Figure 8.39, or the figures in the “Power Transistors for Linear Amplifiers” section of Chapter 3x, all of which plot data we collected with a SMU.

# TRANSMISSION LINES AND IMPEDANCE MATCHING

## APPENDIX

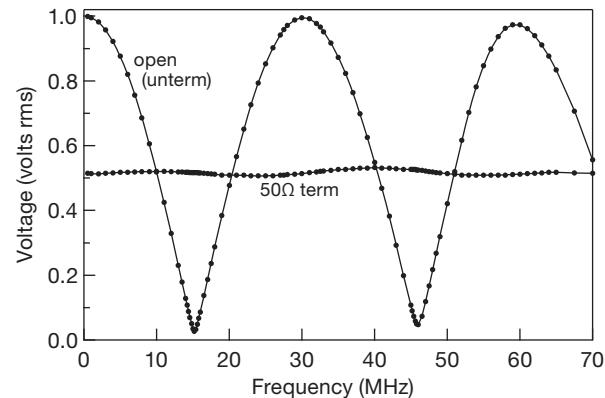
## H

### H.1 Some properties of transmission lines

In §12.9 we introduced transmission lines, which most commonly take the form of *coaxial cable* (“coax”), for example the ubiquitous “BNC cables” (RG-58 cables with male BNC connectors at each end) that are used to run all manner of signals between instruments. As we remarked there, for low-frequency applications it is common (and correct) to think of such a cable simply as a well-shielded wire with  $\sim 30 \text{ pF/ft}$  of capacitance. However, at high frequencies (say those for which the cable is at least 1/20 of the wavelength) the behavior is fundamentally different: as a bizarre example, an open-ended cable ironically looks like a *short circuit* at a frequency for which the cable’s electrical length is  $\lambda/4$ . For a 5-foot length of coax, that happens at about 32 MHz. An important consequence is that you can’t just hook such a BNC cable from a signal generator to some high-impedance circuit under test and assume that it will provide a nice signal source at the circuit’s input; instead you will see huge dips and bumps as you tune the frequency, because the generator sees a load impedance that varies from a short circuit (at odd multiples of 32 MHz) to an open circuit (at even multiples of 32 MHz). Perhaps surprisingly, if you were instead to connect a resistor of exactly  $50 \Omega$  across the circuit end of the cable, you would find that it now delivered a constant signal amplitude (equal to half the signal generator’s open-circuit output amplitude) as you varied the frequency. This nonintuitive behavior is nicely illustrated in the measured data shown in Figure H.1. And even more nonintuitively, at the driving end of such a “terminated” cable the capacitance disappears entirely – you see instead a pure resistive load of  $50 \Omega$ !

#### H.1.1 Characteristic impedance

This simple example illustrates the importance of *termination*: coaxial cable is a form of *transmission line*, with a *characteristic impedance*  $Z_0$  (which is always real: a resis-



**Figure H.1.** Measured amplitude at the output connector of a sinewave oscillator of 1 V amplitude (open circuit), under two conditions: driving 10 feet of RG-58 ( $50 \Omega$ ) coaxial cable, open at the far end; and driving the same cable with a  $50 \Omega$  resistor connected across the far end.

tance) that depends on only its physical construction:

$$Z_0 = \sqrt{L/C} = \frac{138}{\sqrt{\epsilon}} \log_{10} \frac{b}{a} \text{ ohms},$$

where  $L$  and  $C$  are the inductance and capacitance per unit length, which as indicated depend on only the outer diameter,  $a$ , of the inner conductor, the inner diameter,  $b$ , of the outer conductor, and the dielectric constant,  $\epsilon$  (relative to free space), of the insulating material that separates them. For a wave propagating along a transmission line,  $Z_0$  is the ratio of signal voltage to signal current. The most popular coax line for general purposes is RG-58, with an impedance of  $50 \Omega$  (its dimensions are  $a=0.81 \text{ mm}$ ,  $b=2.95 \text{ mm}$ , and  $\epsilon=2.3$ , for which the above equation gives  $Z_0=51 \Omega$ ). This impedance has become the standard for radiofrequency use, except for video applications where the standard is  $Z_0=75 \Omega$ ; the corresponding popular cable type is called RG-59. In pulse electronics you sometimes see  $93 \Omega$  cable (RG-62).

The signal propagates along the cable at a speed

$$v_{\text{wave}} = \frac{c}{\sqrt{\epsilon}} = \frac{1}{\sqrt{LC}},$$

which is a fraction  $1/\sqrt{\epsilon}$  times  $c$  (where  $c$  is the speed of light in vacuum,  $3 \times 10^8$  m/s). The factor  $1/\sqrt{\epsilon}$  is called the *velocity factor*, and ranges from 0.66 (solid polyethylene) to 0.80 (polyethylene foam) for available flexible coaxial cables. In the absence of a dielectric the velocity factor is 1.0, i.e., waves travel at the speed of light in an air-spaced coaxial line. The “electrical length” seen by a propagating signal in a cable of physical length  $L$  is larger by the factor  $\sqrt{\epsilon}$ , i.e.,  $L_{\text{elec}} = L_{\text{physical}}\sqrt{\epsilon}$ .

Note that the inductance and capacitance of the cable cannot take on any old values, because they are constrained such that their product  $LC$  is related to the speed of light. From this it is easy to show that if you know the characteristic impedance and the velocity factor then you can find the capacitance per unit length (or vice versa) by

$$C = \frac{\sqrt{\epsilon}}{cZ_0} \text{ Farads/meter.}$$

For example, RG-8 has a specified impedance of  $52\Omega$ , and a velocity factor of 0.66; the above equation then gives  $C=97.1\text{ pF/m}$ , or  $29.6\text{ pF/ft}$ , in good agreement with the specified value of  $29.5\text{ pF/ft}$ .

### A. Twisted pair and PCB traces

Transmission lines are not *required* to be of coaxial geometry. An extremely popular form of contemporary transmission line is the *twisted pair*, which is just what it sounds like: a pair of insulated wires, gently twisted, and enclosed in an overall insulating jacket (often without any overall shield conductor). These may well be the dominant species of transmission line in our time, because they are the basic stuff of local area networks (LANs). You usually see four twisted pairs bundled into a single unshielded jacket; this is called “UTP” (unshielded twisted pair), and is the most common form of LAN cable. It is available also with a shield (shielded twisted pair, “STP”). Contemporary UTP and STP cables are  $100\Omega$  nominal impedance, and are characterized (in terms of impedance and attenuation) for operation up to 10 Mbps (megabits per second) or 100 Mbps; these are called Category 3 and Category 5, respectively, and appear to differ primarily in the pitch of the twist. Those in the know refer to these casually as cat-3 and cat-5.<sup>1</sup> Ethernet LANs using these data rates

are called “10baseT” and “100baseTX,” the “T” standing for “twisted”; the corresponding designation for “thinnet” coaxial Ethernet is “10base2.”

In high-speed electronics on a printed circuit board it’s often necessary to treat connection traces as transmission lines. See the discussion in §1x.1, where we described the *microstrip* transmission line, which consists of a thin conducting strip on an outer printed circuit layer, with an underlying ground-plane layer. A popular variant adds a pair of ground-trace shepherds on either side – that’s called a *grounded coplanar waveguide* (GCPW). There’s also the completely enclosed *stripline* geometry, where the trace(s) are sandwiched between ground-plane layers.

### H.1.2 Termination: pulses

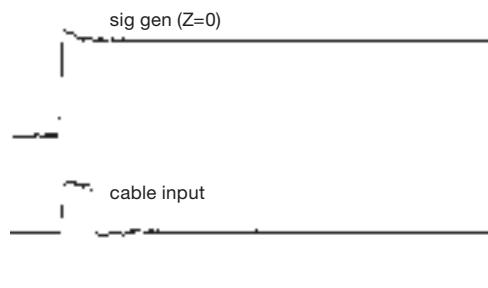
A transmission line at low frequencies (wavelength much longer than the cable length) looks simply like a capacitance, typically of order  $30\text{ pF/ft}$ . However, at high frequencies, or, equivalently, when dealing with signals with fast rise times, the behavior is different. In order to understand the curious behavior illustrated in Figure H.1, it’s helpful to look first at what happens when a simple *pulse* is applied to a length of transmission line. Suppose we connect a fast pulse generator with  $50\Omega$  output impedance (the standard output impedance of signal generators, function generators, and pulse generators) to a length of  $50\Omega$  coax, shorted at the far end. The pulse at first disappears into an impedance  $Z_0$  (thus the signal amplitude is half that of the unloaded generator), but after a round-trip travel time a reflected pulse of opposite polarity returns (Figure H.2). If the signal applied is instead a fast step, the effect of the reflection is to convert the step into a pulse (Figure H.3). An open-ended line produces a reflection of the *same* polarity, with the effects shown in Figure H.4. For arbitrary load resistance  $R$  the ratio of reflected to incident wave amplitude (the reflection coefficient) is given by

$$\rho \equiv A_r/A_i = (R - Z_0)/(R + Z_0).$$

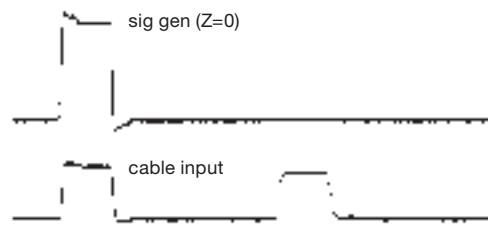
Note that a termination resistance of  $R=Z_0$  produces no reflection. A signal applied to such a terminated line is absorbed by the terminating resistor (as heat) and disappears forever. The signal source sees a loading resistance equal to  $Z_0$ . (It is for this reason that we did not have to worry earlier about the reflected pulses reflecting again from the pulse

<sup>1</sup> Higher performance standards include Category 5e (“e” for *enhanced*), and Category 6, propelled by the development of Gigabit Ethernet –

literally 1 gigabit/sec over unshielded twisted pair – also known as 1000baseT. To achieve this data rate, all four pairs are used simultaneously with 5-level amplitude encoding.



**Figure H.3.** 'Scope trace of a step waveform applied to an 8-foot length of RG-58A/U (solid polyethylene dielectric, velocity factor of 66%), shorted at the end. Vertical; 1 V/div; horizontal; 40 ns/div.

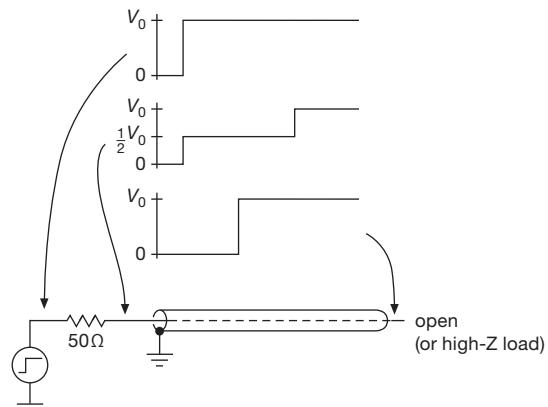


**Figure H.4.** 'Scope trace showing reflection from open-ended coax line. Same parameters as those in Figure H.2B.

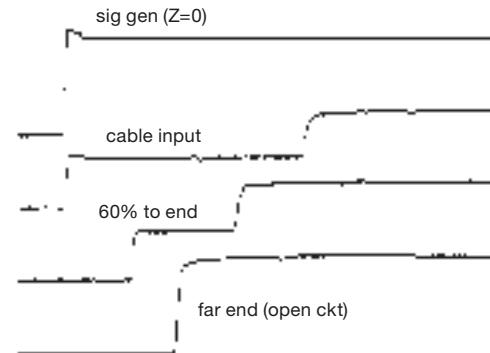
generator – its  $50\Omega$  source impedance swallows any signals returning from an improperly terminated cable, which is the reason most signal sources are standardized at  $50\Omega$  impedance.)

### A. Series termination

This last point – that returning (reflected) signals are completely absorbed if the signal source's impedance matches the line – leads to a nice technique called *series termination* (or *back termination*), frequently used for high-speed logic signals (and in other situations where the load has a high input impedance). Look at Figure H.5: a signal source in series with a resistor (equal to the line's impedance) drives a transmission line whose far end is unterminated (i.e., open). Now imagine a step input of amplitude  $V_{\text{sig}}$  at the signal source; it propagates down the line at half-amplitude, then reflects back from the far end at the full  $V_{\text{sig}}$  amplitude. Although any point along the line sees a two-step waveform, the surprising fact is that the waveform seen at the far end makes a single step from zero to the full



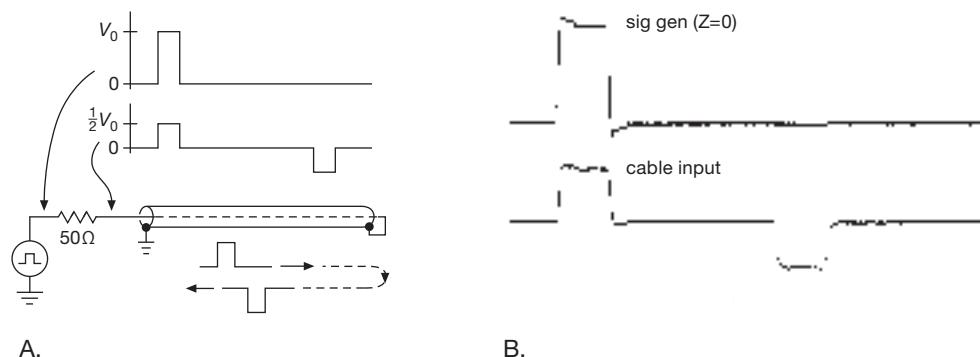
**Figure H.5.** Series termination of an open-ended line: the line is fed from a signal source of matched impedance; the half-sized step at the input propagates to the far end, from which it reflects in-phase to produce a returning step equal in amplitude to the *unloaded* amplitude of the generator. A high-impedance load at the far end sees only a single full-sized step.



**Figure H.6.** 'Scope trace showing waveforms at cable input, midpoint, and far end, for series-terminated step input. The generator's zero-impedance signal is also shown; it was set to produce a 2 V step into an open circuit. The cable is 60 ft of RG-58/U (velocity factor of 66%), tapped at 36 ft with a high-impedance voltage probe. Vertical; 1 V/div; horizontal; 40 ns/div.

$V_{\text{sig}}$ ; at that place the half-sized incident wave arrives at the same time that the half-sized reflected wave departs. This interesting behavior is demonstrated in the 'scope traces in Figure H.6.

You can use this technique for sending CMOS logic signals through a length of coax: three paralleled 74HC buffers (for low source impedance, roughly  $15\Omega$ ) in series with a  $33\Omega$  resistor nicely drives lengths of RG-174 or RG-316 (thin  $50\Omega$  coax line), connected to the receiving gate at the far end *without termination*. The receiving gate sees



**Figure H.2.** A. A pulse driving a length of shorted transmission line reflects off the short and returns as a pulse of opposite polarity. B. Scope trace taken with 70 ft of RG-8/U with foam dielectric (velocity factor of 78%), shorted at end. Vertical; 1 V/div; horizontal; 40 ns/div. For this and following figures we used a high-impedance 'scope probe to avoid introducing additional transmission-line effects.

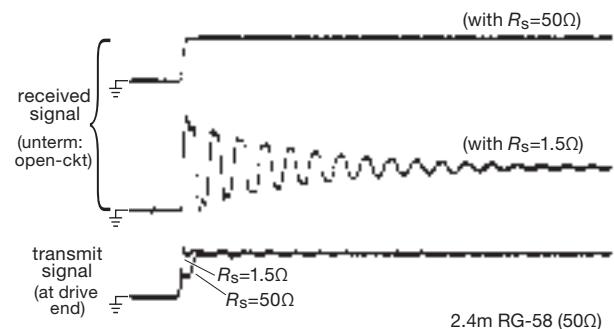
full-swing logic signals. This technique is often preferable to the matched-load alternative – directly driving a line that is terminated in  $50\Omega$  – because with series termination the driver sees a load resistance twice as high ( $100\Omega$  in this case), and that only for the round-trip duration of the signal (after which the load becomes an open-circuit).<sup>2</sup> For very fast logic signals (e.g., ECL 100K, or contemporary high-speed CMOS processors, memory, and peripherals), it may be necessary to treat a circuit trace of just a few inches as a transmission line. Typically printed circuit board (PCB) trace impedances are in the range of  $50\Omega$  to  $100\Omega$ , but can be tailored to a specific impedance by proper choice of trace width and spacing above the ground plane; this specialty art is known as *microstrip* technique,<sup>3</sup> useful both for fast digital signals and for signals at frequencies above about 100 MHz (UHF and microwave).

This is all very nice in theory – but in practice you have to contend with sources of fast digital signals that are not matched to the line impedance. This happens often on digital PCBs where the digital output ports of speedy microprocessors and FPGAs are poorly matched to the PCB trace impedances. For example, a line driven by a signal of source impedance  $Z_0/2$  produces lots of ringing at both ends of an unterminated line, including 33% overshoot at the far end; such ringing can produce false clocking.

To illustrate what this looks like, we drove a 2.4 m length of unterminated RG-58/U ( $Z_0=50\Omega$ ) coax with a step input, probing both the input and output signals.

<sup>2</sup> See §12.10, where several methods of driving cables with logic levels are described and illustrated.

<sup>3</sup> If you sandwich the signal-carrying conductors between a pair of ground planes, you've got a *stripline*; see §1x.1.



**Figure H.7.** Signals seen at the far end of an unterminated 8ft length of  $50\Omega$  line when driven with a unit step from a  $50\Omega$  source (top trace) and from a low-impedance source (middle trace). The corresponding signals at the driving end are shown at bottom. Horizontal; 100 ns/div.

We did this under two conditions: (a) when driven with a  $50\Omega$  series termination at the input (i.e., a “back-terminated” line); and (b) when driven with a low-impedance ( $R_s=1.5\Omega$ ) voltage step.

Figure H.7 shows the results, where the drive signal’s open-circuit amplitude (call it  $V_{OC}$ ) is displayed as one vertical division. The matched series termination generates a clean received step to  $V_{OC}$  (and an input waveform with initial step first to  $V_{OC}/2$ , then to  $V_{OC}$  after a round-trip delay, just as seen in expanded scale in Figure H.6). By contrast, the low-Z drive imposes a full  $V_{OC}$  step at the input, which is first seen at the far end at  $2V_{OC}$  (because the non-inverted reflected wave doubles the amplitude of the arriving wave), subsequently brought nearly back to zero (one round-trip time later) by the inverted wave reflected from

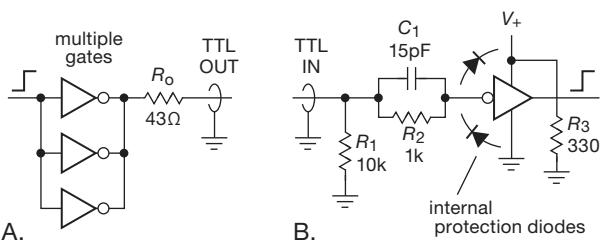
the low-impedance driver. This alternating pattern continues, damped both by cable loss and resistive loss in the  $1.5\ \Omega$  driver. This signal is a mess! That's why it's worth some effort to match driver impedances to the characteristic impedance of lines whose length (given the signal rise times) qualifies them as transmission lines.<sup>4</sup>

### B. A robust logic link

The sort of overshoot behavior seen in Figure H.7 can destroy logic circuits at the receiving end. You see this vulnerability particularly in the laboratory, where a low-impedance logic signal (or pulse-generator output) travels through a length of coax to a logic input on some instrument. The latter is often unterminated, to keep the input impedance relatively high (to prevent heavy loading and attenuation for a signal source unable to drive  $50\ \Omega$ ).

If you want to make your own designs bulletproof against this hazard, you can add a few components as shown in Figure H.8. At the receiving end, series resistor  $R_2$  limits the current, safely clamped by the logic gate's internal protection diodes; the "speed-up capacitor"  $C_1$  prevents loss of speed ( $1\ k\Omega$  into a typical input and wiring capacitance of  $10\ pF$  is an  $RC$  time constant of  $10\ ns$ , a near-eternity in the frenetic world of digital logic). It's always a good idea to include an input pull-down resistor ( $R_1$ ) to ensure a defined logic level when the input is disconnected. In an abundance of caution we've added resistor  $R_3$ , whose job is to prevent a large positive overdrive from forcing the  $V_+$  rail to a positive voltage that can damage other ICs; this could normally be omitted, but it would be a good idea in an instrument with an inviting BNC connector on the front and that contains only a few ICs powered from a small regulator (like a 78L05 – see §9.3.2, Figure 9.6, and Table 9.1) whose dc output can be easily overdriven.

At the driver end the parallel connection of several logic gates generates a drive impedance down in the neighborhood of  $5\text{--}10\ \Omega$ , which the added series resistor  $R_o$  brings up to a driving resistance close to the cable's  $50\ \Omega$  characteristic impedance. That's what you want, of course, and that alone is enough to give you peace of mind as a respectable series-terminated driver. But it never hurts to protect the receiver end, as just discussed: you never know when someone will drive it with a pulse generator, inadvertently set to deliver *negative* pulses, or  $20\ V$  positive pulses (as happened in our laboratory recently).



**Figure H.8.** A. Simple logic-level driver for series-terminated cable; B. Logic-level receiver protected against overdrive. Connect these together to form a complete signal path.

There's further discussion of cable driving and logic interfacing in Chapter 12, beginning at §12.10.

### H.1.3 Termination: sinusoidal signals

We have been talking about signals propagating along transmission lines, for clarity using the particular case of pulses or voltage steps. Of course, a *sinewave* applied to a length of cable also produces reflections, unless of course the cable is properly terminated. The effect is to alter the input current, for an applied input voltage, in a way that depends on the (mismatched) load impedance  $Z_L$ , and also on the ratio of the signal's wavelength in the cable  $\lambda$  to the physical length of the cable  $l$ . The final effect is to produce an input impedance (complex in general) given by

$$Z_{in} = Z_0 \frac{Z_L + jZ_0 \tan(2\pi l/\lambda)}{Z_0 + jZ_L \tan(2\pi l/\lambda)}.$$

From this one can see that:

- (a) a matched termination ( $Z_L = Z_0$  = [usually]  $50\ \Omega$ ) results in an input impedance equal to the characteristic impedance of the cable, independent of length or frequency;
- (b) a quarter-wave line inverts the load impedance, i.e.,  $Z_{in} = Z_0^2/Z_L$ ;
- (c) a half-wave line preserves the load impedance, i.e.,  $Z_{in} = Z_L$ ;
- (d) a short length of open-circuited line  $l \ll \lambda$  looks capacitive, viz.,  $Z_{in} \approx -j/\omega C'$ , where  $C'$  (the effective capacitance) is the constant  $l/cZ_0$ ;
- (e) a short length of short-circuited line  $l \ll \lambda$  looks inductive, viz.,  $Z_{in} \approx j\omega L'$ , where  $L'$  (the effective inductance) is the constant  $Z_0l/c$ .

The impedance-changing properties of transmission lines can be used to match impedances, though any such scheme will be frequency dependent; when you hear words like "stubs," you're dealing with transmission-line impedance matching. Virtuosos in this area make heavy use of network analyzers, and they will try to dazzle you with their

<sup>4</sup> An analogous effect in long-distance power transmission lines is known as the Ferranti effect; it is said that overvoltages caused by the Ferranti effect, if not properly compensated, can cause damage to power-line switch gear.

handsome “Smith Charts” (which are well beyond the humble scope of this book<sup>5</sup>).

When you have sinusoidal signals – with reflections – on a transmission line, you generate *standing waves*. That is, you can picture the net result of waves propagating in both directions (at the same frequency) as the sum of a nonpropagating (hence “standing”) wave and some additional propagating wave. For example, an open-ended line produces a reflected wave of full amplitude; the result is a pure standing wave of the same frequency and twice the amplitude, with a maximum amplitude of oscillation at the open end (and repeating every half-wavelength), and complete nulls (“nodes” – places with no voltage) midway between. For a shorted-end line a similar thing happens, but the reflected wave is of opposite amplitude, producing a null at the far end (and repeating every half-wavelength), with maxima in between. (You get the same pattern if you tie a length of clothesline to a fence, then wiggle the end up and down at the right rate.) With a smaller termination mismatch you don’t get complete cancellation anywhere.

Standing waves are not necessarily bad (though they are never<sup>6</sup> *good!*). But they do increase both the peak voltages and the resistive losses (see next section), relative to a matched line, for the same power transmitted. They are ordinarily seen as the symptom of a mismatched line. So in communications systems people try to minimize the *standing wave ratio* (abbreviated SWR, or sometimes VSWR – for *voltage* standing wave ratio – pronounced “VIZ-wahr”), which is defined as the ratio of the maximum amplitude to minimum amplitude:

$$\text{VSWR} = \frac{V_{\max}}{V_{\min}} = \frac{A_f + A_r}{A_f - A_r},$$

<sup>5</sup> Smith charts are treated in the excellent reference *Fields and Waves in Communication Electronics* by Ramo, Whinnery, and Van Duzer (Wiley, 1994), as well as in the insightful and refreshing *Radio-Frequency Electronics* by Hagen (Cambridge University Press, 2009).

<sup>6</sup> Well, *hardly ever!* For operation over a narrow frequency range you sometimes exploit the impedance-changing properties of mismatched lines, which necessarily have standing waves. Examples are (a) the use of open or shorted lengths of line as high-*Q* capacitors or inductors, (b) a shorted half-wave or an open quarter-wave line used as an RF bypass capacitor, (c) an open half-wave or a shorted quarter-wave line used as an RF choke, (d) matching two different impedances (cables, signal sources, or loads) by interposing a quarter-wave section of transmission line whose characteristic impedance is the geometric mean of the two impedances being matched (this is analogous to a quarter-wave anti-reflection coating in optics), (e) the use of a slotted line and high-impedance probe for a direct measurement of wavelength, and (f) the use of transmission lines to make “ring hybrids” and “rat-race hybrids.” We thank Jon Hagen and Darren Leigh for these applications of “good standing waves.”

where  $V$  is the ac (signal) voltage, measured at points along the line; and  $A_f$  and  $A_r$  are the amplitudes of the forward and reflected waves, respectively. Voltage measurements along a cable with no standing waves will give a constant amplitude (hence VSWR=1.0).

The VSWR is a real number, between 1.0 (perfect match, no reflected wave) and  $\infty$  (“perfect mismatch,” reflected wave amplitude equal to incident wave amplitude). In terms of the reflection coefficient  $\rho$ , the VSWR is just

$$\text{VSWR} = \frac{1 + |\rho|}{1 - |\rho|}.$$

For a purely resistive mismatch we can use our earlier expression for  $\rho$  to find that

$$\text{VSWR} = \begin{cases} R/Z_0 & \text{if } R > Z_0 \\ Z_0/R & \text{if } Z_0 > R. \end{cases}$$

From the VSWR you can find the magnitude (but not the phase) of the reflection coefficient:

$$|\rho| = \frac{\text{VSWR} - 1}{\text{VSWR} + 1}.$$

The VSWR is measured with a directional power meter. Knowing the values of forward and reflected power, you know from the defining equation above that

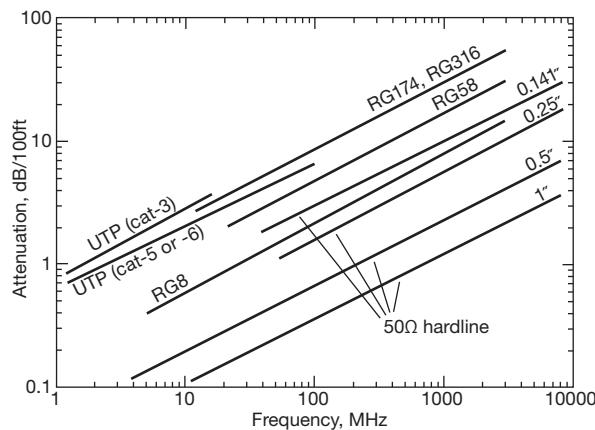
$$\text{VSWR} = \frac{1 + \sqrt{P_r/P_f}}{1 - \sqrt{P_r/P_f}}.$$

#### H.1.4 Loss in transmission lines

In the real world of non-ideal transmission lines, things aren’t quite as nice as we’ve led you to believe. Real transmission lines are *lossy*, meaning that signals are attenuated as they travel down the line; they are also slightly *dispersive*, meaning that signals of different frequency travel with slightly different speeds. The loss is *frequency dependent*: its value (often specified as attenuation in “dB per 100 ft”) increases proportional to  $\sqrt{f}$ ; i.e., a quadrupling of frequency doubles the loss (in dB) of a given length of line. This happens because the loss is dominated by the *skin effect*: when alternating current flows through a conductor, the current is not uniform throughout the bulk – it is confined to an outer layer (called the *skin depth*) of thickness  $\delta = (\pi \sigma f)^{1/2}$ , where  $\sigma$  is the conductivity and  $f$  is the frequency.<sup>7</sup> Because the skin depth decreases inversely with the square root of frequency, you have to quadruple the frequency to double the resistance (halve the skin depth),

<sup>7</sup> To be precise, the current density decreases exponentially, falling to  $1/e$  (37%) of its surface value at a depth equal to  $\delta$ .

which is equivalent (in terms of loss) to doubling the length of the line. This explains the approximate slope of the attenuation curves for transmission lines (Figure H.9), where lines of larger diameter have lower losses. Dielectric losses contribute additional attenuation at the highest frequencies. The curves shown are for a matched line; if there are reflections (i.e., if the VSWR is greater than 1.0) then the loss, *for a given net power transmitted down the line*, will be greater.

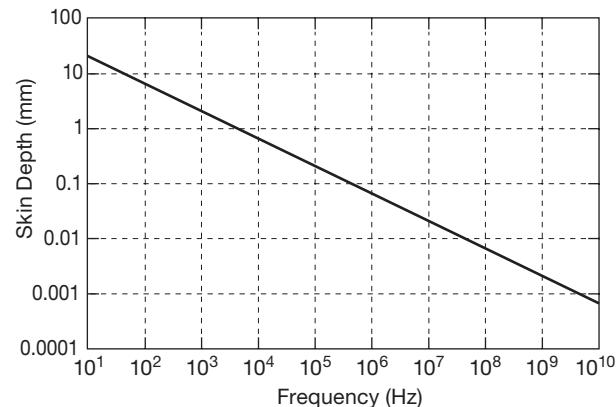


**Figure H.9.** Attenuation (dB/100 ft) as a function of frequency for several representative cable types.

It's useful to realize that skin-depth effects are significant even at low frequencies – current at the power-line frequency of 60 Hz is confined to a surface layer of roughly 1 cm in copper, for which the skin depth (in centimeters) at room temperature is given by  $\delta(\text{Cu})=6.6/\sqrt{f}$ ; you don't reduce power transmission losses much by using wire thicker than that. At radio frequencies the skin depth is so shallow (e.g.,  $\delta=10\ \mu\text{m}$  at 40 MHz) that you can make low-loss connections, inductors, and so on, by silver plating a poor conductor. A common technique for shielding lightweight instruments and computers is to apply a thin metallic plating to a plastic enclosure. Figure H.10 plots the skin depth in copper conductors from 10 Hz to 10 GHz.

## H.2 Impedance matching

Because you get reflections from unterminated (or incorrectly terminated) transmission lines, it's obviously a good idea to make sure you match impedances when you use coax lines whose electrical length is a significant fraction (at least 1/20, say) of the wavelength of the highest frequencies you're using. Stated in terms of time rather than frequency, you have to start worrying about termination



**Figure H.10.** Skin depth in copper as a function of frequency.

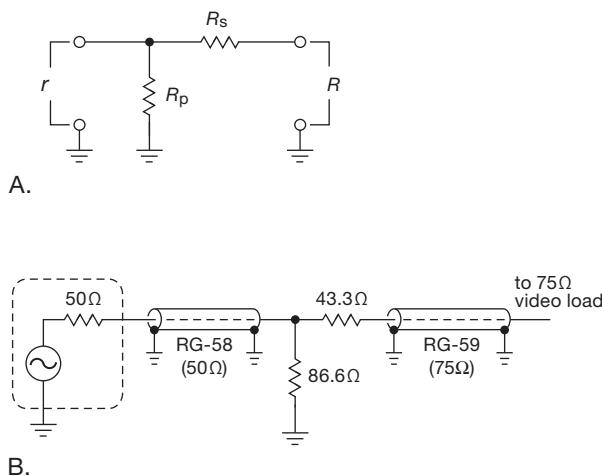
when the round-trip propagation time is about 20% of the signal rise time.

We've seen already that a simple way to do this is to terminate the line in its characteristic impedance (resistance), for example  $50\ \Omega$  for most coax lines. Termination is not required at *both* ends, because a terminated end swallows any incident signals. Thus you can use a mismatched signal-source impedance to drive a line terminated at the far end; or, as we saw above, you can "series terminate" the driven end of a line whose far end is unterminated (i.e., fed to a load of much higher impedance than the line). The usual practice, however, is to terminate *both* ends in the line's characteristic impedance (a conservative instinct that ensures a minimum of reflections). For example, you usually use a  $50\ \Omega$  cable to pipe the signal from a synthesizer or signal generator to a  $50\ \Omega$  matched load at the far end; if your load is high impedance, you place a  $50\ \Omega$  resistor across it (or use a  $50\ \Omega$  coax feedthrough termination).

Sometimes you need to match a line to a load (or source) of a different impedance; for example, you might want to measure the properties of some  $75\ \Omega$  video cable (that's the impedance that the video industry has chosen, much to the chagrin of the rest of the electronics community), and all you've got are  $50\ \Omega$  test instruments. Or, you might want to match the output impedance of a high-frequency amplifier to a length of cable that goes to an antenna.

This brings us to the subject of matching networks. In the following subsections we treat (a) resistive (lossy) networks for broadband impedance matching and attenuation, (b) transformer (lossless) broadband matching, and (c) reactive (lossless) narrowband matching.

### H.2.1 Resistive (lossy) broadband matching network



**Figure H.11.** A. A resistive *L*-network can match any pair of real (resistive) impedances; the parallel resistor  $R_p$  goes across the port of lower impedance  $r$ . B. Example of matching a  $50\Omega$  signal source and cable to a  $75\Omega$  video cable and load (with a loss of 5.72 dB).

You can easily figure out that two resistors (in the form of an “*L*-network,” Figure H.11A) is all it takes to match a pair of impedances  $r$  and  $R$  (assumed resistive, as all cables are); both sides are happy – each sees a matched load. The values of the matching resistors are

$$R_p = r \sqrt{\frac{X}{X-1}},$$

$$R_s = r \sqrt{X(X-1)},$$

where  $r$  is the smaller impedance and  $X$  is the ratio of impedances:  $X=R/r$ . Taking the earlier example, you can match a  $50\Omega$  test instrument to a  $75\Omega$  coax line (the common variety is called RG-59) by putting  $86.6\Omega$  across the  $50\Omega$  port and connecting it to the cable through a  $43.3\Omega$  series resistor (Figure H.11B).

The good news is that such a resistive *L*-network is frequency independent; the bad news is that it’s lossy. It’s easy to show that the loss is

$$\text{loss} = 20 \log_{10} \left( \frac{\sqrt{X}}{X + \sqrt{X(X-1)}} \right) \text{ dB.}$$

For example, the  $50\Omega$ -to- $75\Omega$  *L*-network above has a transmission loss of 5.72 dB for signals going in either direction. With a resistive match you have to accept this attenuation (this is sometimes called a *minimum loss pad*). We’ll see below how to make lossless matching networks with trans-

formers or with networks of *L*’s and *C*’s (“reactive matching networks”).

As you might imagine, you can do even *worse*, in terms of loss, with a network containing more resistors! In particular, you can add another resistor, making either a “*T*” network or a “*Pi*” network, that matches two resistive impedances to each other, with loss greater than the minimum loss we found above. Although this isn’t something you ordinarily want to do, there is a variation on that theme that is often useful; namely, a resistive attenuation network between a pair of already-matched impedances.

### H.2.2 Resistive attenuator

In radiofrequency circuits you sometimes need to attenuate a signal level – for example, to avoid overdriving a stage of gain. In other situations you need to use a resistive attenuator to provide some isolation between an impedance-sensitive component like an amplifier, mixer, or cable, say, and a component that is not impedance matched; an example of the latter is a filter, which typically is impedance matched in its passband, but reflective (a severe mismatch) in its stopband. Some amplifiers will oscillate if their output directly drives a sharp filter.

The solution to these problems is a resistive impedance-matched attenuator. The two topologies are *T* and *Pi*, named for their appearance on a schematic diagram (Figure H.12). It is not difficult to derive the resistor values:

$$R_p = \frac{1+x}{1-x} R_0,$$

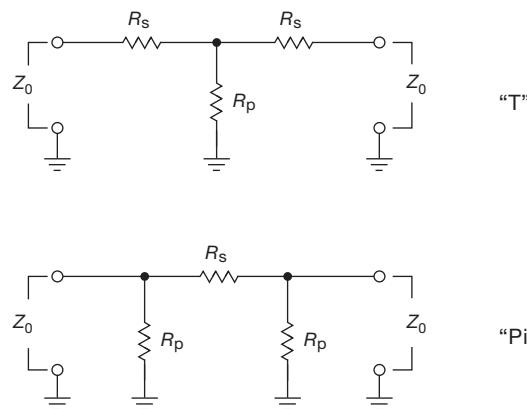
$$R_s = \frac{1-x^2}{2x} R_0, \quad (\text{Pi - network}),$$

and

$$R_s = \frac{1-x}{1+x} R_0,$$

$$R_p = \frac{2x}{1-x^2} R_0, \quad (\text{T - network}),$$

where  $x$  is given by the attenuation:  $\text{atten(dB)} = -20 \log_{10} x$  (or, equivalently,  $x = 10^{-\text{atten(dB)}/20}$ ), and  $Z_0$  (assumed resistive) is the impedance at both input and output. Tabulated values for  $50\Omega$  source and load impedances are given in Table H.1.



**Figure H.12.** Resistive T and Pi attenuators for equal input and output impedances.

### H.2.3 Transformer (lossless) broadband matching network

If the unavoidable loss of a resistive matching network is unimportant in your application, that certainly is the simplest method. However, in many applications it is essential to minimize loss – for example, in a communications transmitter or in low-level circuits whose performance is limited by amplifier or thermal noise.

In that case you can use either a transformer or a reactive matching network; neither method provides coupling all the way down to dc, however. Transformer coupling is relatively broadband, but it is limited in impedance ratios; reactive matching, by contrast, flexibly matches impedances (including reactive impedances), but only around some chosen center frequency. We treat reactive matching in the next subsection.

Transformers for use at signal frequencies are similar in principle to ordinary ac power transformers, that is, they use a pair of windings that are coupled magnetically, and whose turns ratio is the desired voltage ratio. The impedance ratio is then the square of the turns ratio. (For example, a transformer with a 1:4 primary:secondary turns ratio, driven with a  $50\ \Omega$  signal source, would present an output impedance of  $800\ \Omega$ , and should be loaded with that resistance.) However, because of the higher signal frequencies, it is necessary to use magnetic cores that do not have significant conductive paths for eddy currents. At audio frequencies the solution is to use the same sort of laminated metal stacks used in power transformers, but with much thinner laminations. At still higher frequencies, laminated cores are replaced either by powdered iron cores or by completely nonconducting magnetic “ferrite” materials. Because of the devastating effects of parasitic capac-

**Table H.1**  $50\Omega$  T and Pi Attenuators

Attenuation (dB)	Pi		T	
	$R_P$ ( $\Omega$ )	$R_S$ ( $\Omega$ )	$R_P$ ( $\Omega$ )	$R_S$ ( $\Omega$ )
0	$\infty$	0	$\infty$	0
0.25	3.47k	1.44	1.74k	0.72
0.50	1.74k	2.88	868	1.44
0.75	1.16k	4.32	578	2.16
1.00	870	5.77	433	2.88
1.25	696	7.22	346	3.59
1.50	581	8.68	288	4.31
1.75	498	10.1	247	5.02
2.0	436	11.6	215	5.73
2.5	350	14.6	171	7.15
3	292	17.6	142	8.55
4	221	23.9	105	11.3
5	178	30.4	82.2	14.0
6	150	37.4	66.9	16.6
7	131	44.8	55.8	19.1
8	116	52.8	47.3	21.5
9	105	61.6	40.6	23.8
10	96.3	71.1	35.1	26.0
15	71.6	136	18.4	34.9
20	61.1	248	10.1	40.9
25	56.0	443	5.64	44.7
30	53.3	790	3.17	46.9
35	51.8	1.41k	1.78	48.3
40	51.0	2.50k	1.00	49.0
45	50.6	4.45k	0.56	49.4
50	50.3	7.91k	0.32	49.7
55	50.2	14.1k	0.18	49.8
60	50.1	25.0k	0.10	49.9

Resistor values for T and Pi attenuators for use with  $50\ \Omega$  source and load. The values shown can be scaled for use at some other impedance, assuming equal input and output impedances.

itance and inductance, transformers for use at high radio frequencies (say above 10 MHz) generally are constructed from transmission lines (coaxial or parallel) wound around a magnetic core.

Impedance-matching transformers are widely available commercially, though for special applications you may need to design and wind your own. At audio frequencies many manufacturers offer miniature impedance matching transformers with “telephone” bandwidths (200 Hz–4 kHz), or full audio bandwidth (20 Hz–20 kHz); impedances go from loudspeaker and microphone impedances ( $8$ – $600\ \Omega$ ) up to “hi-Z” values of  $10k$ – $50k\Omega$ . There’s further discussion in §8.10.

A nice series of radiofrequency transformers is made by North Hills, including models that transform  $50\ \Omega$  or  $75\ \Omega$  to impedances up to  $1200\ \Omega$ ; these cover the frequency range between 20 Hz and 100 MHz, with a typical

frequency range of 1000:1 or more for a given transformer. For higher frequencies you can get broadband matching transformers from Mini-Circuits, covering the range of 4 kHz to 2 GHz with impedance ratios from 1:1 to 16:1, and with frequency ranges of 1000:1 or more for a given transformer. These are constructed with transmission-line techniques.

It is worth noting that transformer coupling provides *galvanic isolation*: input and output circuits need not share the same ground. This is particularly useful when you need to send a signal (or distribute a “house clock”) between instruments whose individual cases are each grounded through their power cords or enclosure rails. We have seen several instances in which instrument “ground,” in the same laboratory, differed by as much as *several volts* of 60 Hz ac. Here an isolated 50 Ω:50 Ω broadband transformer is ideal, for example the Mini-Circuits FTB1-6 (10 kHz–125 MHz) or the North Hills 0016PA (20 Hz–20 MHz).

#### H.2.4 Reactive (lossless) narrowband matching networks

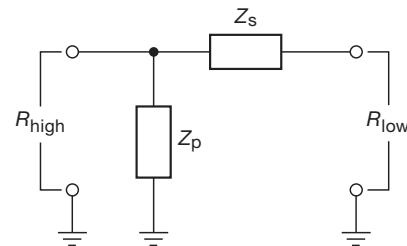
You can match *any* pair of impedances, real or complex, with just two reactive components. The resulting match is perfect only at a single frequency, but “good enough” over some modest band of frequencies. This can be considered an alternative to (broadband) transformer matching, with considerably greater flexibility in target impedances. It is worth noting that a lossless match between impedances that are not purely real (i.e., that have a reactive component) will always be narrowband.

The simplest reactive matching network is an L-network with one inductor and one capacitor (Figure H.13). You can choose either the inductor or the capacitor as the parallel element, but the network must have the parallel reactance located across the port with the larger impedance. The design procedure is straightforward, and is nicely motivated and explained in Hagen (see Appendix N). It amounts to choosing the parallel reactance to produce (in combination with the higher port’s impedance  $R_{\text{high}}$ ) the correct lower resistance  $R_{\text{low}}$  as its real part, then using the series reactance to cancel the resulting reactance.

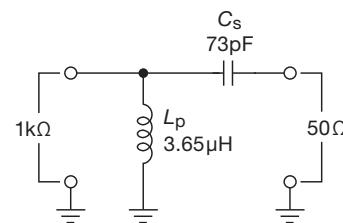
The procedure goes like this.

1. Calculate the quantity

$$Q_{\text{EL}} = \sqrt{\frac{R_{\text{high}}}{R_{\text{low}}} - 1}$$



**Figure H.13.** Lossless reactive impedance-matching L-network.



**Figure H.14.** Example of lossless network to match a 1 kΩ source impedance to a 50 Ω load, at a center frequency of 10 MHz.

(this will be twice the actual  $Q$  value, or frequency selectivity, of the matching network).

2. Now select the form of parallel reactance (i.e., inductor or capacitor), and set the magnitude of its reactance equal to  $R_{\text{high}}/Q_{\text{EL}}$  at the center frequency. In other words,  $L_{\text{parallel}}=R_{\text{high}}/2\pi f Q_{\text{EL}}$  or  $C_{\text{parallel}}=Q_{\text{EL}}/2\pi f R_{\text{high}}$ , respectively.
3. Finally, set the magnitude of the series reactance (i.e., capacitor or inductor, respectively) equal to  $Q_{\text{EL}} R_{\text{low}}$  at the center frequency. In other words,  $C_{\text{series}}=1/2\pi f Q_{\text{EL}} R_{\text{low}}$  or  $L_{\text{series}}=Q_{\text{EL}} R_{\text{low}}/2\pi f$ , respectively.

As an example, let us match a 1000 Ω source (an amplifier output) to a 50 Ω load (an antenna) at a frequency of 10 MHz. We find  $Q_{\text{EL}}=4.36$ , and, choosing a parallel inductor at the input,  $L_{\text{parallel}}=3.65 \mu\text{H}$  and  $C_{\text{series}}=73 \text{ pF}$  (Figure H.14). The  $Q$  of the resultant coupled network is equal to  $Q_{\text{EL}}/2$ , roughly  $Q \approx 2$ ; its bandwidth is thus about 50% between half-power points, though the match is perfect only at the center frequency. Note that the  $Q$  rises with increasing impedance ratio, and that you have no control over it. The network becomes a narrow bandpass filter for very large impedance-transformation ratios.

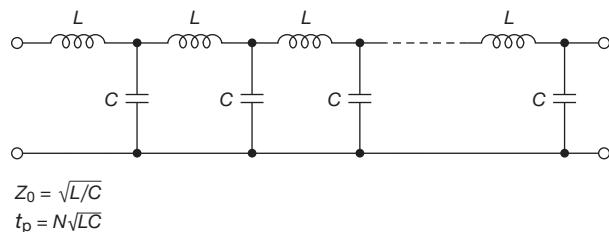
If you want a higher  $Q$ , you can get it by adding another reactive component, to form a Pi (or T) network. You can think of this as a pair of L-networks, going to an intermediate impedance that is much lower (or higher) than either port impedance. Each L-section then has an impedance

ratio greater than the final transformation, hence the higher  $Q$  value. You might think narrow bandwidth is bad, but in fact it is often desirable in communications circuits where you want to suppress out-of-band signal energy.

Alternatively, you can get lower  $Q$  than the simple L-network gives you by cascading a pair of Ls – a “double L.” Here the impedance transformation is taken in two half-steps – each section’s ratio is smaller than the final ratio, hence a lower  $Q$ .

### H.3 Lumped-element delay lines and pulse-forming networks

The continuous transmission line with inductance and capacitance per unit length of  $L/l$  and  $C/l$ , respectively, can be approximated by an array of  $N$  discrete series inductors  $L$  and shunt capacitors  $C$  (Figure H.15). It is easy to show that the resultant circuit approximates a transmission line whose propagation time per element is  $\tau_i = \sqrt{LC}$ , and whose impedance is  $Z_0 = \sqrt{L/C}$ ; the total propagation time is  $t_p = N\tau_i = N\sqrt{LC}$ .

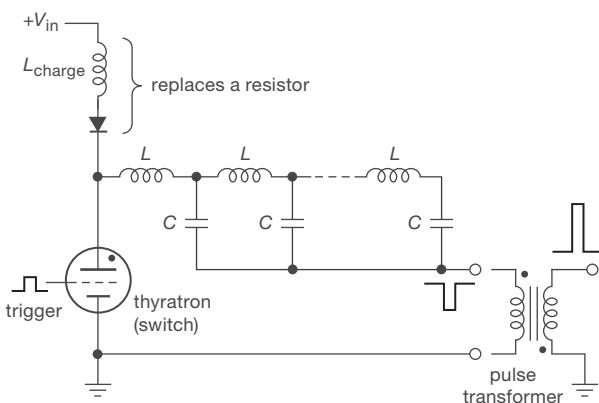


**Figure H.15.** Lumped-element delay line formed with an array of equal-value capacitors and inductors.

You can make a delay line this way, approximating a long transmission line. Roughly speaking, such a discrete approximation to a continuous transmission line will preserve details of the waveform only down to time scales of  $\tau_i$ , or  $1/N$ th of the total propagation time. For example, a  $1\ \mu\text{s}$  lumped delay line with 20  $LC$  sections will swallow details shorter than about 50 ns. It is a lowpass filter that attenuates frequencies above  $f=1/2\pi\sqrt{LC}$ .

Lumped-element delay lines were used in early analog oscilloscopes to allow time for the sweep to begin before the (delayed) signal reached the deflection circuitry; this let you see the triggering event (and a bit before). Later analog scopes used lengths of a helical-conductor coaxial line for the same purpose. Take a look at §H.4.3 for a bit more on this fascinating application.

Lumped delay lines are useful as *pulse-forming networks*, as shown in Figure H.16. Here the parallel capaci-



**Figure H.16.** Pulse-forming network for producing a high-voltage pulse of high energy. The thyatron is a special type of vacuum tube, containing a small amount of hydrogen or other gas, designed for switching really high voltages and high currents (10s of kV, 1000s of amperes: 10s of megawatts!). The inductor-diode fragment shown is a way to implement efficient “resonant charging” of a capacitor from a fixed dc voltage.

tance of a set of  $LC$  delay sections is charged to a high positive voltage; then the “center conductor” of the coax equivalent is switched to ground with a high-voltage switching element such as a thyatron. The common terminal (analogous to the coax “shield”) then produces a negative voltage pulse, of duration equal to *twice* the delay-line propagation time; its source impedance is just that of the delay line. This can drive a load directly; often it is converted to a different amplitude (and perhaps opposite polarity) with a pulse transformer, as shown. Pulse-forming networks find use in radar and other applications in which the pulse voltage and/or duration are inconvenient to produce with the analogous transmission-line circuit.

Shown also in this diagram is the method of “resonant charging,” in which an inductor  $L_{\text{charge}}$  plus diode replaces the conventional charging resistor for the purpose of charging a capacitor  $C_{\text{total}}$  (the  $N$  capacitors in parallel). This has several benefits: (a) the charging wastes no energy, whereas resistive charging wastes exactly 50%; (b) it is complete after a time equal to half the period of the resonant circuit formed by  $L_{\text{charge}}$  and  $C_{\text{total}}$ ; and (c) it charges the capacitor(s) to twice the supply voltage. Resonant charging is a clever technique, and it is used also in switching converters and power supplies, flashlamp circuits, and elsewhere.

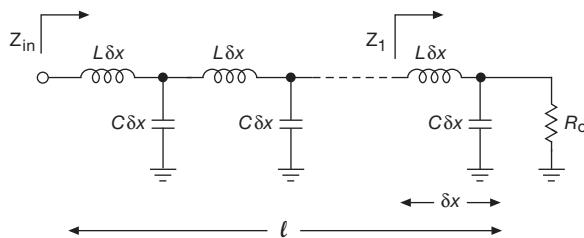
#### H.4 Epilogue: ladder derivation of characteristic impedance

In a formal course on “waves” you are usually subjected to an analysis that uses Maxwell’s equations to derive the relationship between the traveling  $\vec{E}$  field and  $\vec{B}$  field, from which follows the relation between voltage and current, and hence impedance. As a bonus you get the capacitance and inductance per unit length, and the speed of propagation.

But there’s a nice “circuit” way to convince yourself that a properly terminated transmission line looks like a pure resistance (equal in value to its “characteristic impedance,” for example  $50\ \Omega$ ), namely to model it as a discrete  $LC$  ladder (Figure H.17) consisting of little increments of length  $\delta x$ , each having a series inductance  $\mathcal{L}\ \delta x$  and shunt capacitance  $\mathcal{C}\ \delta x$ , where  $\mathcal{L}$  and  $\mathcal{C}$  are the inductance and capacitance per unit length of the coaxial line. There are  $l/\delta x$  of these in the whole length  $l$  of the line.

##### H.4.1 First method: terminated line

We start<sup>8</sup> by writing down the impedance looking into the last section of a terminated line ( $Z_1$  in Figure H.17), which will give us a condition on the ratio  $\mathcal{L}/\mathcal{C}$  in order for  $Z_1$  to equal, approximately,  $R_0$ . Then we’ll see that the impedance  $Z_{in}$  looking into the whole ladder converges exactly to  $R_0$  as we convert the discrete ladder approximation to a continuous transmission line by taking the limit as  $\delta x$  goes to zero.



**Figure H.17.** LC ladder model of a transmission line of length  $l$ . We’re ultimately interested in the limit  $\delta x \rightarrow 0$ , where the number of sections  $N=l/\delta x \rightarrow \infty$ .

Let’s do it.  $Z_1$  is just the impedance of  $\mathcal{L}\ \delta x$  in series with the parallel impedance of  $R_0$  and  $\mathcal{C}\ \delta x$ :

$$Z_1 = j\omega\mathcal{L}\delta x + \frac{R_0 \cdot (-j/\omega\mathcal{C}\delta x)}{R_0 - j/\omega\mathcal{C}\delta x}$$

<sup>8</sup> This treatment was inspired by Hagen’s *Radio Frequency Electronics*; see Appendix N.

$$\begin{aligned} &= j\omega\mathcal{L}\delta x + \frac{R_0}{1 + j\omega\mathcal{C}R_0\delta x} \\ &\approx R_0 + j\omega\delta x(\mathcal{L} - R_0^2\mathcal{C}), \end{aligned}$$

where in the last step we’ve kept only the first term of the binomial expansion, i.e.,  $1/(1+\varepsilon) \approx 1-\varepsilon$ .

The second term vanishes when  $\sqrt{\mathcal{L}/\mathcal{C}}=R_0$ , which is the formula for the characteristic impedance of a transmission line. But... not so fast – that term would have vanished anyway as  $\delta x \rightarrow 0$ ; however, our transmission line would have vanished as well! What we need to do is to include the next-order binomial terms, and see what happens as we let  $\delta x$  go to zero, while holding the total line length  $l$  constant; the number of sections then increases, as  $N=l/\delta x$ .

You can do the math. You’ll find that the next two terms add contributions<sup>9</sup> of order  $\delta x^2$  and  $\delta x^3$ , making  $Z_1$  look like

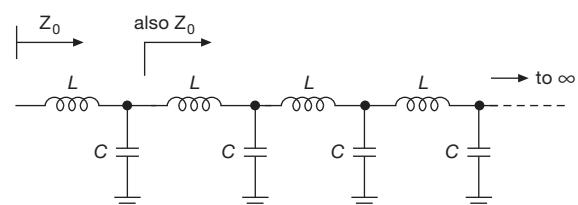
$$Z_1 \approx \left\{ R_0 + \mathcal{O}(\delta x^2) \right\} + j \left\{ \omega\delta x(\mathcal{L} - R_0^2\mathcal{C}) + \mathcal{O}(\delta x^3) \right\}$$

and so, cascading  $N$  such sections (where  $N=l/\delta x$ ), with the condition  $\sqrt{\mathcal{L}/\mathcal{C}}=R_0$ , the higher order terms vanish as  $\mathcal{O}(\delta x)$  and  $\mathcal{O}(\delta x^2)$  (for the real and imaginary parts, respectively) as  $\delta x \rightarrow 0$ . Thus the input impedance of a transmission line of length  $l$ , terminated in its characteristic impedance (resistance)  $R_0$ , is pure resistive and equals  $R_0$ .

##### H.4.2 Second method: semi-infinite line

Here’s a clever method<sup>10</sup> that doesn’t require approximation or worries about convergence. The idea is to look into one end of a lumped-element transmission line that extends to infinity (Figure H.18), noticing that it looks just the same if we step one notch to the right. So, calling the (complex) input impedance  $Z_0$ , we have, simply,

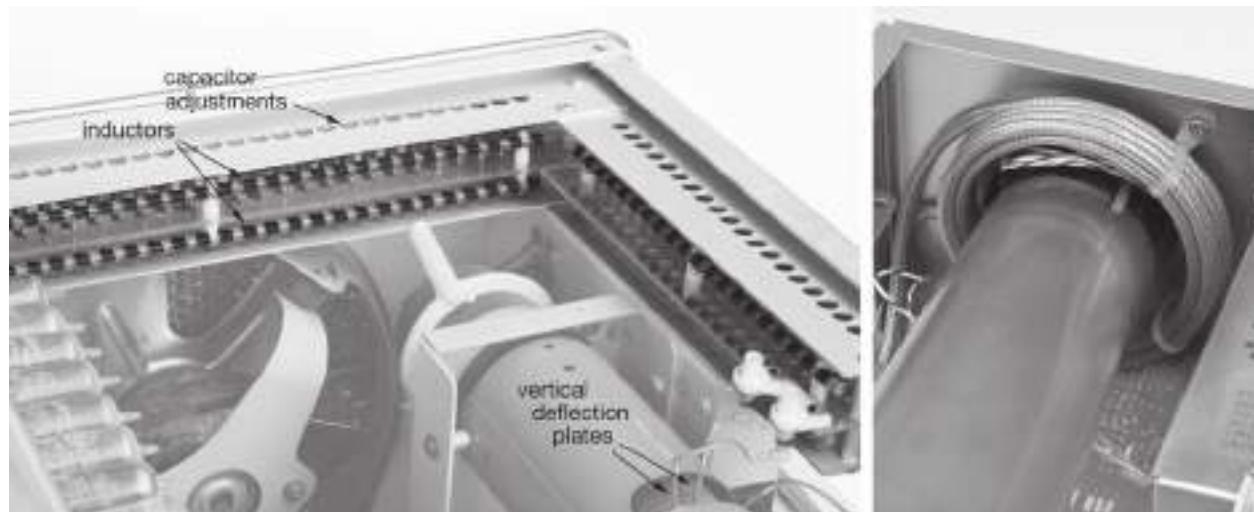
$$Z_0 = j\omega L + Z_0 \parallel Z_C = j\omega L + \frac{Z_0 \cdot (-j/\omega C)}{Z_0 - j/\omega C}.$$



**Figure H.18.** Semi-infinite LC ladder model makes for an easy calculation.

<sup>9</sup> They are  $-R_0^3\omega^2C^2\delta x^2 + jR_0^4\omega^3C^3\delta x^3$ , if you want to check your math (or ours!).

<sup>10</sup> Suggested to us by Jene Golovchenko.



**Figure H.19.** Analog oscilloscopes used delay lines in the signal path, so you could see the triggering event. The 1959-vintage vacuum-tube Tektronix 545A (left) had a 30 MHz bandwidth, and used a two-channel lumped-element 200 ns delay line consisting of 50 inductor pairs and 50 (adjustable!) trimmer capacitors; (Figure H.20); their 1982 vintage solid-state 2213 'scope (right) had 60 MHz bandwidth, and used a 2.5 m length of spiral-conductor coaxial cable for its 100 ns delay line.

Multiplying through by the denominator of the last term and rearranging terms, we get a quadratic equation for  $Z_0$ :

$$Z_0^2 - j\omega L Z_0 - L/C = 0,$$

with the solution

$$Z_0 = j\omega L \pm \frac{\sqrt{4L/C - \omega^2 L^2}}{2}.$$

Now for the *coup de grâce*: we let the individual segments shrink toward zero while keeping the full line length. The individual  $L$ 's and  $C$ 's go to zero, but their ratio remains constant. Only the  $4L/C$  term survives, giving us the (real) impedance  $Z_0 = \sqrt{L/C}$ . No approximations!

#### H.4.3 Postscript: lumped-element delay lines

The clever designers of analog oscilloscopes, back in the dark ages of electronics, found a way to get the horizontal trace going *before* the triggering event, namely by delaying the displayed signal by  $\sim 100$  ns, using a delay line.<sup>11</sup> The designers of early vacuum-tube 'scopes (like the legendary Tektronix 545A) used a lumped-element transmission line like the one in Figure H.19 to achieve the delay (200 ns in this case) that otherwise would have required more than

100 feet of cable; presumably they also felt some satisfaction in exploiting the theory they learned in an electronics course they had taken years earlier.<sup>12</sup> Later analog 'scopes replaced the lumped-element delay line with a cleverly engineered coax delay line with a pair of spiral inner conductors; its larger inductance per unit length increased the signal delay,<sup>13</sup> and, happily, increased the characteristic impedance as well. The photograph shows an example, where the cable has been stuffed into an empty space at the rear, conveniently wrapped about the CRT.<sup>14</sup> Figure H.21 reveals the inner secrets of this elegant differential-pair delay line, whose crisscrossing counterwound helical "center conductor" produces a delay of 12 ns/ft. And Figure H.22

<sup>12</sup> The capacitors in Figure H.19's lumped-element delay line are in fact connected to the *midpoint* of each inductor, as can be seen in the official schematic of Figure H.20. It turns out that this is a more efficient implementation of a finite lumped line, as explained to us over a dinner of fine conversation and Persian cuisine by Larry Baxter ("Mr. Capacitive Sensors"; see the book by the same name).

<sup>13</sup> By a factor of  $\pi n D$ , approximately, where  $n$  is the number of turns per unit length and  $D$  is the spiral diameter. Because the coarse-pitch spiral is closely surrounded by the shield, you can think of the signal as propagating, corkscrew fashion, *along* the spiral; that approximation then gives you this simple expression, without the need to calculate the inductance and capacitance per unit length.

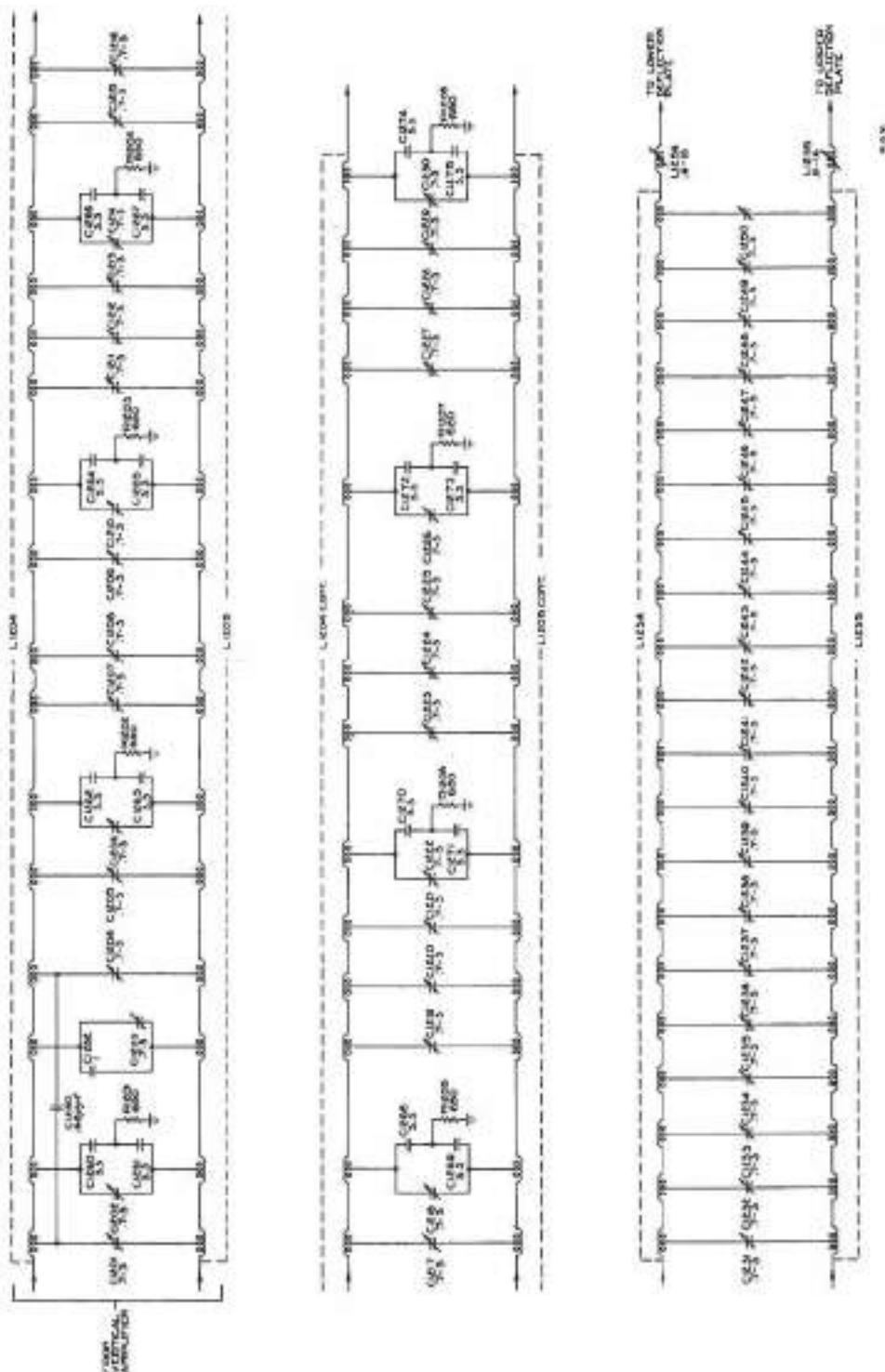
<sup>14</sup> That's *cathode ray tube*, for those born in this millennium, and thus deprived of the opportunity to admire one.

<sup>11</sup> Digital 'scopes finesse this problem by using digital memory to store some pretrigger digitized samples.

DELAY LINE NETWORK

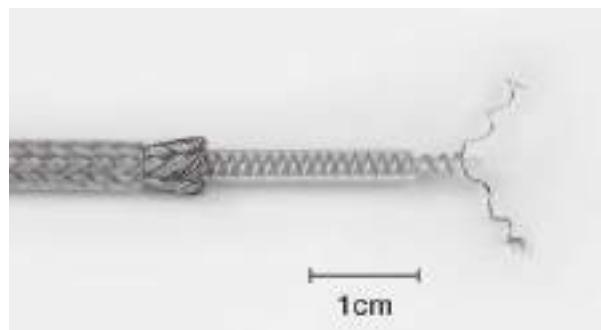
C

TYPE 545 OSCILLOSCOPE

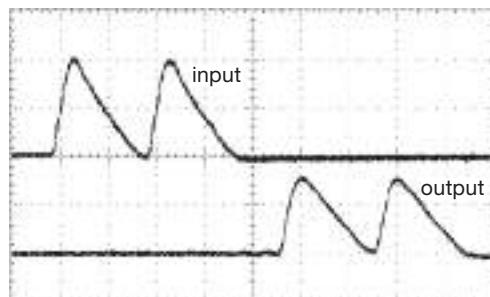


**Figure H.20.** Working hard to achieve perfection: Tektronix used fifty inductor pairs and fifty trimmers for the delay line in their type 545 'scope (pictured in Figure H.19) from the 1950s. Reproduced with permission of Tektronix Inc.

shows the observed delay when a differential triangular pulse-pair waveform, launched into one end, is received at the other end.



**Figure H.21.** We borrowed a Tektronix 2213 carcass from Brian Shaban, and, after a bit of surgery, look what we found inside! We measured a 100 ns delay, and a differential impedance of  $155\Omega$ , for the 8.5 ft cable that used to live in this 'scope. The counterwound helix consists of two insulated 30 ga wires, with a pitch of 1.125 mm and an average diameter of 2mm, insulated from the surrounding braided shield of 3.25 mm inside diameter.



**Figure H.22.** We launched this analog waveform down the cable of Figure H.21 (well, what remained of it, anyway) to see the signal delay of  $\sim 95$  ns. The delayed signal is of good fidelity, with a test signal bandwidth somewhat greater than that of the 60 MHz 'scope. Vertical: 1 V/div; horizontal: 20 ns/div.

# TELEVISION: A COMPACT TUTORIAL

## APPENDIX

This appendix evolved from a tutorial, originally written for a nontechnical audience.<sup>1</sup> It is organized as follows:

### **Television: video plus audio**

#### **Combining and sending the audio + video:**

##### **Modulation**

#### **Recording analog-format broadcast or cable television**

##### **Digital Television: what Is it?**

##### **Digital Television: broadcast and cable delivery**

##### **Direct satellite television**

##### **Digital video streaming over internet**

##### **Digital cable: premium services and conditional access**

##### **Digital cable: video-on-demand**

##### **Digital cable: switched broadcast**

##### **Recording digital television**

##### **Display devices (CRT; LCD; plasma; OLED)**

##### **Video connections (analog; DVI/HDMI; DisplayPort)**

### **I.1 Television: video plus audio**

¶ 1. Television involves the remote delivery of a moving picture plus sound. It is accurate to think of the *sound* as continuous; however, the *picture* is captured, and then delivered, as a succession of still images, at a rate fast enough that the viewer perceives a scene of continuous motion.<sup>2</sup>

¶ 2. Television is distinguished further, of course, by the transmission of this movie-like content to the remote viewer. Originally this was carried out exclusively by terrestrial transmission, via radio waves, to the viewer's antenna and television set. Over time, other methods of transmission have been added – electrical cable,<sup>3</sup> optical fiber, direct satellite transmission via microwaves, and, of course,

the Internet – along with recording methods such as magnetic videotape (Betamax, VHS, D-VHS), and optical discs (Laserdisc, VideoCD, DVD,<sup>4</sup> HD-DVD, Blu-Ray, and others).

#### **I.1.1 The audio**

¶ 3. The *audio* portion of television is perhaps more easily understood, as it differs little from ordinary sound-recording techniques. A microphone converts the instantaneous sound pressure variations into an electrical signal; that is, it creates as its output an electrical voltage that at each moment is proportional to the pressure of the sound wave to which it is exposed. Contemporary audio recording and delivery usually employs two or more microphones, creating “stereo” sound (i.e., two channels), or multichannel sound (e.g., “5.1 channel sound”).

¶ 4. Traditionally these signals were processed, stored, and delivered by “analog” methods, which means simply that the signals were treated as smoothly varying voltages as they passed through the electronic innards of the amplifiers, recorders, modulators, and so on.<sup>5</sup> Contemporary “digital” technology does it differently: almost as soon as possible, the microphone’s signal (the varying voltage that represents the sound) is converted to a succession of numbers (it is *digitized*), and everything that follows is some form of arithmetic on this torrent of numbers that come tumbling out. Only at the final stage – recreating the recorded sound for the listener – is the digital representation converted back to an analog voltage, and then, in the loudspeaker, to a reproduction of the original sound pressure wave.

¶ 5. Just to give a sense of the quantity of numbers involved, in the standardized recording technology of the

<sup>1</sup> Following the stylistic conventions of that audience, the paragraphs are numbered sequentially.

<sup>2</sup> For conventional cinema-style movies, the rate is 24 frames/second; television in the United States uses a rate of approximately 30 frames/second.

<sup>3</sup> Known technically as *coaxial transmission line*.

<sup>4</sup> “Digital Versatile Disc.”

<sup>5</sup> Common analog recording technologies, now nearly obsolete, include the vinyl record (where the audio signal waveform is embossed as small displacements of a fine groove), and the audio cassette tape (where the audio signal waveform is recorded as patterns of magnetization on a thin layer of magnetic oxide coating on a flexible plastic tape).

compact disc (CD), the instantaneous sound is *sampled* at a rate of 44,100 times per second (in both stereo channels simultaneously), and each such sample pair is converted (“digitized”) to a 16-bit binary number.<sup>6</sup> So the bits are tumbling out at a rate of  $2 \times 44,100 \times 16 = 1,411,200$  bits per second, or nearly 100 million bits per minute.<sup>7</sup>

¶ 6. One might ask why any sane person would want to deal with such a quantity of numbers, when the original analog representation of the sound was so much simpler – just a pair of voltages that were varying at most 20,000 times per second.<sup>8</sup> The reasons are several, but they boil down to the contemporary ease and economy of digital processing, combined with the higher efficiency and quality of storage and transmission of audio (and video) that has been properly digitized. To get a sense of those advantages, one need only marvel at the gorgeous images transmitted daily from planetary probes visiting Mars and Saturn – images that are free of “snow” and other artifacts irreparably added to analog transmission by the effects of unavoidable electrical interference – to appreciate the benefits of error-free digital transmission. And, to get a sense of the density of digital storage, we note that a contemporary 5" optical disc (dual-layer Blu-ray disc) holds 80 hours of CD-quality audio, or ten times that amount if modestly “compressed,” compared with a mere hour’s storage of analog audio on the 12" vinyl LPs of yesteryear.<sup>9</sup>

### I.1.2 The video

¶ 7. The *video* is by far the more complicated part of television. The challenge is to reproduce a scene with motion, in color, while preserving adequate fidelity and absence of artifacts. And this must be done within the resources of the storage and delivery channels – that is, with finite disk storage and speed, and with finite transmission (via broadcast tower, cable, Internet, or satellite) *bandwidth*.<sup>10</sup>

<sup>6</sup> That is, a number ranging from -32768 to +32767, those bracketing the “full-scale” range of the recorded sound.

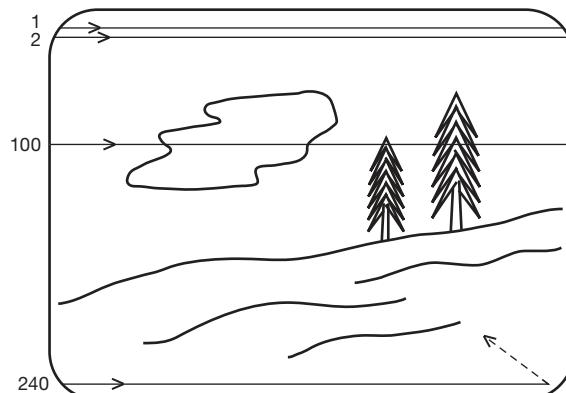
<sup>7</sup> The *recorded* bitrate is roughly triple this figure because of coding, error-correction redundancy, and the like.

<sup>8</sup> Or 20 kHz, the upper limit of human hearing; and that only for one of relative youth, such youth further possessed of sufficient wisdom to avoid deafening rock concerts.

<sup>9</sup> And a contemporary 3 terabyte 3" magnetic hard drive that you can hold in your hand holds yet another factor of 60, or 50,000 hours of excellent quality (128 kbps AAC) compressed stereo audio; that’s 15 years of 40-hour per week music!

<sup>10</sup> *Bandwidth* refers to the range of frequencies that can be carried on the cable or other transmission medium. It is technically accurate to think

¶ 8. Video systems begin with a camera that has an electronic sensor (analogous to a digital camera) and that converts the two-dimensional color scene on that sensor into a succession of *frames*, each of which represents the image at those successive times (for US TV, the rate is approximately 30 frames per second). In traditional *analog* television, the two-dimensional image is converted into an electrical signal by the following method: imagine a single frame, that is, a still picture. To keep it simple, imagine further that it is monochrome; that is, “black and white.”<sup>11</sup> We begin at the upper left, and move horizontally across the picture, generating an electrical voltage proportional to the brightness at each point as we pass by. When we reach the right-hand border, we jump back to the left edge, continuing with another horizontal path, slightly below the first. See Figure I.1. We continue in this way until we reach the bottom right-hand corner, at which time we have scanned the entire frame once, in what is known as a *raster* (“grid” in German) pattern.<sup>12</sup>



**Figure I.1.** A static two-dimensional image is “raster scanned” to create a video waveform (Figure I.2) representing intensity along the scan lines.

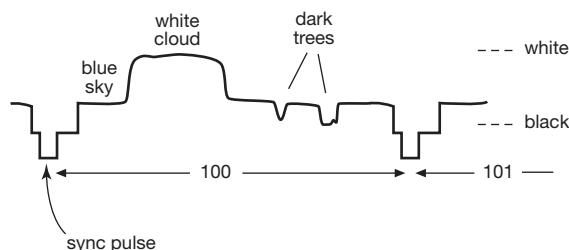
¶ 9. What we have done, then, is to generate an elec-

of this, for example, as the range of stations on the radio dial that could be carried with fidelity by a single electrical cable (or other medium). The term is sometimes used loosely to refer to rate of data transfer.

<sup>11</sup> Or, more accurately, *grayscale*.

<sup>12</sup> Traditional standard definition television (SDTV, usually called “NTSC,” for National Television System Committee, and going back to the 1940s) in the United States divides the whole picture into 480 horizontal lines, along each of which roughly 640 features (picture elements, or *pixels*) could be resolved; a computer user would not be terribly impressed – he or she would say that standard NTSC television has only “VGA” resolution (i.e., 640×480).

trical representation, in time (a varying voltage proportional to brightness at each point in the image) of a single two-dimensional image; that is, we've converted a two-dimensional image into a one-dimensional output voltage. See Figure I.2.



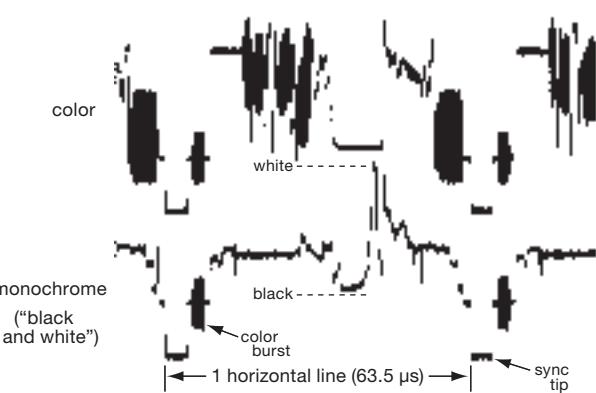
**Figure I.2.** A portion of the video waveform from Figure I.1, representing one of the 240 horizontal scan lines.

¶ 10. This time-varying voltage is called the *video signal*, and it is the first step in creating a television image. In traditional NTSC analog television, this signal was transmitted by analog methods, after a process called *modulation* (more later), and was recovered and used by the television set to paint the picture on the screen, performing the same raster scan (left to right, top to bottom). Each frame follows in sequence, presenting a succession of 30 pictures per second on the television set's viewing screen.<sup>13</sup>

¶ 11. To complete the video signal, it is necessary to add some synchronizing information, so the television set knows when to begin painting a frame, and also when each horizontal line begins. In traditional NTSC television this is done by adding a horizontal *sync pulse* at the beginning of each horizontal line, which is just a short<sup>14</sup> voltage pulse that, if it were in the middle of a picture, would be interpreted as "blacker than black." See Figure I.3. The television set detects these pulses and uses them to synchronize its scanning across each line. Likewise, a unique *vertical sync pulse* is transmitted for each field which informs the

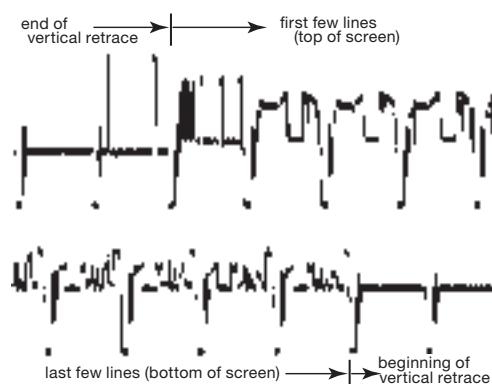
<sup>13</sup> To complicate things, NTSC uses a method known as *interlacing*, in which a coarser raster – omitting every other horizontal line – is performed at twice the rate. Thus, in standard NTSC US television, the viewer sees 60 pictures ("fields") per second, each of which has only 240 horizontal lines; two such fields, with their interlaced lines, form one complete 480-line frame. This is sometimes called a "480i" format, to distinguish it from formats with higher resolution (e.g., high-definition TV, HDTV, with 720 lines or 1080 lines, or "4K Ultra HDTV," with 2160 lines), or from those without interlacing (which are known as *progressive*; e.g., 720p).

<sup>14</sup> About 4.5 millionths of a second.



**Figure I.3.** Composite analog video signal of one horizontal line, framed by horizontal sync pulses. The brightness ("luminance") is represented by its amplitude. Color is accommodated by adding a modulated 3.58 MHz "chrominance" subcarrier, whose amplitude represents degree of saturation and whose phase encodes the colors.

television set when to return to the top to begin painting the next field or frame (see Figure I.4). The complete video picture signal, with its added sync pulses, is called *composite video*.



**Figure I.4.** The vertical retrace (beginning of a new field) is signaled by a set of tailored sync pulses, the first and last of which are shown here.

## I.2 Combining and sending the audio + video: modulation

¶ 12. Continuing for the moment with traditional NTSC television (as opposed to *digital* television, whose standards are known as ATSC, set by the Advanced Television Systems Committee, and which will be explained later), the composite video, along with the audio, must now be

sent, via transmitting tower or cable, to the home viewer. Naively one might think of simply transmitting these signals “as is.” This is not done, however, for at least two reasons: first, if the composite video signals were transmitted directly, then any two television signals would overlap and jam each other (because they would all share the same frequency band, namely that of the raw video signal itself); second, some wavelengths are more conveniently generated and propagated than others. For these reasons, the audio and video content of television signals (and, indeed, all communications and broadcast signals) are instead used to vary some aspect of a “carrier” wave, chosen at some specified wavelength. That carrier wavelength (or, equivalently, frequency) defines the “channel”; and the process of impressing the information (video and audio) onto the carrier wave is known as *modulation*.

**¶ 13.** Radio stations use the same technique: AM stations vary the strength (amplitude) of the carrier (hence “amplitude modulation”), whereas FM stations vary the frequency (“frequency modulation”). The carrier frequency itself defines the channel: in the US, AM stations are assigned to carrier frequencies between 520 and 1710 kilohertz (kHz, thousands of cycles per second), while FM is assigned to the band of carriers from 88 to 108 megahertz (MHz, millions of cycles per second). In the US, broadcast television begins at 54 MHz (Channel 2), and ends at 698 MHz (Channel 51), with gaps for FM, aeronautical, and other services.<sup>15</sup>

**¶ 14.** When information (video, for example) is modulated upon a carrier, the resultant signal spreads out and occupies a small band of frequencies. For example, when an FM station varies the frequency of its assigned carrier to carry its audio signal, the resulting signal occupies about 150 kHz. So FM stations are assigned channels separated by 200 kHz (to allow a “guard band” of 50 kHz in addition to their 150 kHz signal) – and that is why the frequencies of FM stations always end in an odd number after the decimal point (for example New York City’s WNYC is at 93.9 MHz), ensuring a minimum spacing of 0.2 MHz (= 200 kHz).

**¶ 15.** Traditional analog NTSC broadcast television used a variant of AM for the picture signal (composite video) and, separately, FM for the sound signal.<sup>16</sup> The assigned TV channels are spaced apart by 6 MHz, each station being

permitted to occupy nearly that amount, after allowing for a small guard band. Television sets “know” the frequencies allocated for each channel and tune to the correct frequency when the user chooses the channel number. For example, if (during television’s analog era) one tuned to Channel 13 in the New York City area, the television set’s electronics selected the station transmitting on 210 MHz (assigned by the FCC as Channel 13), namely WNET. The electronics in the set *demodulates* the received signal, recovering composite video and audio. The video, with its embedded sync signals, is used to paint the picture, frame after frame, while the audio is sent to the loudspeakers.<sup>17</sup>

**¶ 16.** Broadcast television (and radio) takes place on what is often called the “public airwaves.” One needs only a television set (or radio) and an antenna to receive these over-the-air (OTA) public transmissions. Although some countries require licensing of receiving devices (radios and television sets), in the United States the broadcast services are freely available to anyone within range of a transmitting tower.

**¶ 17.** Depending on the distance and path from the broadcast station to the viewer, the “antenna” can be as simple as an indoor “bowtie” or pair of “rabbit ears,” or as elaborate as a roof-mounted multi-element structure. Whatever its form, the antenna’s function is to create an electrical signal on the feedline, induced by the speed-of-light broadcast signal passing by the antenna’s site. Receiving antennas intended for broadcast television are designed to work over the range of frequencies used by broadcasters (see ¶13); thus the electrical signal delivered to the television set includes multiple stations, and it is the job of the TV tuner to select and process the channel to be viewed.

**¶ 18.** *Cable* television sends traditional analog TV channel signals in almost exactly the same way as broadcast. An evident difference, however, is that the channelized signals are received at the viewer’s end from a coaxial cable (rather than being received by the viewer’s television antenna), and then connected to the television set directly (i.e., to its normal antenna connector on the rear). Alternatively, for additional cable services (such as premium channels) the incoming cable connects to a “set-top box” (STB) provided by the cable company, the output of which is connected to the viewer’s television set (or flat-screen monitor,

<sup>15</sup> You can download a gorgeous multicolor wall-sized spectrum allocation figure from [http://www.ntia.doc.gov/files/ntia/publications/spectrum\\_wall\\_chart\\_aug2011.pdf](http://www.ntia.doc.gov/files/ntia/publications/spectrum_wall_chart_aug2011.pdf).

<sup>16</sup> That is, the picture and sound signals are carried simultaneously on a

pair of designated carrier frequencies within the single assigned television channel.

<sup>17</sup> In this primer we have ignored details associated with reproduction of color (versus black and white).

projector, etc.; see ¶56ff).<sup>18</sup> The channel *frequencies* are also somewhat different, with Channels 2–13 chosen the same as for terrestrial broadcast, but with the remaining channels reassigned to eliminate gaps; this can be done because the cable is a private world of its own, isolated so as not to interfere with, or be interfered by, other broadcast services.

¶ 19. A third difference is that some of the cable content is delivered as a subscription, for which the viewer pays additional fees; examples are premium services such as Home Box Office (HBO). These require some method for permitting or denying viewable delivery of selected channels or programs. Continuing for the moment with analog cable (whose days are numbered!), this can be done in several ways: the simplest is by installing filters (to block un-subscribed channel frequencies) at the utility pole, where the subscriber's cable splits off from the trunk running along the street;<sup>19</sup> A more sophisticated method involves scrambling the cable-borne analog signals of subscription programs,<sup>20</sup> and then instructing the set-top box (via digital communication from the cable provider to the individual STB) as to which programs may be unscrambled.

¶ 20. It is worth noting that cable companies have been required to carry the broadcast stations in their area, normally as analog cable channels.<sup>21</sup> Each such program occupies a cable channel (frequency). However, they may distribute additional services via digital methods ("digital cable") on additional channels (frequencies), which they much prefer; that is because, with digital methods, it is possible to carry up to *ten* NTSC-quality programs (i.e., SDTV, for Standard-definition TV) on a single channel. This is called *multicast*: the ability to carry multiple *programs* on a single channel (i.e., frequency). And, note that a cable can carry more than 100 such carriers – permitting more than 1000 simultaneous programs.

<sup>18</sup> For better picture quality, the latter connection is usually made not to the set's antenna input (called "RF," for Radio frequency, meaning the modulated channels discussed above, in ¶¶12–15), but to special audio-video inputs, with names like *s-video*, *component video*, *composite video*, *DVI*, or *HDMI*; see ¶64ff. The latter pair are *digital* connections, discussed below in connection with digital TV.

<sup>19</sup> Vintage cable subscribers will remember calling the cable company to add a movie channel, whereupon a cable truck appeared, the cable guy went up the pole to fiddle with something (changing the filter), and, voilà, movies on your television!

<sup>20</sup> For example, by suppressing the horizontal sync pulses or inverting the video (interchanging black and white).

<sup>21</sup> Unless all subscribers are provided with STBs that can receive digital delivery.

¶ 21. Previewing some additional characteristic of digital television: digital cable permits flexible subscriptions, with a program being authorized on-the-fly (e.g., pay-per-view, or video-on-demand). It also provides a natural mechanism for content protection via encryption. It allows for interactive participation, via a reverse channel back to the cable provider. It permits the delivery of high-definition content, with more than the 480 lines of NTSC (up to 1080 lines, at the highest quality currently supported). Finally, it provides a natural way to time-shift, pause, or replay live programs, via computer-type hard-disk storage.

¶ 22. Analog broadcast was sent into retirement in the United States in June of 2009, and all television broadcast delivery is now done by digital methods (more to follow). This conversion-to-digital process is taking place worldwide and will likely be complete by 2020 or earlier.

### I.3 Recording analog-format broadcast or cable television

¶ 23. Video recording was complex and costly (and therefore confined to the broadcast studios) until 1975, when home video-recording devices were introduced in the US by Sony ("Betamax") and its competitors ("VHS," for video home system). These devices replicate the "front end" of a television set, to recover video and audio from the incoming signal (broadcast or cable), and use a clever spinning tape head arrangement<sup>22</sup> to capture on magnetic tape a reasonable replica of an NTSC analog television program. The technique is entirely analog (no digitization, no numbers) and records only onto special video-tape media (no computer media, no hard disks, etc.), as a magnetic recording (analogous to an analog audio tape recording; see the footnote at ¶4).

¶ 24. Videotape technology has been upstaged by digital alternatives such as optical disc recording (most famously in the form of DVDs and Blu-Ray discs – whether sold with prerecorded content, or recorded with a disk recorder), which creates a permanent copy of the video material; or by recording to a computer-type hard-disk drive (hdd), where the video copy is stored as a computer file. These digital methods require that the program material be converted from analog to digital form, if it is not already. (This is done internally and automatically in devices like TiVo®)

<sup>22</sup> This is known as a "helical" tape head, which creates successive narrow slanted tracks across the slowly moving tape, each one holding one field of video. The use of a rapidly moving tape *head* eliminates the need for rapidly moving *tape*.

and other personal video recorders.) Digital television and digital video are discussed next.

#### I.4 Digital television: what is it?

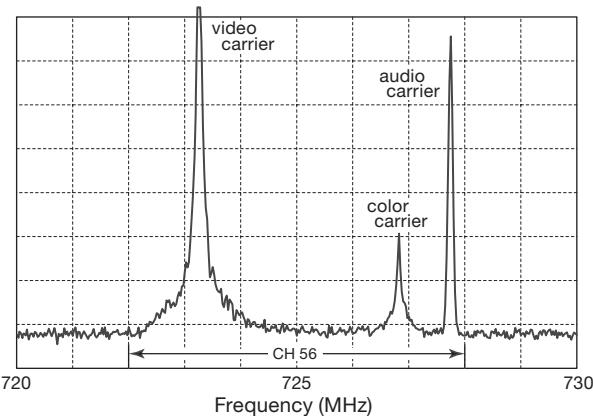
¶ 25. Just as an *audio* signal can be digitized (i.e., its instantaneous amplitude is measured, at rapid intervals, and converted to a succession of numbers), and subsequently transmitted, stored, or processed (¶¶4–6, above), so it is possible to digitize the *video* signal that represents successive frames of picture. Although one could imagine simply sending the digitized version of traditional NTSC as “digital TV,” in practice one takes advantage of the enormous processing finesse of contemporary digital electronics to economize by *compressing* the raw video signal to a small fraction of its native size before it is delivered. The use of compression, along with the fact that a digital signal is “just numbers,” permits the delivery of multiple programs on what would otherwise carry just a single video signal (program), typically by a factor of five to ten.

¶ 26. There are several reasons for this improvement. One is the ability to detect and correct transmission errors by numerical techniques, allowing one to operate with received signal levels that are close to the “noise” (from interference, or signal loss owing to range or obstructions); with purely analog transmission a large received signal/noise ratio is necessary to reduce the visible effects of noise (“snow”) to acceptable levels.

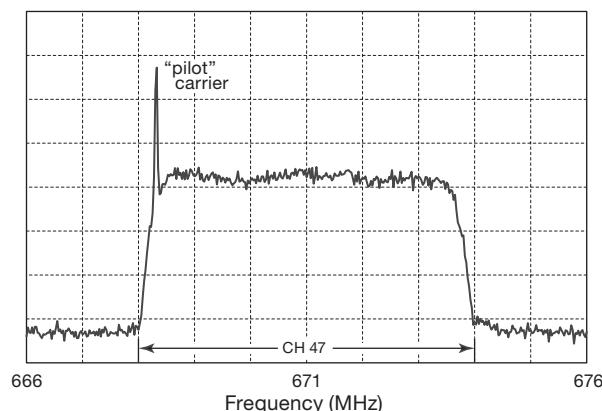
¶ 27. A second reason is the spectral efficiency of digital transmission – or, more accurately, its improvement compared with the *inefficiency* of analog signaling. This can be seen in Figures I.5 and I.6, a pair of spectra taken directly off the home antenna of one of the authors in March of 2009, a time during the switchover to digital when both analog and digital broadcasts were taking place (see also Figure I.7).

¶ 28. Compression aims to reduce by a large factor (ten-fold or more) the quantity of numbers needed to describe the succession of picture frames, without significantly degrading the image quality. This seemingly impossible task takes advantage of redundancies in a moving picture, and of limitations in human visual perception.

¶ 29. Contemporary digital-video compression is a rich and mathematically complex subject, the result of enormous effort in the applied mathematics and electrical engineering communities over the last several decades. But the basic tricks are easy enough to understand. The pro-



**Figure I.5.** The spectrum of 6 MHz-wide analog Channel 56 in Boston, as seen in May 2009. The video information resides in the sideband tails, while most of the transmitted power is wasted in the non-informational video carriers.

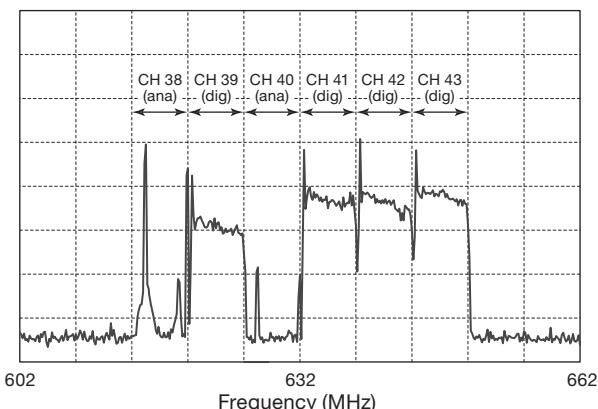


**Figure I.6.** Digital Channel 47, seen also in May 2009, fills its 6 MHz spectrum allocation with digitized video. It carries five times as many programs, with comparable (or better) picture quality.

cess begins by exploiting the fact that portions of an image near each other tend to be similar; so one can encode and send the (smaller) *differences* of brightness and color from a set of reference points, rather than the full description of brightness and color at each point. Likewise, successive frames tend to be similar, so one can define a sparse collection of index frames and send only the differences for intervening frames.<sup>23</sup> A further trick exploits the fact that the image often contains moving objects or a panning camera; so it is efficient to calculate “motion vectors” predicting

<sup>23</sup> More precisely, it is the corrections from an *interpolated guess* between index frames (or reference points within a frame) that are sent.

the approximate motions, and then send only the (smaller) corrections from the predicted values.



**Figure I.7.** A country in transition: RF spectrum of Channels 36–45 (each permitted 6 MHz of spectrum) as seen on our antenna feedline in Cambridge, Massachusetts, on May 6, 2009. Analog Channels 38 and 40 each carry one standard-definition (SDTV) program, whereas Digital Channels 39 and 41–43 can carry up to five SDTV programs each (though it's more common to see one HDTV and one SDTV program).

**¶ 30.** These methods greatly reduce the needed bitrate (number of numbers per second), and they do so *without any loss of picture quality whatsoever* – they are “lossless.” That is because the original digital image can be fully and exactly recovered by applying the numerical differences in the reverse order. However, further bitrate reduction is desirable (and often necessary), and this is accomplished by *lossy* compression. This consists essentially of discarding the less-important image information (from a psycho-visual standpoint); the tradeoff is a somewhat degraded image (the degree of degradation depending on the degree of compression), which, however, can differ from the pristine original in ways that are hardly perceptible to the viewer.<sup>24</sup> The mathematics involves methods with names like discrete cosine transform, variable quantization, and Huffman coding; but the bottom line is that these methods permit a large reduction in bitrate with a relatively small reduction in perceived image quality.<sup>25</sup>

<sup>24</sup> If such effects are noticeable, they are called *compression artifacts*; these are sometimes seen in over-compressed “jpeg” still photographs, as the patchy blocks or the “mosquito noise” around edges. Similar considerations apply to lossy audio compression, for example highly compressed “MP3” music files.

<sup>25</sup> The video-compression recipe currently used for all digital TV broad-

**¶ 31.** The tradeoff of image quality with bitrate is gradual, and somewhere in the process a decision is made as to the desired final bitrate.<sup>26</sup> A major constraint is imposed by the fact that both digital broadcast and digital cable television in the US is sent on channels that conform to the same 6 MHz channel spacing that has been used for television since the 1940s. In practice (see ¶35, below) it is possible to send about 20 million bits per second (Mbps) on an over-the-air digital-broadcast channel, and nearly double that on a digital cable or satellite channel. A typical compressed bitrate for over-the-air NTSC-quality (SDTV) digital video is about 4 Mbps; thus digital-broadcast television stations are able to combine up to 5 or so NTSC-quality programs on a channel. (Recall that a frequency “channel” is no longer a single “program,” because of multiplexing. More on this beginning at ¶33, below.) High-definition (HDTV) content requires nearly the full broadcast bitrate, so only one HDTV program can be broadcast on a channel. By contrast, cable or satellite systems, which are not constrained to MPEG-2 compression, are able to combine as many as eight HDTV programs onto one channel when using efficient H.264/MPEG-4 encoding. The second revision of the broadcast standard (ATSC 2.0) incorporates these more efficient codecs, as well as a host of transport and delivery enhancements that are aimed at mobile and interactive viewing, thus allowing over-the-air broadcasting to compete with services available on the internet.

**¶ 32.** It is worth admiring the impressive bitrate reductions that these methods are achieving: a simple calculation<sup>27</sup> shows that digitizing an HDTV program without any compression would produce a bitrate of roughly 1000 Mbps, whereas contemporary compression methods reduce this to a modest (and deliverable) 20 Mbps, a 50-fold reduction! And comparable reductions are routinely achieved with SDTV.

casting in the US is named “MPEG-2” and described in the Advanced Television Systems Committee documents A/53 and A/54 (see [www.atsc.org](http://www.atsc.org)). An improved set of compression methods is incorporated in the set of standards known as MPEG-4; these are widely used by the cable and direct broadcast satellite services, as well as for video streaming over the internet.

<sup>26</sup> Which is permitted to vary, as program content changes. This is known as *variable bitrate*, or VBR, as distinguished from *constant bitrate*, or CBR.

<sup>27</sup> Bitrate  $\approx 1080 \text{ lines} \times 1920 \text{ pixels/line} \times 30 \text{ frames/second} \times 16 \text{ bits/pixel} = 995,328,000 \text{ bits/second}$ .

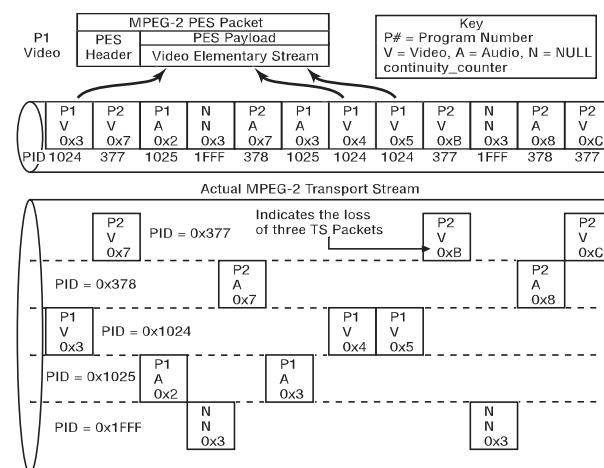
## I.5 Digital television: broadcast and cable delivery

**¶ 33.** Over-the-air digital television broadcast and digital cable television both use traditional frequency “channels,” upon which they put a stream of numbers (the compressed video described in ¶¶25–32, plus the associated digitized audio<sup>28</sup>), instead of the continuous analog waveform that was used in traditional NTSC television. Because of the economical bitrates produced by compression, there is adequate capacity on a single cable (or broadcast) channel frequency to accommodate several simultaneous programs. This is called multicasting, and permits up to four or five SDTV programs (a “multiplex”) to be carried on a single broadcast channel frequency.<sup>29</sup> One can think of these as subchannels.<sup>30</sup>

**¶ 34.** For either OTA digital broadcast or digital cable, the set-top box (STB) or equivalent hardware within the television set receives the multiple channel frequencies, each with its multiplex of programs. The STB or television knows the program assignments within each channel and is able to pull out the subchannel that the viewer selects, which it identifies by assigning a “virtual channel number.” That is what the viewer chooses – it is displayed on the STB and on the screen during selection. For example, the viewer might select HBO, which is assigned a virtual channel (e.g. 82) and which might actually be just one of ten subchannel programs carried on one digital cable channel frequency. The STB then captures the HBO stream, decrypts and decodes its MPEG-2 encoding, and converts it to displayable video for a television monitor (or flat screen, etc.).

**¶ 35.** In more detail, and in the language of digital television engineering, the delivery channel (digital broadcast or digital cable) is called the “transport stream,” which can be

thought of as a data pipe carrying some 20 Mbps (broadcast) or 38 Mbps (cable) in each frequency channel.<sup>31</sup> The ATSC specification dictates that the data put onto the transport stream must be broken into little *packets* of data, each of length 188 bytes and each belonging to an individual program. When multiple programs are sent on one transport stream, it is called a “multiprogram transport stream,” or MPTS; if a single program, it’s a “single-program transport stream” (SPTS). Repeating what was said earlier: a broadcaster can put five standard-definition programs (or one HD and one SD program) onto a single broadcast channel’s MPTS. The individual packets are identified by program, and they are interleaved in time (see Figure I.8).



**Figure I.8.** Multiple programs can be interleaved into one digital transport stream, as a “multiprogram transport stream.” Their individual video and audio packets are tagged with program identifiers (PIPs), by which they can be selected and reassembled (adapted from Figure 7.1 of ATSC Doc. A/54A, courtesy ATSC; readers are recommended to reference the current version of the standard or recommended practice available on the ATSC website).

<sup>28</sup> And also some “conditional access” (CA) information that enables legitimate subscribers to decrypt and view protected content; see ¶47.

<sup>29</sup> Cable delivery is more efficient, and permits as many as ten SDTV programs on a single channel.

<sup>30</sup> Because the cable (or broadcast channel) has a fixed total bitrate, the bitrates of the individual programs that are being multiplexed must be adjusted such that their total combined bitrate matches the channel capacity. This is called *bitrate grooming* and involves *null padding* (adding nulls, to increase a program’s bitrate), on-the-fly compression (to reduce a program’s bitrate), or even time shifting of program content (to prevent unlucky alignment of peak bitrates of the various programs). Digital television packets include “presentation time stamps,” so it’s OK to move things around a bit as they flow through the various digital pipes on their way to the television screen.

**¶ 36.** To put it another way, the several programs that will be put into a multiprogram transport stream (a “program multiplex”) are cut into short pieces (packets, about 40 µs in length for digital cable), tagged with unique identifiers (called PIDs, for “program identifiers”), and then interleaved with the pieces (packets) from the other programs

<sup>31</sup> The disparity has to do with the particular modulation schemes used: for broadcast it is called “8-VSB,” whereas cable uses the more efficient “256-QAM” (pronounced “quahm”), thereby exploiting cable’s better signal-carrying properties to carry roughly twice the information content.

that share the same transport stream. At the STB or television set the packets belonging to the selected program are identified (“filtered”) by looking for their PIDs, and then reassembled into a single program transport stream to be decoded and displayed. In Figure I.8 there are two programs (P1 and P2), each with video (V) and audio (A), with their corresponding PIDs (1024, 1025, 377, 378); they are shown as the interleaved multiprogram stream at the top, and as filtered into their respective single-program streams at bottom.

### I.6 Direct satellite television

¶ 37. Satellites provide an alternative to over-the-air or cable/fiber delivery of television programming, and satellite delivery is particularly welcome in areas not served by wired broadband connections. This is known variously as direct-to-home (DTH), direct broadcast satellite (DBS), or broadcast satellite services (BSS), and exploits (usually) satellites in the geostationary constellation, i.e., in the equatorial “Clarke” orbit<sup>32</sup> of radius 42,200 km, where a satellite’s period matches Earth’s rotational period of 23h56m4s.<sup>33</sup> As surprising as it may seem, a single satellite with a  $\sim 100$  W transmitter can deliver a half-dozen high-definition programs (or 30 “standard-definition” programs) simultaneously to small dish antennas on houses everywhere in the continental US. Typical direct-broadcast satellites are equipped with a dozen or more such “transponders,” and the receiving dish antennas use multiple “feeds” (as many as four, for DirecTV or DISH Network) to point at several satellites, making available many hundreds of television programs.

¶ 38. Early DBS systems used a 4 GHz (C-band) downlink, and required large dishes ( $>3$ m diameter) and expensive RF electronics. Contemporary systems operate around 12 GHz (Ku-band), with mass-produced oval receiving dishes (typically  $0.6\text{m} \times 0.8\text{m}$ ) that incorporate several low-noise RF amplifiers, each with local oscillator, downconverting mixer, and IF amplifier in an integrated LNB (low-noise block-downconverter plus feed) unit that sits at the focus of the parabolic dish. That’s an impressive amount of hardware for \$100.

<sup>32</sup> See “Extra-terrestrial relays – can rocket stations give world-wide radio coverage?” [Sir] Arthur C. Clarke, *Wireless World*, October 1945.

<sup>33</sup> Ha! You thought it was 24 hours. Not so – we’re in orbit around the sun, which adds an extra 4 minutes (1/365th part of a day) to Earth’s true rotational period (the *sideral* day) to arrive at the 24-hour *solar* day (the average time from solar noon to solar noon).

¶ 39. One may be puzzled by the peculiar orientation of home dish antennas – why are they sometimes pointed so low that they seem to be aimed at or below the horizon? There are two parts to the answer: first, the constellation of geostationary satellites spans an arc across the southern sky, populated worldwide with some 200 satellites; over the longitudes of the US alone there are some 35 satellites parked<sup>34</sup> in geostationary orbit. The line of satellites across the southern sky dips down to the horizon at its eastern and western ends.<sup>35</sup> Second (and technologically more interesting), the geometrical arrangement of the receiving dish is what’s called an “offset-feed paraboloid.” That is, the little conical feeds are offset below so they do not block the incoming signal. This makes the dish appear to be pointing some  $25^\circ$  lower than it is, thus the explanation for the apparently “subterranean” satellites. This peculiar arrangement is used to eliminate blockage of incoming signal by the feedhorns, and also to reduce the encroachment of thermal radio noise that is emitted by the surrounding environment – another bit of thoughtful design that enables successful direct satellite broadcasting.

¶ 40. The individual transponder RF channels are 27 MHz wide, with as many as 32 such transponders on a satellite, which (with guard bands) adds up to a total downlink bandwidth of  $\sim 1000$  MHz. Satellite transponders typically use phase-shift keying (QPSK, 8PSK), with downlink bitrates of 40 Mbps per channel, adequate to deliver a half-dozen or more 1080i HD programs with efficient H.264/MPEG-4 encoding. The transponder channels are divided into two sets, one of each circular polarization at each transponder frequency, so that the total downlink spectrum occupied by a given satellite is about 500 MHz. This is downconverted in the LNB to a pair of 500 MHz-wide IF bands, centered at 1.2 GHz and 1.9 GHz.

¶ 41. Rather more hardware resources are expended on the uplink, with steerable dishes of  $\sim 10$  m-diameter class (Figure I.9) illuminating the uplink receiving dishes at the satellite, with transmitted power of the order of hundreds of watts per transponder; dishes of that size produce diffraction-limited beam diameters of about  $0.15^\circ$ .

<sup>34</sup> Hardly “parked,” of course – they are in equatorial orbits, whizzing around the Earth at nearly 7000 miles per hour, to keep up with Earth’s rotation.

<sup>35</sup> You can get a good sense of the satellite arc with mobile apps like DishPointer or Satellite Locator: when you point your smartphone into the sky it shows the satellites as bright red circles (with their locations) on a red arc, superimposed on the camera view (with obstructing trees, etc.) seen by the mobile device.

( $\theta \approx \lambda/D$ ), small enough to prevent illumination of a satellite in an adjacent orbital slot, but requiring some active steering to maintain alignment on the desired satellite. Direct-broadcast providers such as DirecTV and DISH Network like to position their satellites close enough along the Clarke belt so that a single dish with several feeds can capture the downlinked signals from several of them; for example, both DirecTV and DISH Network offer an oval dish with three feeds, targeting satellites separated by  $10^\circ$  (currently at  $110^\circ\text{W}$ ,  $119^\circ\text{W}$ , and  $129^\circ\text{W}$  longitudes).



**Figure I.9.** A few of the several dozen transmitting antennas at EchoStar's Network's uplink facility in Cheyenne, Wyoming.

**¶ 42.** You hook the coax cables from the receiving dish to a set-top box (usually with DVR), similar to what's used for cable TV, but which is designed to power the LNB electronics up there on the roof, to select from the dish's several feeds and polarizations, and to receive the intermediate-frequency (IF) bands coming down from the dish on standard  $75\ \Omega$  video coax (quad-shielded RG-6, usually).<sup>36</sup> As with cable TV, the content providers control your available programming, via subscription-enabled decryption. Satellite systems are not bound by over-the-air terrestrial standards (e.g., MPEG-2 encoding), and they generally use more efficient schemes such as MPEG-4. When compared with cable or fiber television service, the inability to target millions of subscribers individually (along with the absence of a reverse uplink channel to the satellite) limits

the possibilities for interactive services such as video-on-demand.

### I.7 Digital video streaming over internet

**¶ 43.** With steady improvements in internet bandwidth (i.e., speed), it has become practical to deliver video (and associated audio) through the same internet infrastructure that serves personal computers and mobile devices (cell-phones, tablets) with their email, web browsing, and so on. Some familiar examples of internet video streaming are news services such as CNN, government services such as NASA TV, movie streaming services such as Netflix and Hulu, and peer-to-peer services such as Skype. Just as with broadcast or cable delivery, the ultimate payload is a stream of numbers that constitute the video and audio content, in some efficient compressed format that takes advantage of sophisticated encoding schemes with names like "H.264" (also known as AVC, for advanced video coding), one of the current favorites. At the viewer's end the digital content is decoded<sup>37</sup> to recover the video and audio. For some services (e.g., Skype) a dedicated client program must be installed, whereas for others (e.g., NASA TV or CNN) a standard internet browser (such as Internet Explorer, Safari, or Firefox) suffices.

**¶ 44.** Compared with delivery via broadcast or cable, however, internet delivery of the time-critical video data presents some unique challenges. That is because data sent over the internet (via "Internet Protocol," IP) travels as independent packets of data, each typically some thousand bytes in length, and each including headers that specify its destination (IP address). Packets make their way through the multiply-connected nodes of the internet, and (usually) reach their destination. But there is no reserved pathway for a stream of packets (as there is for the "circuit-switched" architecture of the telephone system), and no guarantee of speedy delivery, sequential delivery, error-free delivery, or, indeed, of delivery of any sort. Various schemes are used to circumvent these evident deficiencies, for example by requesting retransmission of missing packets (they are numbered) or of corrupted packets (they include error-revealing "checksums"). These work well, and a data file that is downloaded via "TCP/IP"<sup>38</sup> is effectively guaranteed to be bit perfect.

<sup>37</sup> Hence "codec," a contraction for coder-decoder, usually appended to the name of the compression scheme, e.g., "the H.264 codec."

<sup>38</sup> Transmission Control Protocol/Internet Protocol, a universally used standard for transfer of data that needs to be transported and reconstructed without error.

<sup>36</sup> There may be an intervening module or two, with names like "multi-switch" or "node," to deal with selection of polarizations and IF-band stacking.

¶ 45. Because of this disorganized “packet switching” of internet data, streaming video may suffer interruptions or intervals during which the average delivery rate is reduced. For this reason it is common for the receiving end to store (“buffer”) a few seconds of video beyond what is currently being displayed; and if the internet delivery speed is inadequate (as evidenced by buffer underruns), the sender will reduce the data rate (thus delivering lower-quality video). Contemporary broadcast-quality high-definition TV (known as 1080i, meaning that the picture consists of  $1080 \times 1920$  pixels, delivered 30 frames/sec) requires upward of 2 megabits/sec of download speed, available now in most homes with a broadband internet connection (cable, optical fiber, or telephone-line DSL<sup>39</sup>). That speed is well within the capability of wireless (Wi-Fi) connections, so high-definition video can be delivered to mobile devices such as laptop computers and tablets.

### I.8 Digital cable: premium services and conditional access

¶ 46. Cable television providers offer premium services, such as Showtime, HBO, and pay-per-view, for which the subscriber pays additional monthly fees. A subscriber whose subscription includes Showtime, for example, is able to view (and record, if the STB includes a DVR) programming delivered on Showtime channels. The cable provider needs a method to control each subscriber’s access to the full channel lineup. Although each customer has a cable coming into the home (and therefore one could imagine that different content is sent to each home), in fact the same signal is sent to all the homes within a neighborhood group, known as a “service group.”<sup>40</sup>

¶ 47. To limit access within the full suite of distributed programming, the cable provider includes *conditional access* information along with the video and audio. This is done by including “CA” packets, along with the usual V (video) and A (audio) packets that comprise one program within the multiprogram transport stream. The STB includes decryption hardware, and uses the CA packets to provide the key information to unlock the encryption that is imposed on the video and audio packets by the cable company.

#### I.8.1 Digital cable: video-on-demand

¶ 48. How is it possible to provide program material specifically for an individual subscriber, for example “on-demand” delivery of a previously broadcast program, or of a movie, on a cable network that serves an entire city?<sup>41</sup> Such services go by names like “video-on-demand” (VOD), and are made possible by the fact that the cable provider is able to tailor the actual signals carried on its cables that go to different groups of cable subscribers in an area.

¶ 49. A cable network in a metropolitan area is more complex than one might at first imagine: rather than a city-wide distribution of common program material, the network is organized into smaller groupings of *nodes* and *service groups*. A service group consists of segments of cable, typically running past no more than 500 homes, carrying identical material; those signals are fed into the cable at a “node” in the neighborhood, which in turn is fed via optical fiber from a more distant “hub” at which the provider inserts the group of channels that is to go to that particular service group.

¶ 50. The trick to providing individual on-demand material, then, is first to ensure that there are at least a few extra channels (in addition to the standard lineup) available to carry such content; and, second, to divide up the city into many service groups, so that those extra channels can carry a different suite of on-demand content to the different service groups. For example, suppose there are five channels available for on-demand material and that each can carry ten programs (as described earlier); if a service group includes 500 houses, of which 200 are cable subscribers, then the cable provider can satisfy 25% of those subscribers with simultaneous on-demand programming (because its five extra channels, each sending ten custom programs, deliver 50 simultaneous programs).

¶ 51. On-demand programming requires also a reverse channel for each subscriber, so that the subscriber can select programming, and also pause (or fast-forward or rewind) the material. However, these reverse channels need carry only a few simple commands up to the provider (as opposed to the high-bitrate video coming down in response), and are easily accommodated in an “interactive” uplink signaling band of the cable network.

<sup>39</sup> Digital subscriber line, technology for bidirectional digital transport over analog telephone lines.

<sup>40</sup> See ¶49, below, for more details.

<sup>41</sup> The term *unicast* is sometimes used to distinguish such individual delivery from *broadcast*.

### I.8.2 Digital cable: switched broadcast

¶ 52. Some cable providers use another service that exploits the flexibility inherent in the cable network's use of separate service groups, namely "switched broadcast." Switched broadcast delivers programs only when requested by viewers, on a service-group basis, as compared with delivering all programming to all customers. This allows the cable provider to offer more programming choices than could be carried simultaneously on the available number of channels.<sup>42</sup>

¶ 53. Although switched broadcast and VOD use similar methods to deliver their content to a subscriber, it is worth noting a difference: switched broadcast content is delivered whenever that program material is normally scheduled, and not at the whim of the subscriber; what distinguishes it from normal program material is that it is not put onto the service group's cable at all, if no one in that group has tuned to the program. Once a subscriber has tuned that program (causing it to be sent to that service group, on a particular subchannel), that same program is present on the cable serving any additional subscribers in that service group, on the same subchannel. By contrast, VOD is delivered on-demand, at the time requested by the subscriber; likewise, it can be paused, or fast-forwarded, etc. (tasks that are performed on-the-fly by the cable provider, not by the set-top box). That is possible because VOD content is being sent on a particular subchannel on the subscriber's service group only, and is available for the requesting subscriber only (via encryption and user-specific "entitlement control," in the case of protected content). The terms *narrowcast* and *unicast* are sometimes used to refer to service-group-specific cable delivery of these two varieties of content, namely scheduled ("linear") material, and interactive user-specific on-demand material, respectively.

### I.9 Recording digital television

¶ 54. The conversion of analog material (audio and video of the real world) into digital form is hard work – but it makes the task of *recording* straightforward. That is because a single program received at the STB is, in essence, just a stream of numbers, which can be filtered from the multiprogram stream, assembled in a temporary memory "buffer," and written to a hard-disk file just like any computer file. In that sense, a set-top box with DVR is simply a special-purpose computer, with the usual processor, hard drive, etc., and having additional hardware to do the

special STB tasks – receive the cable signal, generate the displayable output, take control commands from the infrared remote "clicker," and so on. A typical contemporary set-top box with DVR contains a dual-processor chip, 64 MB program memory, a 160 GB (minimum) hard drive, and various video-related additional hardware (input tuner, video memory, display and audio drivers, etc.). Along with buffering and storing the video and audio content, the STB controls the access, decryption, and re-encryption of the (protected) video content.

¶ 55. It is worth noting that any digital storage medium of adequate speed and capacity can be used to store digital video content; at the consumer level there are many "personal video recorders" (PVRs) that store programs on recordable DVDs or onto solid-state "flash" memory chips. There are also digital video tape recorders that can store both SDTV and HDTV onto a digital variant of VHS tape.

### I.10 Display technology

¶ 56. For over half a century television images were displayed with a cathode-ray tube (CRT), in which electrons emitted from a hot cathode and accelerated to potentials of kilovolts were deflected (usually magnetically) to paint a raster, at the video frame rate, on a phosphor surface coated on the interior of the viewing face of the evacuated tube. The intensity was modulated with a grid electrode near the cathode. Early CRTs were monochrome (black and white); color tubes used arrays (triads or stripes) of red, green, and blue phosphors, aligned with a metallic mask so that electrons from each of three electron-emitting cathodes (or, in Sony's Trinitron, steered from a single electron gun) struck only one color of phosphor.

¶ 57. Cathode-ray tube displays worked; but they were heavy (over 100 lbs for a TV with 32" display), bulky, and required fussy "convergence" adjustment to get the colors to track. Shortening the tube to reduce cabinet depth exacerbated the convergence and geometry problems. CRTs are now obsolete, replaced by several technologies, among them liquid-crystal displays (LCDs), plasma displays, and organic light-emitting diodes (OLEDs).

¶ 58. In an LCD a liquid-crystal layer is sandwiched between a pair of crossed optical polarizers; an applied electric field alters the polarizing properties of the layer, thus varying the optical transmission. In the classic display there's a uniform white rear illumination (from white LEDs, or from one or more cold-cathode fluorescent lamps – CCFLs – combined with diffusers and light pipe

<sup>42</sup> This is sometimes called *narrowcast*, versus broadcast or unicast.

material); an array of electrodes applies local electric fields to the image pixels, which are overlaid with red, green, and blue color filters. The array of liquid-crystal pixels act as video-rate dimmable shutters; all the light originates in the rear illuminator.

¶ 59. LCDs are dominant in computer displays and popular in televisions. They are thin (a centimeter or so) and bright. But they have somewhat limited dynamic range (or contrast ratio: ratio of maximum brightness to “maximum darkness”) and speed, and their color balance and black level degrades when viewed off-axis. There have been great improvements in speed and in off-axis performance, owing to methods with names like in-plane switching (IPS), fringe-field switching (FFS), and the like. And the dynamic range can be improved by using LED-array backlighting, which can be dimmed locally and rapidly, adapting to the light and dark areas of the displayed image.<sup>43</sup>

¶ 60. For the most realistic rendering of cinematic material, however, the plasma display is superior to the LCD. It consists of an array of tiny cells, in each of which a switchable gas discharge generates ultraviolet light that causes a spot of phosphor to glow. A high-definition display of  $1080 \times 1920$  pixels has three such cells for each pixel (one each with red, green, and blue phosphors) to generate the pixel’s overall emitted color. Unlike an LCD (which filters an underlying light source), the cells of a plasma display generate the emitted light directly. You can think of it as an array of 6 million little CRTs ( $1080 \times 1920 \times 3$ ), each one time-switched to achieve the required light intensity.

¶ 61. Plasma displays retain their color fidelity and contrast ratio regardless of viewing angle, and they have fast response. In larger sizes they are currently somewhat less expensive than LCDs. They are not as bright as LCDs, however, and a static pattern that is displayed for a long time can cause some image retention, or (in extreme cases) phosphor “burn.” Their contrast ratio is very good, but not infinite, because the gas discharge in every pixel must be sustained at a low level (i.e., it cannot be switched off entirely) so that it can be rapidly modulated.

¶ 62. LCD and Plasma have dominated display technology, but the future is likely to be ruled by OLED (organic LED), a direct-emitting array of tiny LEDs (either in three colors, or white LEDs with filters). These emerged first in

<sup>43</sup> For marketing purposes these LCDs are sometimes called “LED TVs.” Don’t be fooled: it’s an LCD, but with LED rear illumination replacing the CCFL. And it may or may not have local dimming – read the fine print.

small displays (e.g., cellphones and camera viewfinders), but by 2014 they had made the big time, with full 4K “Ultra HD” resolution ( $3840 \times 2160$ ), and screen sizes to 65” and beyond. Unlike earlier display technologies, OLEDs can be made flexible, and the current fad is *curved* TV screens (capable of 3D, if you’re interested in that). OLED is likely to be the ultimate winner, because of their very high dynamic range (1,000,000:1), wide viewing angle, low power consumption, elegant form factor (3 mm thick, lightweight, and nearly borderless), and potential manufacture by an inkjet-like process. These have been winning Best of Show awards, and they deserve it.<sup>44</sup>

¶ 63. Two other technologies that looked good, but have fallen onto hard times, are field-emission display (FED) and surface-conduction electron-emitter display (SED). Both involve an array of phosphor cells (like the plasma display), but with electron (rather than ultraviolet) excitation of the phosphor. An SED prototype from Canon generated great enthusiasm<sup>45</sup> at the 2006 CES, but subsequent patent disputes and economic realities took their toll. FED and SED may rise again – but don’t hold your breath.

### I.11 Video connections: analog (composite, component) and digital (HDMI/DVI, DisplayPort)

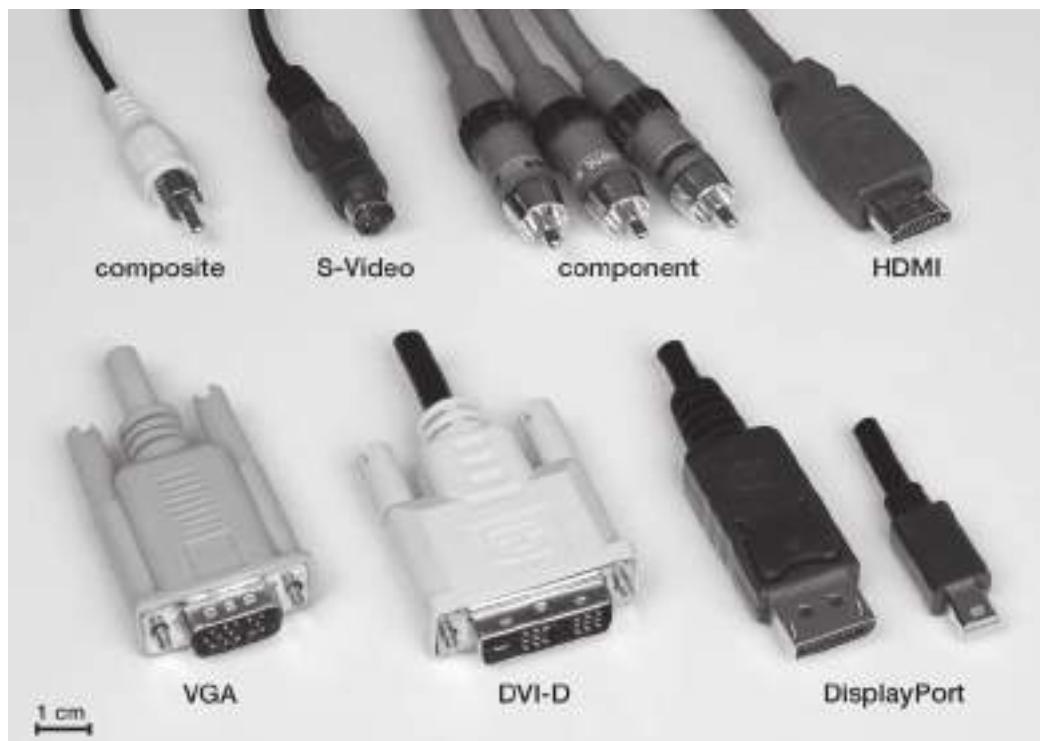
¶ 64. What are all those incompatible cables, anyway? To add to the analog–digital confusion, the *consumer* market (of large-screen TVs and flat panels, etc.) and the *computer* market (of LCD monitors) have gone their (mostly) separate ways. Here’s a quick rundown of the most-used connections; their connectors are pictured in Figure I.10.

¶ 65. In the *consumer television* world there are four types of connections (and connectors), the first two of which are nearly extinct.

**Composite video.** Low quality standard-definition (SDTV: 480i) analog video, recognizable by the yellow RCA-type connector (“phono jack”), usually bundled with an audio pair (red = right, white = left). The single yellow line carries bandwidth-limited luminance

<sup>44</sup> The reviewers are *gushing*: “The best picture I’ve seen on any TV, ever.” (CNET), “The best-looking TV I’ve ever seen. Ever.” (Digital Trends), and “The best direct-view display – of any size, at any price – we’ve ever laid eyes on.” (HDTVtext, UK).

<sup>45</sup> A breathless review in SlashGear (19Oct2006) proclaimed “SED-TV is something that no amount of words can describe. It is something that must be SEEN to be believed; literally.” and “SED-TV is the future of digital image displays; it’s as simple as that.”



**Figure I.10.** Cable connectors for computer monitor and video signals. Top row: TV video connectors, from the legacy analog composite to contemporary HDMI high-definition digital video. Bottom row: computer monitor connectors, from the legacy analog VGA to the popular digital DVI and newer DisplayPort (standard and mini connectors).

(“luma,” abbreviated Y, essentially the grayscale image) and chrominance (“chroma,” abbreviated C, the color difference signal pair that is modulated on a 3.58 MHz color subcarrier), along with line and frame synchronization pulses; it is sometimes called CVBS (composite video, blanking, and sync). Don’t use such a connection unless there’s nothing else available!

**S-video.** Somewhat better SDTV analog video, recognizable by the 4-pin miniDIN connector with fragile pins. It separates the luma (+ sync) and chroma signals, retaining more bandwidth. Avoid this one, too, unless you like fuzzy pictures.

**Component video.** Now we’re talking! This analog format uses three  $75\ \Omega$  coax lines, with (usually) RCA-type connectors (or, occasionally, BNCs) that are colored red, green, and blue, and that can carry full-bandwidth HDTV. The colors are misleading: the green line carries luminance (+ sync), while blue and red carry color *difference* signals. That is, the GBR-colored connectors denote “YPbPr,” where Y is luminance (red+blue+green), Pb is blue-minus-Y, and Pr is red-minus-Y. As with composite

and S-video, component video carries only video; the audio needs its own cables. Component video does not know, or care, about things like content protection; for that reason it is not embraced by content providers, who may prevent full-resolution HDTV ( $1080 \times 1920$ ) output on the component jacks (e.g., on a Blu-Ray player).

**HDMI.** A purely digital format, whose initials stand for high-definition multimedia interface. HDMI is the digital alternative to component video. It is recognizable by the flat 19-pin USB-like connector (sadly without any required latching mechanism), and electrically it is equivalent to the DVI-I computer-monitor format (see below). It carries both audio (up to eight channels, digitized to 24 bits, up to 192 ksp) and video (digitized at 8–16 bits per component, at rates adequate for full HDTV (“1080p,”  $1080 \times 1920$  progressive at 60 Hz; HDMI versions 1.4 and later support full “4K,”  $4096 \times 2160$ , with 60 fps progressive from version 2.0)). The video data is a digitized version of analog video: uncompressed numerical data representing the amplitude of the color

components, sent over four twisted pairs (R, G, B, clock). HDMI supports digital content protection (HDCP, high-bandwidth digital content protection), a protocol by which an HD video source authenticates a display device before sending (encrypted) data, so you are allowed to view the full-resolution video. HDCP seems to work, most of the time anyway (though you may get annoying messages and glitches).

¶ 66. In the *computer monitor* world there are three types of connections (and connectors) in wide use.

**Analog VGA.** Legacy analog link, recognizable by the 15-pin D-type connector with locking screws. VGA (for video graphics array) carries separate RGB analog signals, plus Hsync and Vsync (thus “RGBHV”). In contemporary implementations it also has an I<sup>2</sup>C channel for monitor identification and control. VGA will work up to resolutions of 1600×1200 or so (there’s no specified limit, but you’ll see smearing going rightward of sharp features when pushing the resolution), but many monitors have abandoned VGA altogether in favor of the digital formats: DVI and DisplayPort.

**DVI.** Currently the standard digital interface, recognizable by the 29-pin (maximum) connector with locking screws, looking somewhat like a “VGA on steroids.” It is electrically similar to the more compact and inexpensive HDMI, above, that evolved from it for the consumer tele-

vision market, and that includes audio (thus a single connection from a cable box or Blu-Ray player to the television monitor). It comes in several variants, all using the same connector (in which some pins may not be loaded): DVI-D is video-only, and comes in single-link and dual-link varieties (the latter needed for resolutions greater than 1920×1200 at 60 Hz, for example the 2560×1600 at 60 Hz used in 30" displays); DVI-A is analog video only (for compatibility with analog monitors); and DVI-I (“integrated”) has both digital (single- or dual-link) and analog video. DVI, like VGA, carries no audio.

**DisplayPort.** Newer standard, intended to supersede DVI; it uses a 20-pin USB-like connector with latching mechanism, and supports very high data rates (up to 4.3 Gbps on each of four differential pairs, thus 17.3 Gbps). It departs from the “digitized-raster” scheme of DVI/HDMI, using instead a packetized data protocol; but it’s got enough bandwidth to handle the full video bandwidth of dual-link DVI (which can be converted to DisplayPort protocol). It supports up to 16 bits per color, and 8-channel audio at 24 bits and 192 ksp. It includes provision for fiber optics (instead of copper) for long cable runs (to 50 m or more), and it supports both existing 56-bit HDCP and its own DPCP (DisplayPort content protection, with stronger 128-bit AES encryption). The current revision of DisplayPort can handle 4K 60 fps progressive with ease.

# SPICE PRIMER: GETTING STARTED WITH FREE ICAP/4 DEMO

## APPENDIX

J

Free SPICE is easy, and fun. It's available for several platforms (Macs, Linux, PCs). Here's how to get started in Windows.

### J.1 Setting up ICAP SPICE

1. Download "ICAP/4Windows Demo" zip file (ICAP4Demo.zip) from intusoft.com.
2. Extract zip file to some temporary folder.
3. In that folder, double click (abbreviated hereafter as "cc") on setup.exe.
4. Launch ICAP/4 (you may have to be admin mode):  
(Programs → ICAP\_4 Demo → Start ICAPS).

### J.2 Entering a Diagram

- Add standard components with typing
  - R resistor
  - C capacitor
  - L inductor
  - o ground (the letter "oh," not the number)
  - v signal or voltage source
  - Y voltage test point.  
(These are case-insensitive: r = R, etc.)
- For a library component, Parts → Parts Browser,
  - browse by type, and subtype, or use Find,
  - then Place.
- Moving and orienting:
  - highlight component by click-hold, then can drag;
  - when highlighted, type + to rotate (90°CW), or – to flip 180°.
- Editing a component label:
  - highlight label by click-hold, drag to relocate;
  - when label highlighted, cc on label to edit values.
    - \* Note units multipliers: f, p, n, u, m, k, meg (not M).
    - \* Note that SPICE only needs the units *multiplier*, not the units themselves. For example, for a capacitor "1u" is OK for SPICE; but you can type "1.0uF" for clarity on the schematic, which SPICE will treat the same.

\* You can cc on the component, instead, to bring up the value dialog.

– You can add parameters by entering the value and clicking Apply; this is necessary, e.g., for a voltage source (add dc or ac, give amplitude). Or you can click >>Add>>, which also adds the parameter to the schematic drawing.

- Wires:

- type w, then drag a wire between points;
  - you can drag multiple separate wires; ESC when done.

- Settings:

- set "Rubberbanding" in Options, to maintain connections when dragging components.

### J.3 Running a simulation

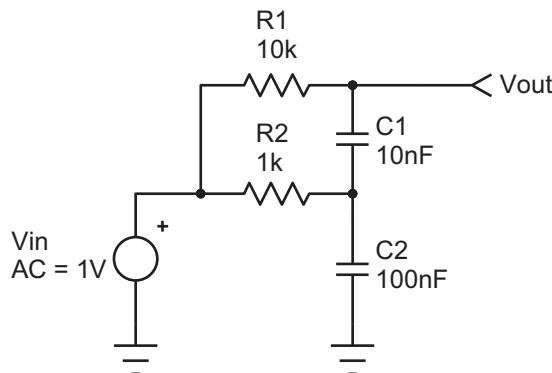
Can a passive circuit consisting only of resistors and capacitors have *voltage gain*? The surprising answer is yes. Let's use this simple example to illustrate an ICAP/4 SPICE entry and simulation.

#### J.3.1 Schematic entry

We launch ICAPS, and place two resistors (type the letter "r"), two capacitors (type the letter "c"), and a voltage source (type the letter "v") with the bottom terminal grounded (type the letter "o" right after typing "v"). If the pieces are too small to see on your screen, you can resize them by using Options→Zoom (or the function keys F6, F7, and F8). We drag them around to connect them as shown in Figure J.1, and add an output-voltage test point (type the letter "y" and change the label to Vout). We assign values by double clicking on each component, then entering the value in the value field (where "???" appears initially); note that pressing the Enter key after each value moves you along to another field in the dialog – you have to click OK to complete the operation.

For the capacitors we could have used microfarads (e.g., 0.01uF instead of 10nF); note that the "F" can be omitted, but including it makes the schematic marking clearer. For the voltage source we enter the value 1V (i.e., 1 volt) in the

“AC” field, and then we click on the ADD button so that the value shows on the schematic (SPICE uses the value, even if we do not ADD it; but it’s nice to show it on the circuit diagram). As before, the unit symbol “V” can be omitted.<sup>1</sup>



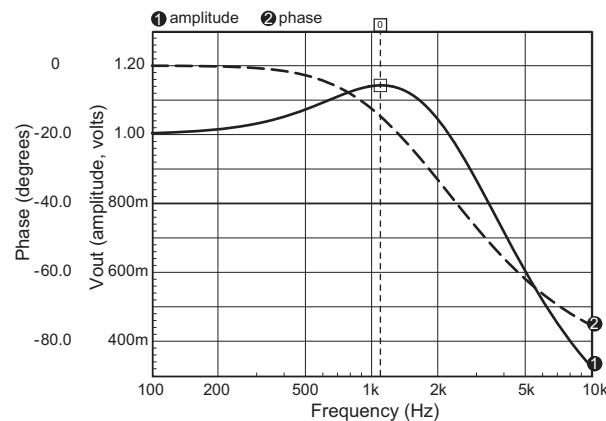
**Figure J.1.** An *RC* circuit with voltage gain! Schematic entry in Intusoft ICAP/4 is swift, taking no more than a couple of minutes. Printing the captured schematic to “Adobe PDF” produced this figure.

### J.3.2 Simulation: frequency sweep

Now for the fun! We begin by saving the project to a folder of our choosing, using the File→Save As menu. Then we set up the simulation by Actions→Simulation setup→Edit (or by clicking on the button with a pencil over a wavy line). Choose and click “AC Analysis,” then enter 20 points per octave, start at 100 Hz, end at 10 kHz (type 10k in the box), and click OK. Then click DONE. Now run the simulation by clicking on the little running person (or Actions→Simulate). This launches the Spice engine, with a small window (“IsSpice4”) showing status and errors, and (if successful) a little graph of the result. Running the simulation also launches the “IntuScope” display program (which you can launch manually by clicking on the ‘scope icon next to the running person, or by Actions→Scope). At this point there are multiple windows piled up, which you should resize and drag to convenient locations on the screen and then do Options→Save Preferences to stick them down. Now click the “Test Pts Only” box in the “Add Waveform” window that belongs to IntuScope, then highlight Vout<sup>2</sup> and click “Add.” Voilà – this produces a nice

plot of  $V_{\text{out}}$  versus frequency (see Figure J.2), with default axis labels, grids, and scaling. You can change the axis labels by double clicking on the label; and you can fiddle with the scaling in the “Scaling” window.

To add a plot of phase shift versus frequency, highlight “ph\_Vout” in the Add Waveform window, click on the box labeled “With Like Traces” (unless you want a separate graph), and click “Add.” We’ve done that to make Figure J.2, where we’ve also used the cursor tools on the IntuScope window to mark the point of maximum voltage gain.



**Figure J.2.** “Scope” output of the *RC* gain circuit, showing amplitude and phase (“AC Analysis”) for the circuit of Figure J.1. The cursor indicates the peak of the voltage gain,  $V_{\text{out}}/V_{\text{in}}=1.142$  at  $f=1.096 \text{ kHz}$ . The original IntuScope output plotted phase in blue, which we’ve converted to a dashed black line; we enlarged the text, also, for readability.

### J.3.3 Simulation: input and output waveforms

Do we really believe this gain-versus-frequency plot? Everyone knows that a simple *RC* network just mushes things out – shifted phase and reduced amplitude. We won’t believe this thing has voltage gain unless SPICE shows us the actual input and output signal waveforms for a sinewave near the frequency of maximum voltage gain (and maybe not even then!).

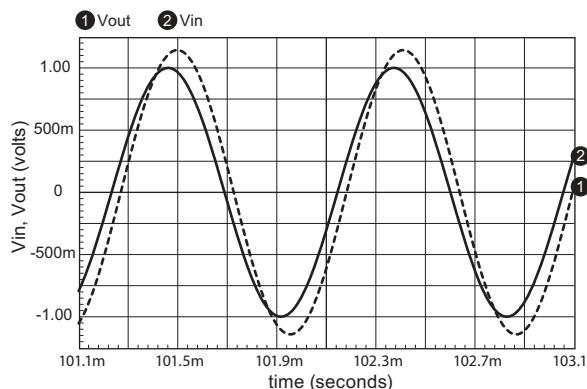
Here’s how you do it: first, add another voltage test point at the input (“y” again), and change the label to Vin. Then double click on the input voltage source and click on the field labeled “Tran Generators” (SPICE calls any old waveform, including a sinewave, a “transient”). Select SIN, set the peak amplitude to 1, the offset to 0, and the frequency to 1.096 kHz. Click OK, and OK. Then go to Simulation Setup (pencil over wavy line), deselect AC Analysis, and select Transient. In the Transient Analysis

<sup>1</sup> For SPICE’s ac analysis, the frequency-sweep gain computation is performed with infinitesimal small-signal amplitudes, which are then normalized to the signal amplitude you specified. Here, for example, SPICE uses a signal amplitude much smaller than 1 V.

<sup>2</sup> Or, if you have not renamed the output to Vout, select from among the signals with names like v2, etc., in the Y Axis output list.

window that pops up, set Data Step Time to 1us, Total Analysis Time to 105ms, and Time to Start Recording Data at 100ms (this gives the circuit time to settle into its steady state). Leave the Maximum Time Step field blank (i.e., accept default value). Click the runner icon to run the simulation. At this point the IsSpice4 window should show a wiggly trace. Now go to the IntuScope window, clear the old graph by pressing the DEL key for each selected trace (you can save it, if you want, or start a new plot with File→New Graph), and Add “vin” and “vout” in the Add Waveform subwindow of IntuScope, with the “With Like Traces” box checked as before.

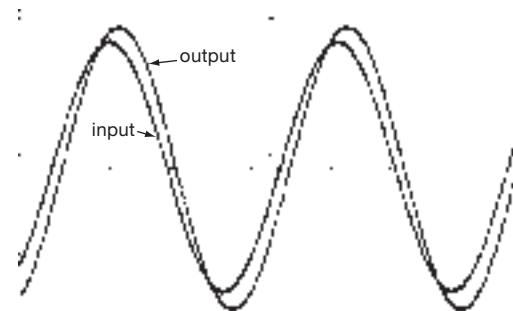
Not bad...but the waveforms look a bit jagged. If you zoom in (drag the box boundaries in the “Scaling” window) you’ll see that there are only about 20 points plotted per cycle. We asked for 1  $\mu$ s steps, but we’re getting output at  $\sim 50 \mu$ s intervals! That’s because we didn’t constrain the Maximum Time Step in the Transient Analysis window. Go back now (Simulation Setup→Transient) and set Maximum Time Step to 1us. Then click on the running person (the simulation will take slightly longer, maybe a full second this time); then clear the old graph and Add the two waveforms in IntuScope (output first, to set scale, then input). You should see silky-smooth plots, as in Figure J.3. If you still have doubts about an *RC* circuit with voltage gain, you can hook up the circuit and check it out on a real-live oscilloscope, as we did (Figure J.4). Evidently this *RC* voltage-gain stuff is real!



**Figure J.3.** “Scope” output of the *RC* gain circuit, showing sinewave response (“Transient Analysis”) at the frequency of maximum gain for the circuit of Figure J.1. The original IntuScope output plotted  $V_{out}$  in blue, which we’ve converted to a dashed black line; we enlarged the text, also, for readability, and adjusted the axis scales to match the scope trace.

#### J.4 Some final points

The ICAP/4 demo program includes ten “tutorial movies” (in the Help menu), which are mercifully short (about



**Figure J.4.** A real ‘scope (Tek TDS3044B) validates that crazy SPICE stuff! Horizontal; 200  $\mu$ s/div; Vertical; 0.5 V/div.

a minute each) and quite helpful for the beginner. Take ten minutes to view them! You’ll learn, for example, that you can create many signal sources by going directly to Parts→Parts Browser→!Generators. You’ll also learn some tricks to simplify schematic entry (copying, zooming), and ways to make the parts labels more readable (“tall,” “wide,” and “split” options). And by experimenting a bit you can discover that both the schematic and the simulation graphs can be copied (“Print”) to the clipboard and then pasted into Word or Wordpad, where they exist as scalable graphics. If you Print to Adobe PDF you’ll get a scalable graphic that can be cleaned up in Adobe Illustrator.

We have been using the Intusoft “ICAP/4” full version of SPICE now for many years; it provides a convenient Windows-based schematic entry and simulation environment, which we use to explore circuit configurations. Typically we begin with a starting circuit design, making explanatory notations in a Word document as we go. We copy and paste the schematics and simulation graphics into the running text, making a “lab notebook” of sorts; see next.

#### J.5 A detailed example: exploring amplifier distortion

In Chapter 2x (“BJT Amplifier Distortion: a SPICE exploration”) we do some serious work with SPICE, in the form of an exploration of distortion in discrete BJT amplifier designs.<sup>3</sup> The amplifier stuff is lots of fun; and there are some nice SPICE tricks there revealed. Take a look!

<sup>3</sup> When we do this in real life, we like to paste circuit diagrams and screen shots of simulated performance into a Word document, annotating as we go. We’ve kept that spirit alive in the discussion in Chapter 2x, albeit in the form of a L<sup>A</sup>T<sub>E</sub>X conversion of the original Word document.

### J.6 Expanding the parts database

You can freely add SPICE models to the ICAP/4 parts library. An easy way to start the process is

to download the folder AoE-PR from our website ([www.artofelectronics.com](http://www.artofelectronics.com)), and place it under the default PR folder. Then do File—Update Part Database.

# “WHERE DO I GO TO BUY ELECTRONIC GOODIES?”

## APPENDIX

K

Good question! Here are some hints, from our experiences.

### I. Mail order and online

**Digi-Key Corp** (Thief River Falls, MN: 1-800-digikey). We used to say “*Get their catalog!!*” – but, sadly, they’ve abandoned paper, replacing it with an impressive search capability. You can get everything here, even in small quantities, with fast delivery. It’s often worth designing with their webpage open in front of you. Online ordering and stock/price checking: [www.digikey.com](http://www.digikey.com).

**Mouser Electronics** ([www.mouser.com](http://www.mouser.com)). Broad stocking distributor, with service comparable to Digi-Key’s, and willingness to ship small quantities. Good selection of precision passives; and they are still printing a comprehensive paper catalog.

**Newark Electronics + Farnell** (1-800-2-newark; [www.newark.com](http://www.newark.com)). Broadest stocking distributor, with service comparable to Digi-Key’s and good selection of tools; paper catalog still in print.

**“Stocking Distributors”**. These are the standard distribution channel for quantity buying; names like Allied (still publishing a paper catalog), Arrow, Avnet, FAI, Heilind, Insight, Pioneer, Wyle. Substantial minimum quantities – not generally useful for prototyping or small production.

**Manufacturers’ Direct**. Many semiconductor manufacturers (Analog Devices, TI, Maxim,...) will not only send free samples with the slightest provocation, they will also sell in small quantities via credit card; check out Mini-Circuits for RF components, and Coilcraft for inductors, transformers, and RF filters.

**Oddballs**. Marlin P Jones, Jameco, B&D, Herbach & Rademan, Omnitron, ABRA, All Electronics; ephemeral collection of “surplus” stuff, some real bargains.

**eBay** ([www.ebay.com](http://www.ebay.com)). If you haven’t been here, you’ve probably just arrived from Mars. LOTS of stuff, literally millions of items, an online auction. You can get plenty of

electronic stuff, but CAVEAT BIGTIME EMPTOR. Feed-back Forum helps.

**Alibaba** Small Pacific-rim companies that have stocks of obsolete components are easily found on Alibaba. You can place a quote request for a part number, and you’ll get dozens of useful reasonably-priced suppliers.

### II. Indexes and Locators

**Octopart**, **FindChips**, **NetComponents** ([octopart.com](http://octopart.com), [findchips.com](http://findchips.com), [netcomponents.com](http://netcomponents.com)). Give it a part number and it searches dozens of distributors, returning (sometimes unreliable) information on availability and pricing.

**WhoMakesIt** ([www.whomakesit.com](http://www.whomakesit.com)). A bit like the EEM catalog, helpful if you know the category of stuff you want, but not the manufacturer, etc.

**Google** ([www.google.com](http://www.google.com)). Our standard “portal,” reads your mind and vectors you to the goodstuff. Helpful, sometimes, in finding parts and equipment manufacturers and vendors.

### III. Local

Sometimes it’s nice to shop in person; here are the sorts of places to go.

**Radio Shack** ([www.radioshack.com](http://www.radioshack.com)). They call themselves “America’s Electronic Supermarket”; we’d call them “America’s Electronic Convenience Store.” Their stores are everywhere, and they stock (pretty reliably) an idiosyncratic collection of parts and supplies, of uncertain quality or duration. However, in the changing marketplace of consumer electronics, their future path is unclear.

**Electronics Flea Markets**. Also known as “swap meets,” perhaps somewhat in decline; two legendary meets are on opposite coasts: De Anza College (Cupertino, every second Saturday, March–October), and MIT (Cambridge, every third Sunday, April–October). What meets

are three cultures (electronics, computers, hams); haggling is mandatory; caveat very emptor.

**Electronics Surplus Supply Stores.** These are incredibly cool! Several well-known haunts are Halted ([www.halted.com](http://www.halted.com); officially “HSC Electronics Supply”), in Santa Clara, Sacramento, and Santa Rosa; and Murphy’s Surplus Warehouse ([www.murphyjunk.bizland.com](http://www.murphyjunk.bizland.com)) in El Cajon.

#### IV. Miscellaneous

**Obsolete ICs.** Your best place to start is Rochester

Electronics ([www.rocelec.com](http://www.rocelec.com)), a wonderful place that apparently buys up inventories of ICs being discontinued. Jameco also has a lot of obsolete parts. Freetradezone has broker lists for obsolete parts. Also try Interfet ([www.interfet.com](http://www.interfet.com)), a manufacturer of small-signal FETs, including ones that the big guys have abandoned.

**PC Board Manufacture.** We like a place called Advanced Circuits, [www.4pcb.com](http://www.4pcb.com); you can get online quotes, and they do a good job and deliver pronto. Another inexpensive and fast PC house is Alberta Printed Circuits in Canada ([www.apcircuits.com](http://www.apcircuits.com)).

# WORKBENCH INSTRUMENTS AND TOOLS

## APPENDIX



Here are some electronic favorites, stuff that we really like using when we design and build electronic circuits. It's best to check current catalogs and websites – much of this sort of equipment becomes obsolete with terrifying speed.

### Soldering iron

Metcal MX-500 (not cheap; it will change your life!)  
Weller WSL (less expensive, variable temp)

### Desoldering station

Pace MBT201-SD

### Surface-mount prototype and rework

Zephyrtronics Airbath and Airpencil  
Metcal offers SMT "tweezers" and other goodies

### Bench DMM

Keithley 2100 (6½-digit, with USB control & readout)  
Agilent 34410A (6½-digit, with LAN & USB control & readout)

### Pocket DMM

Amprobe 37XR-A (cheap, good enough)  
Fluke 289 (not cheap, very good)  
Agilent U1252A/53A (everyone's favorite lately)  
"Smart Tweezers" (auto-ranging SMT tweezer-style meter, neat!)

### Triple bench power supply

HP E3630A (excellent performance, reasonable price)

### High-voltage bench power supply

SRS PS300-series (single and split, to 20 kV)

### Device programmer

(if needed; JTAG in-ckt is taking over, pods from mfgrs)

BP Microsystems 1610 (universal and reliable; lifetime free algorithm updates)

### LCR meter

HP 4263B (a cheaper one is the SRS model SR720)

### Analog oscilloscope

B&K Precision and Hameg still offer some models, to 200 MHz bandwidth

### Digital oscilloscope

Tek DPO2024B (cheap "lunchbox"); DPO/MSO 3k-, 4k-, 5k-series

Agilent DSO/MSO 5k-, 6k-, and 7k-series (Agilent's answer to the "lunchbox")

Lecroy WaveRunner, WaveSurfer, WaveJet series (lot of models, do your homework)

### Arbitrary function generator

Tek AFG3000-series (single and dual channel, to 240MHz and 2Gs/s)

### Low-distortion function generator

SRS DS360 (0.01 Hz–200 kHz, 0.001% distortion)

### RF and microwave synthesizer

Agilent N9310A  
SRS SG380-series (2, 4, and 6 GHz models, low phase noise, see §13.13.6B)

### Low frequency spectrum analyzer

SRS model SR785

### RF spectrum analyzer

Agilent ESA series (depends on frequency range, and \$\$)

### Source measure unit

Agilent B2900-series (single and dual channel)  
Keithley 2600-series (single and dual channel)

### Precision time and frequency standard

Symmetricom 4411A (uses GPS constellation)

### Engineering software

Altium System Designer, OrCad, or Eagle (for schematic capture and layout; includes simulation)

Xilinx WebPack (for PLD and FPGA design), and analogous tools from Altera, Lattice, Actel, etc.

ICAP/4, LTspice (\$0!), MicroCap 9, MMICAD, PSpice (for simulation)

FilterCAD (\$0!, from LTC), FilterPro (\$0!, from TI) (simple analog filter design)

MathCAD, MATLAB, Mathematica (engineering/math worksheets)

LabVIEW™ (virtual instruments; control of real instruments)

# CATALOGS, MAGAZINES, DATABOOKS

## APPENDIX

M

Here are some recommendations for data books, catalogs, and magazines; some have become on-line services only (e.g., IC Master). You will need some of these if you want to practice electronic design.

### “Master” catalogs

EEM (Electronic Engineer’s Master) *lists all categories of electronic stuff*

IC Master *ditto, for ICs*

Octopart *excellent on-line part (and datasheet) finder*

### Parts and equipment catalogs

Digi-Key Corp (a distributor), *wide range, quick delivery*

Mouser Electronics (a distributor), *wide range, quick delivery*

Newark/Farnell Electronics (a distributor), *best broad-line catalog*

Allied Electronics (a distributor), *similar to Newark*

TechniTool (a distributor), *measurement, assembly, and test tools*

Stanley Supply & Services, (a distributor) *equipment and supplies for assembly*

Keysight Technologies (Agilent, HP), *broadest manufacturer of test and measurement equipment*

Tektronix, *test and measurement equipment, especially scopes*

Fluke, *test and measurement equipment*

Stanford Research Systems (SRS) *test and measurement equipment; great documentation and app notes*

### Magazines and tabloids

EDN, *keeps you up to date on new products and methods*

Electronic Design, *ditto; a bit thin, lately*

Electronic Products, *ditto, emphasis on products only*

EE Times, *tabloid format*

Computer Applications Journal, *microcontroller project mania*

Circuit Cellar Ink, *emphasis on microcontrollers*  
Nuts and Volts, *quirky do-it-yourself, tabloid format*  
Make Magazine, *cool stuff you can build*

**Data books; short-form/design guides** A starter selection, with product emphasis.

Altera: cPLD, FPGA

Analog Devices: all linear functions; converters; DSP

Atmel: PLDs and microcontrollers

Avago (←Agilent←HP←Avantek): opto, rf

Cirrus/Apex: power op-amps etc.

Cypress: memory, processors

Diodes Inc/Zetex: discretes, etc.

Fairchild: discretes, linear, digital

Freescale: processors, DSP, automotive

IDT, Micron, Samsung: memory

Infineon: discrete, power, processors, RF

Intel: microprocessors and microcontrollers

Linear Technology Corp: all linear functions

M/ACom: RF and microwave

Maxim/Dallas: linear, digital,  $\mu$ C

MiniCircuits: RF – broad line, inexpensive

NXP (Philips): logic, microcontroller

ON Semi/Sanyo: linear, logic, discrete

Renesas: discrete, memory, processors

TI/National/Burr-Brown: linear, logic, opto, power, processors, DSP

Xilinx: cPLD, FPGA

Datasheets (in \*.pdf format) are available online from nearly every semiconductor manufacturer’s website; try, e.g., [www.analog.com](http://www.analog.com), [www.maxim-ic.com](http://www.maxim-ic.com), [www.linear.com](http://www.linear.com), [www.ti.com](http://www.ti.com), etc.

Another way to find datasheets, prices, and availability is via a search website like [octopart.com](http://octopart.com) or [findchips.com](http://findchips.com); these display price and inventory (and sometimes datasheets), and link you through to the stocking distributor.

# FURTHER READING AND REFERENCES

## APPENDIX

N

### General

- Ashby, D., ed., *Circuit Design: Know It All*. Newnes (2008). A collection of fascinating electronics engineering stuff from 14 acclaimed authors.
- Camenzind, H., *Much Ado About Almost Nothing, Man's Encounter with the Electron*. Booklocker.com (2007). Fascinating stories about electronics, by the famed designer of the 555.
- Dobkin, B. and Williams, J., eds., *Analog Circuit Design: A Tutorial Guide to Applications and Solutions*. Newnes (2011). Excellent selection of informative and well-written application notes from Linear Technology. Lively and entertaining, too.
- Dunn, P. C., *Gateways into Electronics*. Wiley (2000). Fascinating physics-based approach to electronics; deep coverage of critical areas.
- Jones, R. V., *Instruments and Experiences: Papers on Measurement and Instrument Design*. Wiley (1988). Classic on instrument design, based on Jones' papers.
- Lee, T. H., *The Design of CMOS Radio-Frequency Integrated Circuits*. Cambridge University Press (2nd ed., 2003). From the originator of gigahertz CMOS comes this delightful volume, covering much more than its humble title suggests. Terrific introductory chapter on the history of radio.
- Pease, R. A., *Troubleshooting Analog Circuits*. Butterworth-Heinemann (1991). The curmudgeon-in-chief reveals his tricks.
- Purcell, E. M., and Morin, D. J., *Electricity and Magnetism*. Cambridge University Press (2013). Excellent textbook on electromagnetic theory. Relevant sections on electrical conduction and analysis of ac circuits with complex numbers.
- Scherz, P. and Monk, S., *Practical Electronics for Inventors*. McGraw-Hill (3rd ed., 2013). The title says it all.
- Sedra, A. S. and Smith, K. C., *Microelectronic Circuits*. Oxford University Press (6th edition, 2009). Popular classic engineering text.
- Senturia, S. D., and Wedlock, B. D., *Electronic Circuits*

and Applications

Wiley (1975). Good introductory

engineering textbook.

Sheingold, D. H., ed., *Nonlinear Circuits Handbook*. Analog Devices (1976). Highly recommended.

Sheingold, D. H., ed., *The Best of Analog Dialog, 1967 to 1991*. Analog Devices (1991). Outstanding collection of analog-engineering techniques.

Termaan, F. E., *Radio Engineers' Handbook*. McGraw-Hill (1943). Three score and ten years later it continues to amaze, with excellent sections on passive circuit elements and other basic engineering.

Tietze, U., and Schenk, Ch., *Electronic circuits: Handbook for Design and Applications*. Springer-Verlag (2nd edition, 2008). Spectacular all-around reference.

Williams, J., ed., *Analog Circuit Design: Art, Science, and Personalities*. Butterworth-Heinemann (1991). Idiosyncratic collection of wisdom from 22 analog gurus.

Williams, J., ed., *The Art and Science of Analog Circuit Design*. Butterworth-Heinemann (1998). The sequel: 16 analog gurus dispense yet more wisdom.

### Handbooks

Fink, D. G., and Beaty, H. W., eds., *Standard Handbook for Electrical Engineers*. New York: McGraw-Hill (16th ed., 2012). Tutorial articles on electrical engineering topics.

Jordan, E., ed., *Reference Data for Engineers: Radio, Electronics, Computer, and Communications*. Howard W. Sams & Co. (9th ed., 2001). General-purpose engineering data.

"Temperature Measurement Handbook." Stamford, CT: Omega Engineering Corp. (revised annually). Thermocouples, thermistors, pyrometers, resistance thermometers.

### BJTs and FETs

Camenzind, H., *Designing Analog Chips*. Virtualbookworm.com and available online (2005). Inspiring

- book by a real world analog IC designer; includes the story of his design of the 555 at Signetics (now NXP). Ebers, J. J., and Moll, J. L., "Large-signal behavior of junction transistors." *Proc. I.R.E.* **42**:1761–1772 (1954). The Ebers–Moll equation is born.
- Gray, P. R., Hurst, P. J., Lewis, S. H., and Meyer, R. G., *Analysis and Design of Analog Integrated Circuits*. Wiley (5th ed., 2009). The classic go-to book for a real understanding of integrated linear circuit design.
- Howe, R.T. and Sodini, C. G., *Microelectronics, an Integrated Approach*. Prentice-Hall (1996). Introductory IC design.
- Mead, C. and Conway, L., *Introduction to VLSI Systems*. Addison-Wesley (1980). Device physics and circuit design; a classic.
- Muller, R. S., and Kamins, T. I., *Device Electronics for Integrated Circuits*. Wiley (1986). Transistor properties in ICs.
- Sze, S. M., *Physics of Semiconductor Devices*. Wiley (1981). The classic.
- Tsividis, Y. P., and McAndrew, C., *Operation and Modeling of the MOS Transistor*. McGraw-Hill (3rd ed., 2010).

## SPICE

- Cheng, Y. and Hu, C., *MOSFET Modeling & BSIM3 User's Guide*. Springer (1999).
- Kielkowski, R., *Inside SPICE*. McGraw-Hill (1998). A short book with hints about SPICE convergence, etc.
- Liu, W., *MOSFET Models for SPICE Simulation: Including BSIM3v3 and BSIM4*. Wiley (2001). If you use SPICE to analyze MOSFET designs you need this book.
- Massobrio, G. and Antognetti, P., *Semiconductor Device Modeling With SPICE*. McGraw-Hill (2nd ed., 1998). Modeling BJTs, JFETs and MOSFETs.
- Ytterdal, T., Cheng, Y., and Fjeldy, T. A., *Device Modeling for Analog and RF CMOS Circuit Design*. Wiley (2003). MOSFET device physics, and SPICE modeling, noise in MOSFETs.

## Amplifiers, Transducers, and Noise

- Buckingham, M. J., *Noise in Electronic Devices and Systems*. Wiley (1983).
- Hollister, A. L., *Wideband Amplifier Design*. Scitech (2007). Wideband amplifier design techniques using BJTs and FETs, with extensive SPICE analysis.

- Morrison, R. *Grounding and Shielding Techniques in Instrumentation*. Wiley (1986).
- Motchenbacher, C. D. and Connelly, J. A., *Low-noise Electronic System Design*. Wiley (1993). A serious in-depth treatment of low-noise amplifier design.
- Netzer, Y., "The design of low-noise amplifiers." *Proc. IEEE* **69**:728–741 (1981). Excellent review.
- Ott, H., *Noise Reduction Techniques in Electronic Systems*. Wiley (1988). Shielding and low-noise design.
- Radeka, V., "Low-noise techniques in detectors." *Ann. Rev. Nucl. and Part. Physics*, **38**:217–277 (1988). Amplifier design, signal processing, and fundamental limits in charge measurement.

## Op-amps

- Applications Manual for Operational Amplifiers, for Modelling, Measuring, Manipulating, & Much Else*. Philbrick/Nexus Research (1965). Charming compendium from the originators of the first commercial op-amp; these are collectors' items, long out of print.<sup>1</sup>
- Carter, B., and Brown, T. R., *Handbook of Operational Amplifier Applications*. Rework of the classic Burr-Brown handbook, described by Carter as a "treasure... some of the finest works on op amp theory that I have ever seen."
- Frederiksen, T. M., *Intuitive IC op-amps*. Santa Clara, CA: National Semiconductor Corp. (1984). Extremely good treatment at all levels.
- Graeme, J. G., *Applications of Operational Amplifiers: Third Generation Techniques*. McGraw-Hill (1987). One of the Burr-Brown series.
- Jung, W. G., ed., *Op Amp Applications Handbook*. Newnes (2004). Fascinating history section, excellent up-to-date detail.
- Jung, W. G., *IC op-amp Cookbook*. Howard W. Sams & Co. (3rd ed., 1986). Lots of circuits, with explanations. See also Jung's *Audio IC Op-amp Applications*.
- Soclof, S., *Analog Integrated Circuits*. Prentice-Hall (1985). Detailed IC-designer information, useful for IC users as well.
- Zumbahlen, H., *Linear Circuit Design Handbook*. Newnes (2008). Things you need to know from Analog Devices engineers.

<sup>1</sup> But happily resurrected at [http://www.analog.com/library/analogdialogue/archives/philbrick/computing\\_amplifiers.html](http://www.analog.com/library/analogdialogue/archives/philbrick/computing_amplifiers.html).

## Audio

- Duncan, B., *High Performance Audio Power Amplifiers*. Newnes (1996). Excellent review of professional audio power amplifier design.
- Hickman, I., *Analog Electronics*. Newnes (2nd ed., 1999). Interesting overview, one wishes he had written more.
- Hood, J. L., *The Art of Linear Electronics*. Newnes (1998). Audio electronics, including FM.
- Pohlman, K. C., *Principles of Digital Audio*. McGraw-Hill (3rd ed., 1995). All aspects of digital audio; non-mathematical, an easy read.
- Self, D., *Small Signal Audio Design*. Focal Press (2010). Audio design basics and tricks, with a special view from an industry master.
- Strawn, J., ed., *Digital Audio Signal Processing*. A-R Editions Inc. (Madison, WI; 1985). The first article (by Moore) is a magnificent introduction to the mathematics of digital signal processing. Sadly, this volume is out of print.
- Watkinson, J., *The Art of Digital Audio*. Focal Press (3rd ed., 2000). Another nice book on digital audio.

## Filters and Oscillators

- Hilburn, J. L., and Johnson, D. E., *Manual of Active Filter Design*. McGraw-Hill (1982).
- Lancaster, D., *Active Filter Cookbook*. Howard W. Sams & Co. (1979). Explicit design procedure; easy to read.
- Matthys, R. J., *Crystal Oscillator Circuits*. Wiley (1983), Krieger Publishing (revised, 1992).
- Parzen, B., *Design of Crystal and Other Harmonic Oscillators*. Wiley (1983). Discrete oscillator circuits.
- Williams, A. and Taylor, F., *Electronic Filter Design Handbook*. McGraw-Hill (4th ed., 2006). Practical filter design, with formulas, tables, and many examples.
- Zverev, A. I., *Handbook of Filter Synthesis*. Wiley (1967). Extensive tables for passive *LC* and crystal filter design.
- See also Graeme, J. G., under op-amp listings.

## Power, Regulation, and Control

- Basso, C. P., *Switch-Mode Power Supplies: SPICE Simulations and Practical Designs*. McGraw-Hill (2008). The title says it all.
- Billings, K. and Morey, T., *Switchmode Power Supply Handbook*. McGraw-Hill (3rd ed., 2010). Highly readable and comprehensive treatment of an often-confusing topic.

- Erickson, R. W. and Maksimovic, D., *Fundamentals of Power Electronics*. Springer (2nd ed 2001). Learn how to compensate an SMPS feedback loop.
- Grover, F. W., *Inductance Calculations*. Dover (2009 reprint of the 1946 classic). Formulas, tables, and graphs for the inductance of just about anything.
- Hnatek, E. R., *Design of Solid-state Power Supplies*. Van Nostrand Reinhold (1989). Switching supplies.
- MacFadyen, K. A., *Small Transformers and Inductors*. Chapman & Hall (1953). Learn how to calculate leakage inductance in your transformers.
- Maniktala, S., *Switching Power Supplies: A to Z*. Newnes (2nd ed, 2012). Filled with useful unusual material, like magnetics design with the all-important AC-resistance loss analysis.
- Pressman, A., Billings, K., and Morey, T., *Switching Power Supply Design*. McGraw-Hill (3rd ed., 2009). Standard comprehensive book for a two-week introductory course in SMPS design.
- Rogers, G. and Mayhew, Y., *Engineering Thermodynamics: Work and Heat Transfer*. Prentice-Hall (4th ed, 1996). Develop a better understanding of thermal management in electronics.
- Snelling, E. C., *Soft Ferrites*. Butterworth-Heinemann (2nd ed, 1988). The bible for inductor and transformer design.

## Optics and Light

- Friedman, E. and Miller, J. L., *Photonics Rules of Thumb: Optics, Electro-Optics, Fiber-Optics, and Lasers*. McGraw-Hill (2003). What's this doing here? Well, it's an amazing collection of coolstuff, both serious and quixotic (e.g., "crickets as thermometers".)
- Graeme, J. G., *Photodiode Amplifiers: Op Amp Solutions*. McGraw-Hill (1995). The low-down on transimpedance amplifiers.
- Hobbs, P. C. D., *Building Electro-Optical Systems: Making It All Work*. Wiley (2nd ed., 2009). Great balance of theory and practice.
- Lenk, R. and Lenk, C. *Practical Lighting Design with LEDs*. Wiley (2011).
- Schubert, E. F. *Light-Emitting Diodes*. Cambridge University Press (2nd ed., 2006). LED device physics, practical devices, color physics.
- Yariv, A., *Introduction to Optical Electronics*. Rinehart & Winston (1976). Physics of opto-electronics, lasers, and detection.

### **High-speed Digital and RF**

- Hagen, J. B., *Radio-Frequency Electronics: Circuits and Applications*. Cambridge University Press (2nd ed., 2009). Refreshingly different, an insight per page.
- Johnson, H. and Graham, M., *High Speed Digital Design: A Handbook of Black Magic*. Prentice-Hall (1993). Ringing, cross-talk, ground bounce, etc. – a must-have if you're doing fast digital design.
- Johnson, H. and Graham, M., *High Speed Signal Propagation: Advanced Black Magic*. Prentice-Hall (2003). Techniques for pushing the limits of high-speed signal transmission.
- Johnson, R. C., ed., *Antenna Engineering Handbook*. McGraw-Hill (3rd ed., 1992). Comprehensive, excellent tables and design information.
- Krauss, J. D. and Marhefka, R. J., *Antennas for All Applications*. McGraw-Hill (3rd ed., 2001). Highly readable and usable text.
- Ramo, R., Whinnery, J. R. and Van Duzer, T., *Fields and Waves in Communication Electronics*. Wiley (3rd ed., 1994). A classic electricity and magnetism text, with emphasis on communications.
- Roy, K. and Prasad, S., *Low-Power CMOS VLSI Circuit Design*. Wiley (2000).
- Sevick, J., *Transmission Line Transformers*. Noble (4th ed., 2001). Practical guide to understanding and building RF transformers.

### **Digital Signal Processing and Communication**

- Bracewell, R. N., *The Fourier transform and its applications*. McGraw-Hill (3rd ed., 1999). The classic in this field.
- Brigham, E. O., *The Fast Fourier Transform and its Applications*. Prentice-Hall (1988). Highly readable.
- Oppenheim, A. V. and Schafer, R. W., *Discrete-Time Signal Processing*. Prentice-Hall (3rd ed., 2009). Well-received classic on digital signal analysis.
- Sklar, B., *Digital Communications: Fundamentals and Applications*. Prentice-Hall (2nd ed., 2001). Fine introduction to all aspects of digital communications.

### **Logic, Conversion, and Mixed-signal**

- Best, R. E., *Phase-locked Loops*. McGraw-Hill (6th ed., 2007). Advanced techniques.
- Brennan, P. V., *Phase-Locked Loops: Principles and Practice*. McGraw-Hill (1966).
- Gardner, F. M., *Phaselock Techniques*. Wiley (1979). The classic PLL book; emphasis on fundamentals.
- Kester, E., ed., *The Data Conversion Handbook*. Newnes (2004). Includes an excellent history of data conversion, and extensive detail on the nuances of conversion, timing, bandwidth, etc.
- Lancaster, D., *CMOS Cookbook*. Howard W. Sams & Co. (2nd ed., 1997). Good reading, down-to-earth applications. Includes widely used (but rarely mentioned) M<sup>2</sup>L (Mickey Mouse logic) technique.
- Rohde, U. L., *Digital PLL Frequency Synthesizers*. Prentice-Hall (1983). Theory and lots of circuit detail.
- Sheingold, D. H., ed., *Transducer Interfacing Handbook*. Analog Devices (1980).

### **Computers and Programming**

- Hancock, L. and Krieger, M., *The C Primer*. McGraw-Hill (1982). Introduction for beginners.
- Harbison, S. P. and Steele, G. L., Jr., *C: A Reference Manual*. Prentice-Hall (1987). Readable and definitive; has ANSI extensions.
- Wescott, T., *Applied Control Theory for Embedded Systems*. Newnes (2006). From the author of *PID without a Ph.D.*

### **Miscellaneous**

- Grätzer, G., *More Math into L<sup>A</sup>T<sub>E</sub>X*. Springer (4th ed., 2007). Best single reference on typesetting with L<sup>A</sup>T<sub>E</sub>X (the software typesetting language in which this book was written).
- Kleppner, D. and Ramsey, N., *Quick Calculus*. Wiley (2nd ed., 1985). The title is honest, it's the fastest way to learn calculus. Don't be put off by the book's vintage (hey, calculus itself goes back almost four centuries).

# THE OSCILLOSCOPE

## APPENDIX

## O

The oscilloscope (“scope” for short) is, by far, the most useful and versatile electronic circuit test instrument.<sup>1</sup> As usually used, it lets you “see” voltages in a circuit as a function of time, triggering on a particular point of the waveform so that a stationary display results. Contemporary scopes are almost invariably *digital* (input signals are digitized, processed, and displayed), and they do (and usually *better*) what their analog ancestors did. To understand how to use an oscilloscope, we think it best to start with the traditional (and nearly extinct) 2-channel *analog* scope, for which we’ve drawn a block diagram (Figure O.1) and typical front panel (Figure O.2). Digital scopes carry forward nearly all of its features, to which they add an impressive array of capabilities (and a few hazards).

### O.1 The analog oscilloscope

#### O.1.1 Vertical

Beginning with the signal inputs, most analog scopes have two channels; that’s very useful, because you often need to see the relationship between signals. Each channel has a calibrated gain switch, which sets the scale of **VOLTS/DIVISION** on the screen.<sup>2</sup> There’s also a **VARIABLE** gain knob (concentric with the gain switch) in case you want to set a given signal to a certain number of divisions. Warning: be sure the variable gain knob is in the “calibrated” position when making voltage measurements! It’s easy to forget. The better scopes have indicator lights to warn you if the variable gain knob is out of the calibrated position.

The scope is dc-coupled, an essential feature: what you see on the screen is the signal voltage, dc value and all. Sometimes you may want to see a small signal riding on a large dc voltage, though; in that case you can switch

the input to **AC COUPLING**, which capacitively couples the input with a time constant of about 0.1 second. Most scopes also have a grounded input position, which lets you see where zero volts is on the screen. (In **GND** position the signal isn’t shorted to ground, just disconnected from the scope, whose input is grounded.) Scope inputs are usually high-impedance ( $1\text{M}\Omega$  in parallel with about  $20\text{ pF}$ ), as any good voltage-measuring instrument should be.<sup>3</sup> The input resistance of  $1\text{ M}\Omega$  is an accurate and universal value, so that high-impedance attenuating probes can be used (as will be described later); unfortunately, the parallel capacitance is not standardized, which is a bit of a nuisance when changing probes.

The vertical amplifiers include a vertical **POSITION** control, an **INVERT** control on at least one of the channels, and an **INPUT MODE** switch. The latter lets you look at either channel, their sum (their difference, when one channel is inverted), or both. There are two ways to see both: **ALTERNATE**, in which alternate inputs are displayed on successive sweeps of the trace, and **CHOPPED**, in which the trace jumps back and forth rapidly (0.1–1 MHz) between the two signals. **ALTERNATE** mode is generally better, except for slow signals. It is often useful to view signals both ways, to make sure you’re not being deceived.

#### O.1.2 Horizontal

The vertical signal is applied to the vertical deflection electronics, moving the dot up and down on the screen. The horizontal sweep signal is generated by an internal ramp generator, giving deflection proportional to time. As with the vertical amplifiers, there’s a calibrated **TIME/DIVISION** switch and a **VARIABLE** concentric knob; the same warning stated earlier applies here. Most scopes have a **10× MAGNIFIER** and also allow you to use one of the input channels for horizontal deflection (this lets you generate those beloved but generally useless

<sup>1</sup> It is sometimes said that practitioners of other engineering disciplines are especially envious of EEs, because we are blessed with such a splendid instrument with which to visualize the happenings in our circuits.

<sup>2</sup> Note that the two channels can be set for different scale factors, offsets, and coupling. This goes also for digital scopes, which commonly have four channels.

<sup>3</sup> Scopes intended for high-frequency measurements, going beyond 100 MHz, say, offer also a  $50\Omega$  input impedance option.

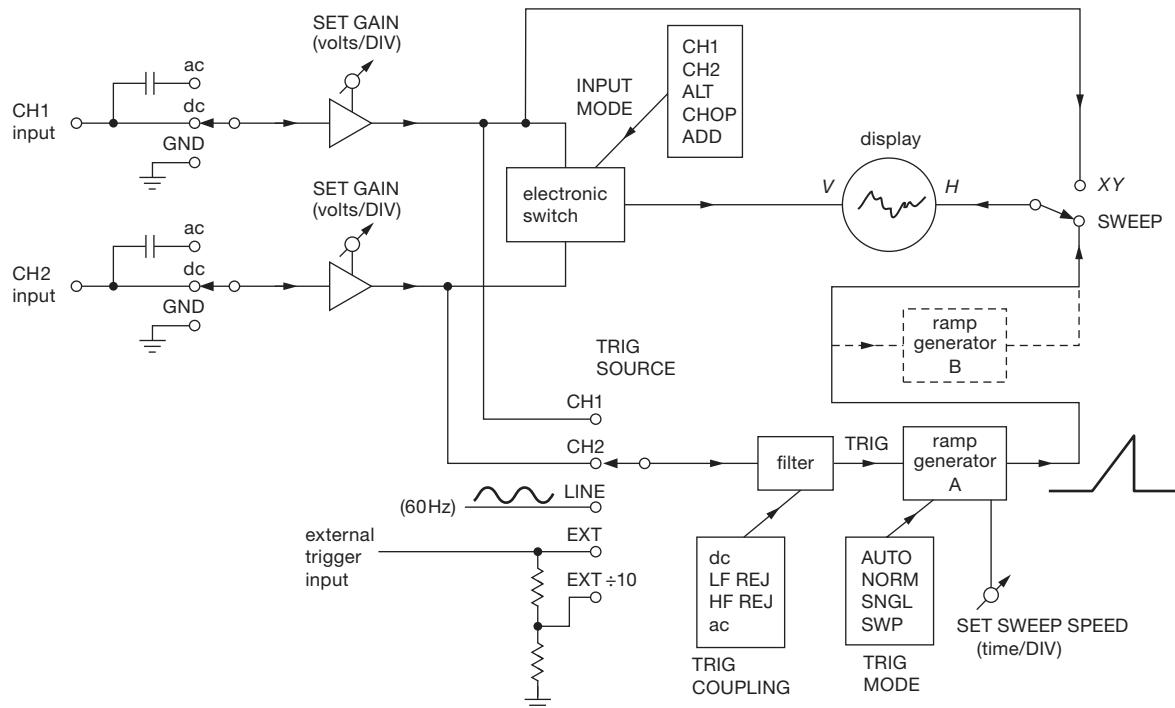


Figure O.1. Block diagram of a 2-channel analog oscilloscope.

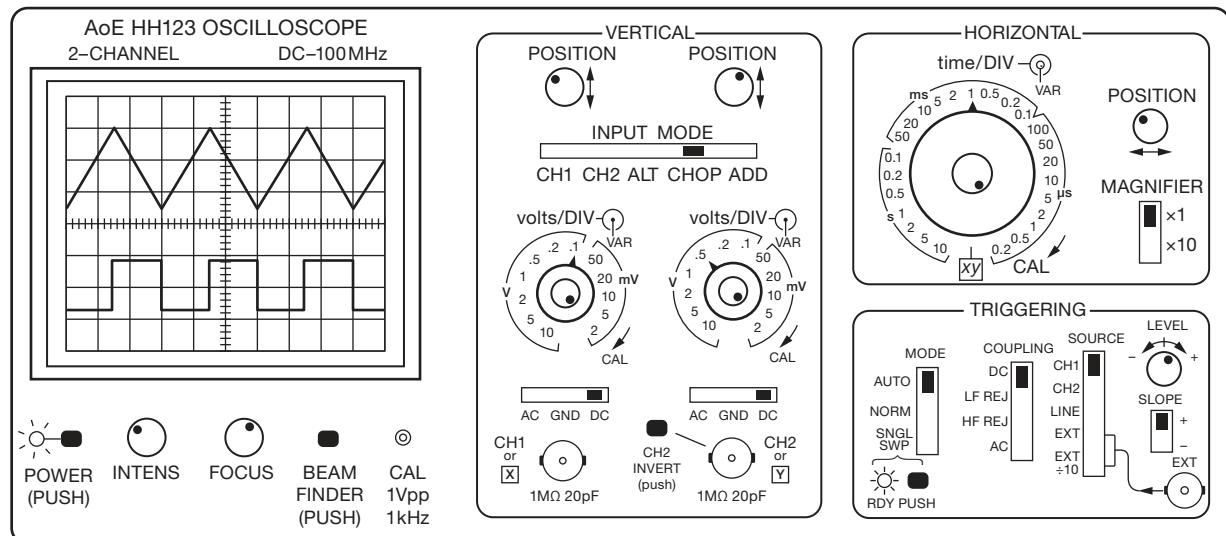


Figure O.2. Portrait of a 2-channel analog scope.

"Lissajous figures" featured in elementary books and science fiction movies).

### O.1.3 Triggering

Now comes the trickiest part: *triggering*. We've got vertical signals and horizontal sweep; that's what's needed for a graph of voltage versus time. But if the horizontal sweep doesn't catch the input signal at the same point in its wave-

form each time (assuming the signal is repetitive), the display will be a mess – a picture of the input waveform superimposed over itself at different times. The trigger circuitry lets you select a **LEVEL** and **SLOPE** (+ or -) on the waveform at which to begin the sweep. You can see from the front panel that you have a number of choices about trigger sources and mode. **NORMAL** mode produces a sweep only when the source selected crosses through the trigger point you have set, moving in the direction (**SLOPE**) you have selected. In practice, you adjust the level control for a stable display. In **AUTO** the sweep will “free run” if no signal is present; this is good if the signal sometimes drops to small values, since the display won’t disappear and make you think the signal has gone away. It’s the best mode to use if you are looking at a bunch of different signals and don’t want to bother setting the trigger each time. **SINGLE SWEEP** is used for nonrepetitive signals. **LINE** causes the sweep to trigger on the ac power line, handy if you’re looking at hum or ripple in a circuit. The **EXTERNAL** trigger inputs are used if you have a clean signal available at the same rate as some “dirty” signal you’re trying to see; it’s often used in situations where you are driving some circuit with a test signal, or in digital circuits where some “clock” signal synchronizes circuit operations. The various coupling modes are useful when viewing composite signals; for instance, you may want to look at an audio signal of a few kilohertz that has some spikes on it. The **HF REJ** position (high-frequency reject) puts a low-pass filter in front of the trigger circuitry, preventing false triggering on the spikes. If the spikes happen to be of interest, you can trigger on them instead in **LF REJ** position.

Many scopes now have **BEAM FINDER** and **TRIGGER VIEW** controls. The beam finder is handy if you’re lost and can’t find the trace; it’s a favorite of beginners. Trigger view displays the trigger signal; it’s especially handy when triggering from external sources.

#### O.1.4 Hints for beginners

Sometimes it’s hard to get *anything* to show on the scope. Begin by turning the scope on; set triggering for **AUTO**, **DC COUPLING**, **CH 1**. Set sweep speed at 1 ms/div, cal, and the magnifier off ( $\times 1$ ). Ground the vertical inputs, turn up the intensity, and wiggle the vertical position control until a horizontal line appears (if you have trouble at this point, try the beam finder).<sup>4</sup> Now you can apply a signal, unground

the input, and fiddle with the trigger. Become familiar with the way things look when the vertical gain is far too high, when the sweep speed is too fast or slow, and when the trigger is adjusted incorrectly.

#### O.1.5 Probes

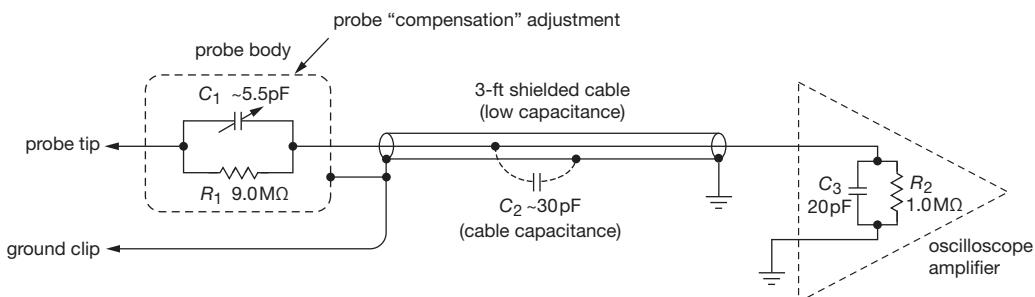
The oscilloscope input capacitance seen by a circuit under test can be undesirably high, especially when the necessary shielded connecting cable is included. The resulting input impedance ( $1M\Omega$  in parallel with  $100\text{ pF}$  or so) is often too low for sensitive circuits and loads it by the usual voltage divider action; for example, at 10 MHz a  $100\text{ pF}$  load looks like  $160\Omega$  – ouch! Worse yet, the capacitance may cause some circuits to misbehave, even to the point of going into oscillation. In such cases the scope obviously is not acting like the “low-profile” measurement instrument we expect; it’s more like a bull in a china shop.

The usual solution is the use of high-impedance “probes.” In simplified form,<sup>5</sup> the popular  $10\times$  probe works as shown in Figure O.3. At dc it’s just a  $10\times$  voltage divider. By adjusting  $C_1$  to be  $\frac{1}{9}$ th the parallel capacitance of  $C_2$  and  $C_3$ , the circuit becomes a  $10\times$  divider at all frequencies, with input impedance of  $10M\Omega$  in parallel with a few picofarads. In practice, you adjust the probe by looking at a square wave of about 1 kHz, available on all scopes as **CALIB**, or **PROBE ADJ**, setting the capacitor on the probe for a clean square wave without overshoot. Sometimes the adjustment is cleverly hidden; on some probes you twist the body of the probe and lock it by tightening a second threaded part. One drawback: a  $10\times$  probe makes it difficult to look at signals of only a few millivolts; for these situations use a “ $1\times$  probe,” which is simply a length of low-capacitance shielded cable with the usual probe hardware (wire “grabber,” ground clip, handsome knurled handle, etc.). The  $10\times$  probe should be the standard probe, left connected to the scope, with the  $1\times$  probe used when necessary. Some probes feature a convenient choice of  $1\times$  or  $10\times$  attenuation, switchable at the probe tip.

Even with a  $10\times$  probe, the circuit loading may be unacceptable; after all, its improvement is just the same factor of ten by which it attenuates the input signal. You *can* get  $100\times$  probes, with correspondingly higher input

<sup>5</sup> In practice the cable itself is made from resistance wire, to damp transmission-line effects (frequency peaking and transient reflections, see Appendix H), an elegant 1959 invention by Kobbe and Polits (US Patent 2,883,619); you also see tricks such as a series *RC* across the scope terminals (e.g.,  $500\Omega$  and a trimmer capacitor), to provide a transmission-line match at high frequencies.

<sup>4</sup> Curiously, some scopes (for example the once-popular Tektronix 400 series) don’t sweep on **AUTO** unless the trigger level is adjusted correctly.



**Figure O.3.** A  $10\times$  passive scope probe attenuates signals by a factor of ten at all frequencies, conveniently raising the input impedance by the same factor. (In practice additional tricks are used to suppress transmission-line effects, see text.)

impedance (e.g., the Tektronix P5100 series), but these are intended primarily for viewing high-voltage signals (the scope itself is usually limited to a maximum of  $\pm 400$  V at the input connector), and they do not excel in important features such as small physical size. What you do, instead, is to use an *active probe*, which uses a FET follower at the tip to achieve an input capacitance  $<1$  pF.<sup>6</sup> Active probes, being intended for wideband use, are intended to drive a  $50\ \Omega$  input (available on most high-speed scopes; if not, attach a  $50\ \Omega$  pass-through terminator); they require a source of power, available at the scope's input connector (on digital scopes), or provided by a stand-alone box like the Tektronix 1103.

Any discussion of probes would be incomplete without a mention of *current probes*: these handy devices, when clipped around a wire in some circuit, convert the circuit's current to a voltage waveform that's displayed on the scope. The simplest current probes are inherently ac-coupled (they wrap a secondary winding around a magnetic split core that surrounds the one-turn wire "primary") and thus do not sense dc current; the fancier types use a combination of Hall effect and transformer coupling to achieve response down to dc. Examples of the latter are the Tektronix A622 (dc to 100 kHz) and the TCP312A (dc to 100 MHz); the latter requires the matching TCPA300 amplifier.

### O.1.6 Grounds

As with most test instruments, the oscilloscope input is referred to the instrument ground (the outer connection of the input BNC connectors), which is usually tied electrically to the case. That, in turn, connects to the ground lead of the ac power line, via the 3-wire power cord. This means that you

cannot measure voltages between the two arbitrary points in a circuit, but are forced to measure signals relative to this universal ground.

An important caution is in order here: if you try to connect the ground clip of an oscilloscope probe to a point in the circuit that is at some voltage relative to ground, you will end up shorting it to ground. This can have disastrous consequences to the circuit under test; in addition, it can be downright dangerous with circuits that are "hot to ground" (for example line-powered switching power supplies). If it is imperative to look at the signal between two points, you can make a differential measurement by inverting one input channel and switching to ADD, or you can use an external differential preamp (e.g., the LeCroy DA1855A). In desperate situations we have been known to "float" the scope by lifting the ground lead at the power cord, but this is *not recommended*, unless you really know what you're doing (and agree to waive any liability on our part).

Another caution about grounds when you're measuring weak signals or high frequencies: be sure the oscilloscope ground is the same as the circuit ground where you're measuring. The best way to do this is by connecting the short ground wire on the probe body directly to the circuit ground,<sup>7</sup> then checking by measuring the voltage of "ground" with the probe, observing no signal. One problem with this scheme is that those short ground clips are usually missing, lost! Keep your probe accessories in a drawer somewhere.

### O.1.7 Other analog scope features

Many scopes have a **DELAYED SWEEP** that lets you see a segment of a waveform occurring some time after the trigger point. You can dial the delay accurately with a multi-turn adjustment and a second sweep-speed switch. A delay

<sup>6</sup> One of our favorites is the Tektronix P6243,  $<1$  pF and 1 GHz bandwidth.

<sup>7</sup> See the illustrations in Figure 12.32

mode known as **A INTENSIFIED BY B** lets you display the whole waveform at the first sweep speed, with the delayed segment brightened; this is handy during setup. Scopes with delayed sweep sometimes have “mixed sweep,” in which the trace begins at one sweep speed, then switches to a second (usually faster) speed after the selected delay. Another option is to begin the delayed sweep either immediately after the selected delay or at the next trigger point after the delay; there are two sets of trigger controls, so the two trigger points can be set individually. (Don’t confuse delayed sweep with “signal delay.” All good analog scopes have a delay in the signal channel, so you can display the event that caused the trigger; it lets you look a little bit backward in time! See the photographs of the analog delay lines in Figures H.19 and H.21).

A common feature of analog scopes is a **TRIGGER HOLD OFF** control; it inhibits triggering for an adjustable interval after each sweep, and it is very useful when viewing complicated waveforms without the simple periodicity of, say, a sinewave. The usual case is a digital waveform with a complicated sequence of 1s and 0s, which won’t generate a stable display otherwise (except by adjustment of the sweep-speed vernier, which means you don’t get a calibrated sweep).

All scopes (analog and digital) include some **BANDWIDTH LIMIT** vertical amplifier options (for simplicity, not shown in Figures O.1 and O.2), useful for reducing the amount of wideband “fuzz” on the displayed trace when you’re working with relatively slow signals.

During the height of the analog scope era, you could get scopes with on-screen “storage” (for single-shot capture) and scopes with an impressive array of plug-in modules that let you do lots of interesting stuff, including display of eight traces, or spectrum analysis, or accurate (digital) measurements of voltage and time on waveforms, and so on. Happily, these functions and many others (e.g., looking far backward in time from the triggering event) are now embodied in the dominant oscilloscope species, the *digital* oscilloscope. Let’s take a look.

## 0.2 The digital oscilloscope

Analog scopes are easy to use, but they are seriously limiting in what you can do. For example, (a) it’s hard to see a “single-shot” event; (b) you can’t store a trace, or compare a live trace with an earlier trace; (c) you can’t extract a trace for measurement or illustration; and (d) you can’t look back in time to see what happened before the triggering event.

Digital scopes effortlessly provide these and many other

capabilities; and, because of the stunning capability and low cost of digital conversion and processing, they are, ironically, less expensive than an analog scope (if you can find one) of comparable bandwidth. The transition to usable and friendly digital scopes was rocky at first, but they are now ubiquitous and universal.

The basic scheme (Figure O.4) is to digitize the incoming signal after the frontend stages of programmable gain and bandwidth limiting, capture the samples in a fast circular buffer memory, and then use a processor (or multiple processors) to do all the signal processing, measurements, conversion to a meaningful display, user interface, and I/O. We’ll keep this section brief, and merely run through some capabilities of digital scopes.

### 0.2.1 What’s different?

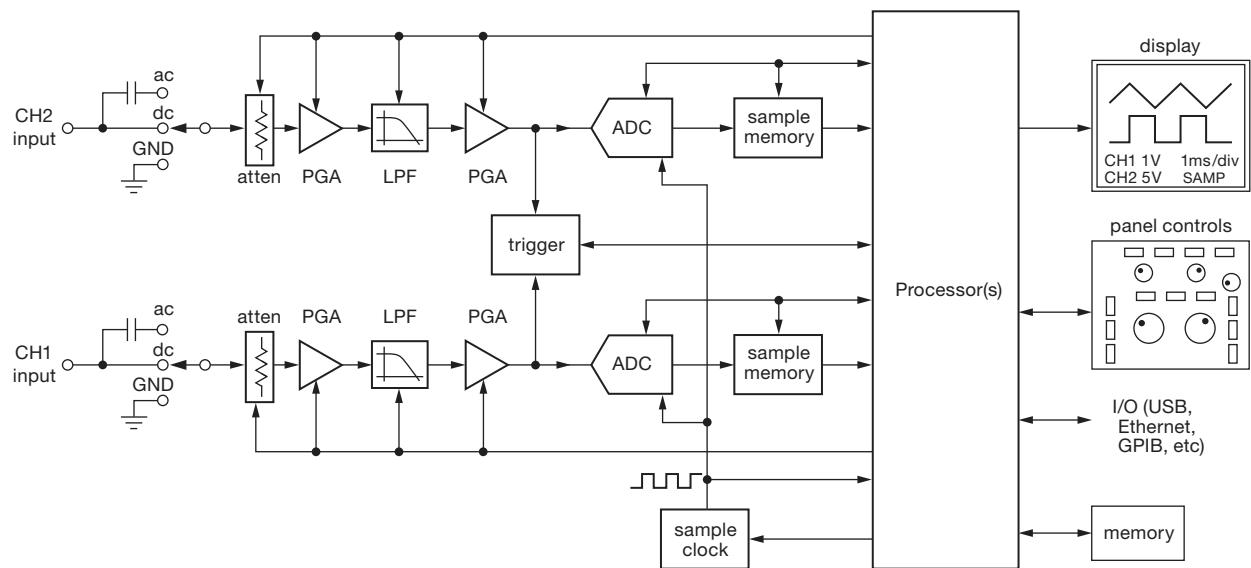
In no particular order:

**Front-end:** The signal emerging from the (variable-gain) input amplifiers is digitized at some sampling rate  $f_{\text{samp}}$  (typically 1 Gsps or more, but always above the minimum Nyquist rate of  $2f_{\text{max}}$  when the scope is set at fast enough sweep rates to resolve the scope’s bandwidth  $f_{\text{max}}$ ). But – **important** – see “aliasing,” below). The digital samples, typically of 8 bits resolution, are stored at full speed into a *sample memory* (or “capture memory”), often of length 1 Mpt or more per channel (and reaching 1 Gpt at the high end). Note that, even though digital scopes let you zoom in after a trace is captured, the resolution depends on the vertical scale factor, because of the fixed bit depth of conversion.

**Simultaneous on all channels:** Digital scopes digitize all channels simultaneously; there’s no “alternate” or “chop.” Most digital scopes come in 2- or 4-channel flavors, augmented in “mixed-signal” scopes by sixteen or more 1-bit (i.e., logic-level) channels.

**Pre-trigger:** Because digitized input signals are pouring into memory, you can set a trigger condition (most simply, level and slope; but see “Smart trigger,” below) and, when it is satisfied, you’ve got substantial pre-history in the sample memory. From the user’s point of view, you can simply set the displayed trigger pointer to the right portion of the screen to reveal what came before. And you can walk backward or forward to your heart’s content through a saved single-shot capture (see “Single-shot capture,” below).

**Display:** The time interval between points on the *displayed* waveform (the “waveform interval”) is typically longer (often much longer) than the sampling interval



**Figure O.4.** Block diagram of a 2-channel digital oscilloscope.

$1/f_{\text{samp}}$ . This allows for various modes of processing the sampled points to produce the displayed waveform, in particular:

**sample** one sample point per waveform interval is displayed; the rest are discarded. Simple, but susceptible to aliasing, see below.

**peak detect** the highest and lowest sampled points in two successive waveform intervals are displayed. Produces a thicker waveform, but no short spikes are lost.

**envelope** similar to peak detect, but combines min/max from multiple triggered acquisitions. Useful for seeing excursions from an ideal repetitive waveform.

**average** each point in the displayed waveform is the calculated average of single samples (as in sample mode) over many (a settable number, e.g., 2, 4, 8, ... 512) triggered acquisitions. Greatly reduces noise, without reducing bandwidth, but requires a repetitive signal.

**high resolution** each displayed point is calculated as the average value of the multiple samples captured within one waveform interval. Provides higher resolution, and does not require a repetitive signal, but reduces bandwidth.<sup>8</sup>

**Persistence:** Engineers with gray beards wax poetic about the beautiful gradations of intensity with which analog scopes display waveforms. Digital scopes took a while to catch up, but now they trumpet their ability to do the same, with terms like “digital phosphor,” “persistence trace,” and “digital persistence.”

**Single-shot capture:** Digital scopes excel in single-shot capture. You can troll through the sampled data post-capture, either manually (by turning knobs for scrolling and magnifying) or with some helpful automated search tools (e.g., Tektronix’s “Wave Inspector” – the name says it all).

**Slow sweep:** Analog scopes are hopeless when you want to view a waveform that takes many seconds; digital scopes couldn’t care less. Use “rolling mode” at slow sweep rates.

**Save/Recall:** You can save one or more waveforms to memory, bring them back for comparison, etc. You can also save the scope’s *state* (i.e., settings).

**Measurements:** The data’s all there, so digital scopes have no problem measuring period, frequency, amplitude, time interval, duty cycle, etc. These measurements usually update continuously, and you can use settable horizontal and vertical cursors to define the measurement regions and intervals.

**Math:** Going further, digital scopes let you calculate products (e.g., to measure power from voltage and current), quotients (to normalize a waveform), jitter, histograms, frequency spectra, etc. Almost limitless possibilities,

<sup>8</sup> You can think of this as a “horizontal average” along one waveform capture, as compared with the “vertical averaging” of single sample points in successive stacked waveforms in *average* mode.

but you may prefer to extract the data and do the math offline.

**I/O:** You can send waveforms and data out (via connected Ethernet or whatever), and you can control the scope's operations remotely. A networked data-acquisition system!

**Mixed-signal:** Many digital scopes come with a bunch of 2-level channels (typically 16 or 32) along with the usual 2 or 4 full-resolution channels; so it works as a logic analyzer, but augmented by a few channels of clear waveform view. As with a traditional logic analyzer, you can trigger on a defined set of levels, and it can do bus decoding, bus triggering, and other fancy stuff for you.

**Smart trigger:** Good digital scopes let you trigger on just about any condition you can imagine: pulse width  $<$ ,  $>$ ,  $=$ , or  $\neq$  to some value; runt pulses and glitches; setup or hold time violations; specified range of rise-times or falltimes; specified conditions or violations on serial buses; trigger after  $n$  events; and so on. (check-out the enjoyable reading in a datasheet from Tektronix, LeCroy, Keysight/Agilent, or Rohde & Schwarz)

**Limit/Mask testing:** You can set up a template and detect out-of-spec waveforms, for Go/No-Go testing on a production line; ditto for jitter and other measurable parameters.

**Autoset:** It's easy to get lost in this multi-dimensional wilderness; digital scopes provide a rescue button (autoset, autoscale, or some linguistic variant, depending on manufacturer), which will at least get something going on the screen (but see the Cautions, next).

**Probe skew:** When you're using several different probing systems (e.g., passive  $10\times$  probe, active FET probe, current probe) the signal delays can vary by tens of nanoseconds or more, completely disrupting the fidelity of the multichannel display (hey, you may be tricked into thinking you've got a violation of causality – effect precedes cause!<sup>9</sup>).

**Probe readout, Probe power:** Probes for use with contemporary digital scopes have extra connections by which they communicate their attenuation factor ( $\times 1$ ,  $\times 10$ ,  $\times 100$ ) and other useful scale factor information (e.g., the amps/div of a current probe); they use such connections also to send power to the probe (needed, for example, with FET active voltage probes or Hall-effect current probes). This can be an annoyance, though, if your scope's input connectors are of the wrong format

(which happens even within one manufacturer's scope offerings).

## O.2.2 Some cautions

There's not much not-to-like about digital oscilloscopes. But here are a few cautions, ways in which a digital scope can trick the unwary.

### A. Aliasing

This one can fool even the experienced scope user: digital scopes are designed such that the *maximum* sample rate is always adequate for signals up to the scope's full bandwidth; but when you are running at a slower sweep rate in "sample" mode (i.e., one sample per displayed waveform point) the effective sample rate is much lower. So you may see some serious nonsense (jittery unstable signal, inability to trigger, weird change of waveform when sweep rate is changed, etc.) if there's a high frequency signal present.<sup>10</sup>

If you suspect aliasing, try speeding up the sweep, or changing to **PEAK DETECT** mode. Aliasing can be really annoying when you're dealing with signals that combine timescales (the classic one was analog television, with a 3.59 MHz color carrier on a  $\sim 15$  kHz horizontal line frequency).

### B. Dead time

For human visual perception it's necessary to update the scope display at only  $\sim 100$  times per second or so. If the scope captures input waveform data only at that rate, the fraction of time that it's sensitive to important signal events (like a glitch or timing violation) may be exceedingly low. For example, at a middling sweep rate of  $1\ \mu\text{s}/\text{div}$  (thus  $10\ \mu\text{s}$  per sweep) a scope that's updating 100 times/second is active only 0.1% of the time.

When scope users became aware of this ugly secret, scope manufacturers went at it, and they now provide some measure of true update rate (usually in the form of "waveforms per second," typically in the range of 100,000–1,000,000). Be careful when evaluating such metrics, because there's more than a little "specsmanship" going on.<sup>11</sup>

<sup>10</sup> Dare we admit? One of the authors was testing a circuit that operates at frequencies in the kilohertz range, driving it with a digital function generator at 1.0 kHz. Go out to lunch, come back, look at the waveform – the scope is broken, won't trigger, jittery waveform sliding left and right. Weird. Tried everything. Then noticed that the generator had defaulted to its 1.0 megahertz setting. Ha! Problem solved (and I won't ever tell anyone how dumb I was).

<sup>11</sup> Ask your scope salesperson, they love to flame the lying competition.

<sup>9</sup> As does the protagonist in Asimov's delightful short story from 1960: *Thiotimoline and the Space Age*.

### C. Lost in a multidimensional vector space

Analog scopes are simple, and you can see the full state of the instrument just by looking at the knobs and indicators. No such luck with the immense capabilities of digital scopes. Early digital scopes were particularly troublesome, lacking annunciators and (mostly) knobs. They've improved enormously,<sup>12</sup> but it's still awfully easy to be sitting in front of a scope that just isn't triggering, or showing significant vital signs. It takes some keen intuition to know which menu to pull down (horizontal? trigger? mode?) to get to the problem. It may be as simple as triggering on the wrong channel; or it might be that you've left the display in **AVERAGING** mode, and lack of a stable repetitive trig-

ger produces a bunch of garbage. And you can even waste a minute not realizing that the thing is in **SINGLE-SWEEP** or **STOP** mode.

### D. The scope is lying to you

When observing signals with a digital scope, you may be victimized by a blessing (a vast array of measurement capabilities and settings) that becomes a curse (the scope's settings are not what you think). It's easy to forget some obscure but important settings that falsify the measurements you think you're making. For example, it's easy to forget (we've done it, often) that you've left an earlier **PROBE SKEW** compensation in effect, or that some channels still have **BANDWIDTH LIMIT** set. Such oversights corrupt your measurements in less-than-obvious ways that you may not notice for quite a while; and when you do, you're sentenced to serve time repeating the measurements properly.

<sup>12</sup> We are particularly fond of the "QuickMenu" feature that was introduced by Tektronix in their original TDS3000-series "lunchbox" scopes. Inexplicably, this highly useful feature has been eliminated (despite our howls of protest) in Tek's successor scopes. We've been badgering them ever since.

# ACRONYMS AND ABBREVIATIONS

## APPENDIX

## P

Electrical engineers are fond (*too* fond, some would say) of acronyms and abbreviations, a familiarity that every educated circuit designer must necessarily acquire. To assist in that education, and for handy reference, we here provide a lightly annotated list of terms used in this book.

ac: literally “alternating current” (i.e., alternating *voltage*); more generally a varying signal  
AC(T): advanced CMOS (logic family)  
A/D: analog-to-digital  
ADC: analog-to-digital converter  
ADI: Analog Devices Inc.  
AES: Audio Engineering Society  
AFC: automatic frequency control  
AGC: automatic gain control  
AGF: all-glass fiber  
AHC(T): advanced high-speed CMOS (logic family)  
ALS: advanced low-power Schottky (logic family)  
ALU: arithmetic logic unit (in a processor)  
ALV: advanced low-voltage (logic family)  
AM: amplitude modulation  
ANSI: American National Standards Institute  
APD: avalanche photodiode  
APF: all-plastic fiber  
ARM: a popular processor architecture from ARM Holdings  
ASIC: application-specific full-custom integrated circuit  
ASCII: American Standard Code for Information Interchange  
ASF: all-silica fiber  
ASSP: application-specific standard product  
ATA: advanced technology attachment (a disk interface; see PATA, SATA)  
ATAPI: ATA packet interface (a generalized ATA)  
ATE: automated test equipment  
ATM: asynchronous transfer mode  
ATSC: Advanced Television Systems Committee (digital TV standards)  
AUC: advanced ultra-low-voltage CMOS (logic family)  
AVC: advanced (low)-voltage CMOS (logic family)  
AVR: a microcontroller family from Atmel Corp.

AWG: American wire gauge  
AZ: auto-zero  
BBM: break-before-make (switch)  
BCD: binary-coded decimal  
BGA: ball-grid array (an IC package)  
BJT: bipolar junction transistor  
BNC: bayonet Neill-Concelman (connector)  
BPSK: binary phase-shifting keying  
BRT: bias resistor transistor  
BSS: broadcast satellite services  
BV: breakdown voltage  
BW: bandwidth  
C0G: low tempco (stable) ceramic dielectric  
CA: conditional access  
CAN: controller area network (bus)  
CANH: controller area network high (a CAN signal)  
CANL: controller area network low (a CAN signal)  
CBR: constant bitrate (coding)  
CCD: charge-coupled device  
CCFL: cold-cathode fluorescent lamp  
CCM: continuous-conduction mode (in a power converter)  
CD: compact disc (optical storage)  
CDMA: code-division multiple access  
CDR: clock and data recovery  
CES: Consumer Electronics Show  
CF: compact flash (memory card)  
CFB: current feedback (op-amp)  
CHE: channel hot-electron  
CLB: configurable logic block  
CML: current-mode logic  
CMOS: complementary metal-oxide semiconductor  
CMRR: common-mode rejection ratio  
codec: coder-decoder  
CPLD: complex programmable-logic device  
CPU: central processing unit  
CR: carriage return  
CRC: cyclic redundancy checksum  
CRT: cathode-ray tube  
CSMA: carrier-sense multiple-access  
CSP: chip-scale package

CTR: current transfer ratio	eSATA: external SATA interface
CTS: clear to send (in a serial link)	ESD: electrostatic discharge
CVBS: composite video, blanking, and sync	ESL: equivalent series inductance
CVSD: continuously variable-slope delta-modulation	ESR: equivalent series resistance
DA: dielectric absorption	ETF: 8-to-14 (digital coding)
D/A: digital-to-analog	FCC: Federal Communications Commission
DAC: digital-to-analog converter	FDNR: frequency-dependent negative resistor
DAQ: data-acquisition system	FED: field-emission display
DBS: direct broadcast satellite	FET: field-effect transistor
dc: direct current (i.e., a fixed voltage)	FFS: fringe-field switching (an LCD display technology)
DCE: data communications equipment (in a serial link)	FFT: fast Fourier transform
DCM: discontinuous-conduction mode (in a power converter)	FG: floating gate
DDR: double data rate (memory)	FGA: floating-gate array
DDS: direct digital synthesis	FIFO: first-in-first-out (memory)
DFC: digital frequency converter	FIR: finite-impulse-response (filter)
DIN: Deutsches Institut für Normung (a German standards organization); a connector series	FM: frequency modulation
DIP: dual in-line package	F-N Fowler–Nordheim tunelling
DIR: direction (a control signal)	FOT: fiber-optic transceiver
DMA: direct memory access	FPBW: full-power bandwidth
DMM: digital multimeter	FPGA: field-programmable gate array
DNL: differential nonlinearity	FR-4: “flame-retardant 4” (glass-epoxy PCB material)
DPCP: display-port content protection	FRAM (also FeRAM, F-RAM): ferroelectric random-access memory
DPDT: double-pole double-throw (switch)	FSB: front-side bus (of a computer processor)
DRAM: dynamic random-access memory	FSE: full-sunlight equivalent (you saw it here first!)
DSBGA: die-size ball-grid array (an IC package)	GAL: generic array logic
DSL: digital subscriber line (for data over telephone line)	GBP, GBW: gain-bandwidth product
DSP: digital signal processing (or processor)	GCC: GNU C-compiler
DSR: data set ready (in a serial link)	GCPW: grounded coplanar waveguide
DTE: data terminal equipment (in a serial link)	GDT: gas-discharge tube
DTH: direct-to-home (satellite TV)	GIC: generalized impedance converter
DTL: diode–transistor logic	GMR: giant magnetoresistance
DTR: data terminal ready (in a serial link)	GND: ground
DUT: device under test	GPIB: general-purpose interface bus
DVI: digital visual interface (for digital video)	GPL: graphical programming language; general public license (in GNU)
DVM: digital voltmeter	GPS: global positioning system
DVR: digital video recorder	GPU: graphics-processor unit
EAROM: electrically alterable read-only memory	GUI: graphical user interface
EAS: avalanche energy specification	HAPD: hybrid avalanche photodiode (detector)
ECL: emitter-coupled logic	HBM: human body model
EEPROM: electrically erasable programmable read-only memory	HC(T): high-speed CMOS (logic family)
EIA: Electronic Industries Alliance (standards and trade organization)	HDCP: high-bandwidth digital content protection
EMF: electromotive force ( $\sim$ voltage)	HDD: hard-disk drive
EMI: electromagnetic interference	HDL: hardware description language
ENOB: effective number of bits	HDMI: high-definition multimedia interface (for digital display)
EPROM: erasable programmable read-only memory	HDTV: high-definition television
	HI: human interface

HP: Hewlett-Packard	LVDS: low-voltage differential signaling
HV: high voltage	LVPECL: low-voltage positive emitter-coupled logic
IC: integrated circuit	LVX: low-voltage crossvolt (logic family)
ICSP: in-circuit serial programming	LXI: LAN eXtensions for Instrumentation
IDC: insulation displacement connector	
IDE: integrated development environment (for coding)	$\mu$ C: microcontroller
IEC: International Electrotechnical Commission	MAC: multiplier-accumulator; media-access control
IEEE: Institute of Electrical and Electronic Engineers	MBB: make-before-break (switch)
IF: intermediate frequency (in RF receiver)	MCU: microcontroller unit
IGBT: insulated-gate bipolar transistor	MDAC: multiplying digital-to-analog converter
IGFET: insulated-gate field-effect transistor	MEMS: microelectromechanical system
IIC (I <sup>2</sup> C): inter-integrated-circuit (a serial bus)	MFB: multiple-feedback (active filter)
IIR: infinite-impulse-response (filter)	MIPS: mega-instructions per second
INA: instrumentation amplifier	MLC: multilevel cell (in nonvolatile memory)
INL: integral nonlinearity (in A/D conversion)	MMU: memory management unit
I/O: input-output	modem: modulator-demodulator
IP: Internet protocol; intellectual property	MOS: metal-oxide semiconductor
IPS: in-plane switching (an LCD display technology)	MOSFET: metal-oxide semiconductor field-effect transistor
IR: infrared	MOV: metal-oxide varistor (surge protector)
ISA: International Society of Automation (a standard-setting organization)	MPTS: multiprogram transport stream (in digital TV)
ISI: intersymbol interference	MPU: microprocessor unit
JFET: junction field-effect transistor	MRAM: magnetoresistive random-access memory
JTAG: Joint Test Action Group (an IC interface)	MRI: magnetic resonance imaging
KCL: Kirchhoff's current law	MSB: most-significant bit
KVL: Kirchhoff's voltage law	MSI: medium-scale integration
LAB: logic array block (in programmable logic)	MUX: multiplexer
LAN: local area network	NAN: not a number
LCD: liquid-crystal display	NC: normally closed (switch)
LCX: low-voltage CMOS crossvolt (logic family)	NECL: negative emitter-coupled logic
LDO: low-dropout (linear voltage regulator)	NEMA: National Electrical Manufacturers Association (a standards-setting organization)
LE: logic element (in programmable logic)	NIC: negative-impedance converter; network interface card
LED: light-emitting diode	NiCd: nickel cadmium (battery)
LFSR: linear feedback shift register	NiMH: nickel metal-hydride (battery)
LIFO: last-in first-out (memory)	NIST: National Institute of Standards and Technology
Li-ion: lithium ion (battery)	NMI: nonmaskable interrupt
LIN: local-interconnect network (bus)	NMR: nuclear magnetic resonance
LNA: low-noise amplifier	nMOS: n-type metal-oxide semiconductor
LNBF: low-noise block-downconverter plus feed (for satellite TV)	NO: normally open (switch)
LO: local oscillator (in RF receiver)	NPO: low tempco (stable) ceramic dielectric
LPF: lowpass filter	NRZ: nonreturn to zero (data code)
LPT: line printer (a parallel port)	NRZI: nonreturn to zero inverted (data code)
LS: low-power Schottky (logic family)	NSC: National Semiconductor Corporation (now part of TI)
LSB: least-significant bit	NTC: negative temperature coefficient
LSI: large-scale integrated circuit	NTSC: National Television System Committee (analog TV standard)
LUT: lookup table	NV: nonvolatile
LV: low-voltage (logic family)	
LVC: low-voltage CMOS (logic family)	

NVM: nonvolatile memory	PRBS: pseudorandom bit sequence
O/C: open-collector (logic output)	PROM: phase-change read-only memory
OCXO: oven-controlled crystal oscillator	PSRAM: pseudostatic random-access memory
O/D: open-drain (logic output)	PSRR: power-supply rejection ratio
OEM: original equipment manufacturer	PTAT: proportional to absolute temperature
OLED: organic light-emitting diode	PUJT: programmable unijunction transistor
op-amp: operational amplifier	PV: photovoltaic (light detector)
OSI: open systems interconnection (network hierarchy)	PVC: polyvinyl chloride (insulator)
OSR: oversampling ratio	PVR: personal video recorder
OTA: over-the-air (broadcasting)	PWM: pulse-width modulation
OTP: one-time-programmable NV memory	QAM: quadrature amplitude modulation
PAL: programmable array logic	QPSK: quadrature phase-shift keying
PARC: Palo Alto Research Center	RAM: random-access memory
PATA: parallel ATA interface	RCO: ripple-clock output
PC: printed circuit; personal computer	RD: receive data (in a serial link)
PCB: printed circuit board	RF: radiofrequency
PCF: plastic-clad fiber	RFI: radiofrequency interference
PCI: peripheral component interface (a computer bus)	RG-nn: “Radio Guide” (coax cable designators)
PCIe (also PCI-E): extended peripheral component interface	RGB: red–green–blue (video signals)
PCM: pulse-code modulation	RISC: reduced instruction set computing
PCMCIA: Personal Computer Memory Card International Association (a card interface standard) <sup>1</sup>	RLL: run-length limited (digital codes)
PDA: personal digital assistant	rms: root-mean-square
PECL: positive emitter-coupled logic	ROM: read-only memory
PEN: polyethylene naphthalate (a capacitor dielectric)	RRI: rail-to-rail input
PFC: power-factor correction (in ac-powered converters)	RRIO: rail-to-rail input and output
PDF: phase-frequency detector	RRO: rail-to-rail output
PFM: pulse-frequency modulation	RTC: real-time clock
PGA: programmable gain amplifier	RTD: resistance temperature detector (or resistive temperature device)
PID: proportional-integral-differential (in control systems); program identifier (in digital TV)	RTI: referred to the input
PIN: positive-intrinsic-negative (diode)	RTL: resistor-transistor logic; register-transfer level (in an HDL)
PI/PO: parallel-in-parallel-out	RTO: referred to the output
PIR: passive infrared (detector)	RTS: request to send (in a serial link)
PIV: peak inverse voltage	SA: sense amplifier
PLA: programmable logic array	SACD: Super Audio compact disc
PLC: powerline cycles	SAD: silicon avalanche device (i.e., a zener TVS)
PLD: programmable logic device	SAR: successive approximation register
PLL: phase-locked loop	SAS: serial attached SCSI (interface)
pMOS: p-type metal-oxide semiconductor	SATA: serial ATA (interface)
PMT: photomultiplier tube	SAW: surface acoustic-wave
POF: plastic optical fiber	SBC: single-board computer
POL: point-of-load	SC: subscriber connector (a fiber-optic connector)
pp: peak-to-peak (voltage)	SCPI: Standard Commands for Programmable Instruments
PPS: polyphenylene sulfide (a capacitor dielectric)	SCR: silicon-controlled rectifier
PRAM: phase-change random-access memory	SCSI: small computer system interface
	SD: secure digital (memory card)
	SDI: serial data in

<sup>1</sup> Whose awkwardness spawned jokes like “Personal Computer Manufacturers Can’t Invent Acronyms.”

SDO: serial data out  
 SDR: single data rate (memory)  
 SDRAM: synchronous dynamic random-access memory  
 SDTV: standard-definition television  
 SE: single-ended  
 SED: surface-conduction electron-emitter display  
 SEPIC: single-ended primary-inductance converter  
 SERDES: serializer–deserializer  
 S/H: sample-and-hold  
 SHV: “safe high voltage” (connector)  
 SI: serial input  
 SIP: single in-line package  
 SMA, SMB, SMC: subminiature RF coax connector series  
 SMI: small media interface (a fiber-optic connector)  
 SMPS: switch-mode power supply  
 SMT: surface-mount technology  
 SMU: source-measure unit  
 SNR: signal-to-noise ratio  
 SO: serial output; small-outline (IC package)  
 SOA: safe-operating area  
 SODIMM: small-outline dual-in-line memory module  
 SOIC: small-outline integrated circuit  
 SOT: small-outline transistor  
 SPDIF: Sony–Philips Digital Interconnect Format (for digital audio)  
 SPICE: “simulation program with integrated circuit emphasis” (analog circuit simulator software)  
 SPDT: single-pole double-throw (switch)  
 SPI: serial peripheral interface (a simple IC bus)  
 SPL: sound pressure level  
 sPLD: simple programmable logic device  
 SPST: single-pole single-throw (switch)  
 SPTS: single-program transport stream (in digital TV)  
 SR: slew rate  
 SRAM: static random-access memory  
 SSD: solid-state drive (an NV memory)  
 SSH: secure shell (a network protocol)  
 SSP: synchronous serial port  
 SSR: solid-state relay  
 ST: straight tip (a fiber-optic connector)  
 STB: set-top box (for cable or satellite TV)  
 STP: shielded twisted pair (cable)  
 SWR: standing-wave ratio (on a transmission line)  
 T&M: test and measurement  
 TAC: time-to-amplitude conversion  
 TBH: take-back-half (a control algorithm)  
 TCP: transmission control protocol (an Internet protocol)  
 TCXO: temperature-compensated crystal oscillator

TD: transmit data (in a serial link)  
 tempco: temperature coefficient  
 THD: total harmonic distortion  
 TI: Texas Instruments  
 TNC: threaded Neill-Concelman (connector)  
 TO: transistor outline (e.g., TO-92, TO-220)  
 TSSOP: thin-shrink small-outline package  
 TTL: transistor–transistor logic  
 TVS: transient voltage suppressor  
 TWI: two-wire interface (a serial bus)  
 UART: universal asynchronous receiver-transmitter  
 UDP: user datagram protocol (an Internet protocol)  
 UHF: ultrahigh frequency; also a legacy coaxial connector  
 UL: Underwriters Laboratories (a safety certification company)  
 UPS: uninterruptible power supply  
 USB: universal serial bus (a data interface)  
 UTP: unshielded twisted pair (cable)  
 UV: ultraviolet  
 VBR: variable bitrate (coding)  
 VCO: voltage-controlled oscillator  
 VCVS: voltage-controlled voltage-source (active filter)  
 VCXO: voltage-controlled crystal oscillator  
 VDE: Verband der Elektrotechnik, Elektronik und Informationstechnik (a German organization whose activities include safety standards)  
 V/F: voltage-to-frequency (converter)  
 VFB: voltage feedback  
 VFD: vacuum fluorescent display  
 VGA: video graphics array ( $640 \times 480$  analog video)  
 VHS: video home system (video recording)  
 VLSI: very-large-scale integration  
 VME: VERSAmodule Eurocard Bus (a card interface)  
 VOD: video-on-demand  
 VOM: volt-ohm-millimeter  
 VSWR: voltage standing-wave ratio (on a transmission line)  
 VU: volume unit (an audio level)  
 WL: write latency (in computer memory)  
 X7R: a ceramic dielectric  
 XLR: a professional audio connector series  
 XO: crystal oscillator  
 Y5V: a ceramic dielectric  
 YIG: yttrium–iron garnet  
 Z5U: a ceramic dielectric  
 ZCS: zero-current switching  
 ZVS: zero-voltage switching

# INDEX

## KEY

<b>page number font</b>	<b>boldface</b>	main subject treatment
	<i>italic</i>	figure
suffixes	ff	“and pages following”
	g	graph
	p	photograph
	s	screen shot (‘scope or spectrum analyzer)
	t	table

## A

abbreviations, 1166  
*ABEL*, *see* programmable logic device (PLD)  
 absolute-value circuit, 257  
*abuse*, *see* argument  
 ac power control  
     microcontroller, **1062–1065**  
 ac powered  
     dc power supply topologies, 630  
 accelerometer, 1078, 1079p, 1082–1084  
 acoustic ranging, 207  
 acronyms, 1166  
 active-HIGH, active-LOW, *see* logic  
 Adafruit, 1082  
 ADC, *see also* converter, AD/DA, **900–956**,  
     1061, 1082  
     ac monitor, 943–944  
     bridge-sensor oriented, 945–946, 947  
     choosing, **938–940**  
     comparison, **938–940**  
     data-acquisition system (DAQ), **946–955**  
         isolated SPI, 950–951  
         multiplexed 16-channel SAR, 948,  
**946–950**  
         parallel 8-channel  $\Delta\Sigma$ , 953, 954, **952–955**  
         parallel 8-channel SAR, 951, 952,  
**950–952**  
     delta-sigma, *see* delta-sigma  
     design examples, **946–955**  
     differential-input, 380  
     flash, 903, **902–907**  
         driving (design example), 906, 904–907  
         folding, 904  
         for thermocouple, 1084  
         half-flash, 904  
     integrating, 902, **912–940**  
         as lowpass filter, 940–941  
         dual-slope, 902, **914–916**  
         multislope, 902, 919, **918–921**, 921t  
         normal-mode rejection, 915

## ADC (cont.)

powerline rejection, 915  
 single-slope, 902, 914, **914**  
 introduction, 879–880  
 micropower, 916t, **941–942**  
 noise  
     filtering out-of-band, 911  
     sampling vs integrating, 940–941  
 parallel, **902–907**  
 resolution  
     effective, 935  
     noise-free, 935  
 RF  
     driving (design example), 906, 904–907  
 selected, 905t  
 sequencer, 945  
 shootout, 939t  
 specialty, 942t  
 subsystem, 1084  
 successive approximation, 902, 908s, 910t,  
**908–913**  
     charge-redistribution, 950  
     design example, 912, 910–913  
 tables, 905t, 910t, 916t, 921t, 935t, 937t,  
 939t, 942t  
 techniques, 902–903  
 timing example, 1088–1089  
 tracking, 909  
 transformer input coupling, 382  
 undersampling  
     deliberate, 907  
     V to f, 902, 912–913  
 adder, *see* logic  
 admittance, 69, 90  
 alias, 901, 931  
     digital oscilloscope, 1164  
     in DACs, 900  
     undersampling conversion, 907  
 AM radio, 55, 56s  
 amp, 1

## amplifier

autonulling (design example), 297  
 input errors, 306  
 balanced audio, 376  
 bandwidth vs flatness, 385g  
 class-A, 106–107  
 class-AB, 108  
 class-B, 108  
 class-D, 109s  
 common-emitter, **87ff**  
     biasing, **95ff**  
     bootstrapping, 111  
     bypassed emitter resistor, 96  
     emitter resistor as feedback, 96  
     gain of, 94  
     input impedance of, 95  
     maximum gain of, 98  
     nonlinearity, 94, 95, 96s  
     revisited (Ebers–Moll), 93  
 composite, 332, 543–547, 920  
     slew rate of, 920  
 current, *see* transimpedance amplifier  
 current-sensing, 278  
 decompensated, 515  
 difference, 353t, **347–356**  
     as current source, 350–351  
     CMRR, 355g  
     CMRR trim, 356  
     common-mode input range, 354  
     filter node, 355  
     offset trim, 355  
 differential, 375t, **372–380**  
     as comparator, 105  
     as dc amplifier, 104  
     biasing, 104  
     BJT, **102ff**  
     choosing, **383–387**  
     CMRR, 376, 383  
     common-mode input range, 376,  
 383

- amplifier, differential (cont.)  
 compensation, 379  
 current-mirror load, 105, 153, 496  
 distortion, 386g  
 gain of, 103  
 gain-setting resistor, 376  
 high-voltage, 209, 607, 694  
 input impedance of, 376  
**JFET, 152ff**  
 noise, 386, 496, 520  
 offset voltage, 383  
 output common-mode voltage, 376  
 phase splitter, 105  
 slew rate of, 385  
 speed, 383  
 supply voltage, 383  
 use with ADCs, 380–382  
 use with single-ended input, 376  
 equalization, 381, 870–871  
 error  
   in linear regulator, 597  
   in switchmode regulator, 646, 651, 651–655  
 fully differential, *see* amplifier, differential  
 gain, predictability of, 117  
 hybrid  
   low-noise, 534  
 input errors, 301  
 input impedance  
   effect of feedback on, 118  
 input protection, 210  
 instrumentation, *see* instrumentation  
   amplifier  
     JFET-input, 512  
 isolation, 583ff, 585ff  
   capacitive, 586  
**JFET, 146ff**  
   differential, 152  
   low noise, **509–520**  
   op-amp hybrid, 152, 153g, 155g, 343–347, 534  
   series-feedback pair, 150, 151g  
   with cascode, 148  
 laboratory, general-purpose, 274  
 logarithmic, *see* AoE 2nd ed., pp 212–216  
 noise, *see also* noise  
   total, vs  $R_s$ , 526g  
   total, vs frequency, 531g  
 op-amp, *see* op-amp  
   BJT hybrid, 534  
   JFET hybrid, 153g, 545  
   low-noise, 522–524t  
   single-supply, 261  
 operational transconductance, *see* OTA  
   (operational transconductance amplifier)  
 output errors, 307  
 output impedance  
   effect of feedback on, 119  
 patch clamp, 552  
 photodiode, 234, 548  
 amplifier (cont.)  
   photomultiplier, 843  
   piezo, 208, 209  
   programmable gain (PGA), 948  
   programmable-gain (PGA), 370, 371t  
   sense, 1019  
   STM, 553  
   summing, 234  
   switchable gain, 182  
   switching, 109s, 673  
   transformer in feedback, 285, 536  
   transimpedance, *see* transimpedance  
   amplifier  
   video and RF, 274, 533  
   wideband analog link, 353, 380  
   world's quietest, 505, 507g  
 amplitude  
   rms, 1  
 analog  
   vs digital, 703  
 analog switch, 176t  
   4053-style, 917t, 916–918  
   charge injection, 176t, 180, 181g  
   vs  $R_{ON}$ , 182g  
**FET, 171–184**  
   applications, 182–184  
   capacitance, 178–181  
   charge injection, 180  
   CMOS, 172  
   feedthrough, 179g  
   JFET, 172  
   latchup, 174  
   limitations, 174  
   multiplexer, 173  
   ON resistance, 171, 173g, 175–178  
   protected, 175  
   speed, 178  
   T-switch, 179  
   in conversion applications, **916–918**  
   logic-family, 916–918  
   transmission gate, *see* logic  
 analog-to-digital, *see* ADC  
 aperture time, 940  
 arbitrary function generator, 18  
 Arduino, 1086, **1092**, *see also* microcontroller  
 argument, *see also* abuse, *see* computer  
 ARM, *see* microcontroller  
 Art of Electronics  
   logo wristwatch, 448p  
 ASCII code, 1039, 1040t  
 assertion-level logic, 713  
 asynchronous, *see* logic, *see also* memory  
 attenuator, 17  
   matched impedance, 1123, 1124t  
 audio  
   analog, 1131  
   balanced driver, 376  
   sampling rate, 902  
 autocorrelation  
   of PRBS, 978  
 autozero, *see* op-amp  
 avalanche, 199, 979, 1069  
 AVR, *see* microcontroller  
**B**  
 bandgap, *see* voltage reference  
 bandwidth  
   narrowing, **575ff**  
   lock-in detection, 398, 562, 576,  
**575–578**, 851  
   synchronous detection, 851  
 of noise, 477  
 of op-amp, 308  
 signal averaging, 576  
 base  
   spreading resistance, *see*  $r_{pb}$   
 battery, 2, **686ff**, 688p  
   backup, 36  
   characteristics, 687, 689t  
   choosing, 688  
   discharge curves, 687g  
   energy density  
    vs capacitor, 690g  
   energy storage  
    vs capacitor, 690t  
 Li-ion  
   charging, 688  
 power density  
   vs capacitor, 690g  
 primary, 686  
 rechargeable, *see* battery, secondary  
 secondary, 686  
   charging, 687  
   series resistance of, 9  
 BCD, *see* number  
 beast, the number of the, 666  
 bench instruments favorites, 1152  
 beta, *see* BJT  
 bias  
   tee, 836  
 biasing  
   bad, 85  
   of emitter follower, 83  
   with compensated  $V_{BE}$  drop, 97  
   with dc feedback, 98  
 bibliography, **1154ff**  
 binary  
   number, *see* number  
   search, 908  
 biology  
   electrical engineer's view of, 805  
 biphasic code, 1041  
 bipolar transistor, *see* BJT  
 bit  
   effective number of, 928, 929  
   grooming, 1138  
 bitstream, *see* delta-sigma  
**BJT (bipolar junction transistor), 71ff**  
   as transconductance amplifier, 91  
   base resistance, *see*  $r_{bb}$   
   basic circuit topologies, 90  
   basic circuits, 91

- BJT (cont.)  
 beta, 72, 74g, 504g, 501–504t  
 current noise, 483, 493  
   vs current, 484g  
   vs frequency, 484g  
 current ratio, 102g  
 current sharing in, 112  
 Darlington, 109ff, 1081  
   beta vs current, 110g  
 Ebers–Moll model, 90  
 emitter follower, *see* emitter follower  
 emitter impedance of, 92  
 gain-bandwidth product, 549g  
 inductive load, 75  
 logic with, 124  
 low-noise, 501–502t  
   challenge, 507g, **505–508**  
   replace with JFET, 500  
   selecting, **500–505**, 507  
 low-noise design with, **492–509**  
 noise, 481  
   from NF, 489–492  
   model, 488  
   resistance, 494  
   test circuit, 557  
   vs JFET, 517  
 paralleling, 112  
 pinouts, 72, 502  
 power, 106t  
 pulse generator, 458, 459s  
 $r_{bb}$ , 481ff, 483g, 488, 501t, 503, 505ff  
 representative, 74t  
 rules of thumb, 91  
 saturation current,  $I_S$ , 91  
 series connection of, 697  
 simple model, 72  
 superbeta, 111, 252, 302t, 323, 326  
 Sziklai connection, 110ff, 148, 207, 214,  
   229, 598, 608  
 table of, 74t, 106t, 501t  
 terminal notation, 71  
 $V_{BE}$  tempco, 92  
 voltage noise, 481, 493  
   calculation (circuit example), 486  
   charting, 494  
   from  $i_n$  through  $R_s$ , 483  
   vs  $r_{bb}$ , 483g  
   vs current, 482–483g  
   vs frequency, 485g  
   vs power MOSFET and IGBT, 201, 202t,  
   208t  
 Black, Harold, 116, 118, *see also* feedback  
 Blackman–Harris, *see* window, sampling  
 Bluetooth, 1061  
 Bode plot, **280–285**, 311, 899, 961, 963  
 boost, *see* switchmode  
 bootstrap, **111ff**  
   cascode, in TIA, 550  
   leakage cancellation, capacitor, 897  
   of op-amp follower, 233
- bootstrap (cont.)  
   of power supply, 359  
   of TIA, 547  
   twin-T filter, 414
- branch  
   you're sitting on, sawing off, 1091
- breakdown, *see also* second breakdown  
   base-emitter, 82  
   MOSFET gate, 199
- brickwall, *see* filter
- buck, *see* switchmode
- buffer  
   logic, 23, 218t, 798g, 817  
   unity-gain, 311, 843
- bus, *see also* computer, 1029t  
   multidrop, 990  
   point-to-point, 990
- byte, 990
- C**
- cable  
   cat-5, 864  
   cat-6, 864  
   coax, **1116–1123**, **1126**  
     counterwound helix, 1130p, 1130s  
     driving with logic, **858–864**, **1120**  
     reflections, **858–864**, 1116s, 1118–1119s  
     terminating, **858–864**, **1117**  
   differential, **864–874**  
   digital signals through, **856–874**  
   driving, **856–874**  
   equalization, 864  
   pre-emphasis, 864  
   shielded, 587, *see also* cable, coax  
   STP, 864  
   television  
     analog, 1134  
   termination  
     back, 860, **1118–1120**, **1122**  
     double-ended, 860, 1122  
     series, 860, **1118–1120**  
   twisted-pair, **864–874**  
   UTP, 864
- calculus, 1099
- CAN bus, **1043**, *see also* computer, data bus, 1061, 1081
- capacitance  
   disappears in terminated cable, 860, 1116  
   feedback, 113, *see also* Miller effect  
   gate, 197g  
   junction, 114  
   multiplier, 508, 557, 578, 579g  
   of analog switch, 178–181  
   sensor, 1086
- capacitive  
   coupling, 581  
   isolator, 586
- capacitor, **18ff**, 20p  
   “X” and “Y” types, *see* capacitor, powerline rated
- capacitor (cont.)  
   blocking, 19, 43  
   bypass, 19, 54  
   ceramic  
     ESR and stability, 616  
   charge pump, 638  
   current through, 19, 46  
   dielectric absorption, 28, 211, 298, 300,  
   301g, 326, 422, 915  
   digital, 1082  
   electrolytic, 20p, 232, 301g, 392, 580, 597,  
   633ff, 635, 661, 690t, 856  
   dc coupling eliminates, 496  
   energy density  
     vs battery, 690g  
   energy storage  
     vs battery, 690t  
   energy storage in, 19, 686
- ESR  
   filter attenuation, effect on, 393g  
   increased ripple from, 639
- film, 19, 20p, 279, 300, 326, 422, 438, 631p
- flying  
   driving MOSFETs, 822, 847  
   in power conversion, 184, **638ff**
- HV discharge circuit, 211
- leakage, 300  
   bootstrap cure, 897
- leakage compensation, 298
- multiplier, *see* capacitance
- mylar, 19, 20p
- non-ideal in filter, 392
- parallel, 21
- piezoelectric effect in, 682
- power density  
   vs battery, 690g
- powerline rated, 631, 631p, 664p, 671p  
   discharge resistor, 631
- RC circuit, 21
- reactance of, 42, 45, 49g
- reactive current, 47g
- ripple current, 634, 635s, 659, *see also* current
- series, 21
- speedup, 808s
- storage, 55  
   ripple current, 634
- storage, in power supply, 633
- switched, filter, **415–418**
- timing, 55
- types, 20
- variable, 64
- CAS, 1019
- cascode, 102, 114, 115, 146, **148ff**, 345, 369,  
   377, 534, 545, 548, 550
- folded, 149, 377, 498, 536
- regulated (RGC), 552
- regulator, for HV, 693, 698  
   series, for HV, 697
- catalogs, 1153
- cathode ray tube (CRT), 1143

- CD audio, 902  
 ceramic resonator, 450  
 cesium, *see* oscillator, atomic  
 chain rule, 1099  
 charge  
     integrator, *see also* AoE 2nd ed., pp 640–643  
     integrator (Coulomb counter), 933  
     MOSFET gate, 197g, 198s  
     pump, 638, *see also* capacitor, flying, *see also* switchmode  
     redistribution  
         ADC, 950  
         DAC, 909  
 charge injection  
     absence of, 843, 849  
     in ADC, 911  
     in analog switch, 176t, 180, 181g, 301, 948  
         vs  $R_{ON}$ , 182g  
     in integrator, 918  
     in sample-and-hold, 256  
     in switched-capacitor filter, 418  
 choke, 29, *see also* inductor  
     common-mode, 584  
 cigar, *see* cigar  
 circuit  
     BJT  
         basic, 91  
     breaker, 630  
     diagram  
         how to draw, 1101ff  
     integrated, *see* IC  
     linear, 14  
     loading, *see* loading  
     parallel, 2  
     RC  
         discharge waveform, 21g  
         Thévenin simplification, 23  
     resistor  
         shortcuts, 6  
     resonant, 52ff  
     series, 2  
         parallel, 5  
         resistor, 5  
 clamp, *see also* SCR  
     active, 257  
     active reset, in SMPS, 657  
     diode, 36  
     input transient, in SMPS, 669  
     op-amp output, 913  
     power supply, 692  
 clearance, 663  
 clipping, 81  
 clock  
     generation with PLL, 972  
     real-time, 1084  
     recovery, 1037–1042  
     skew, 757  
 CML (current-mode logic), 874  
 CMOS, *see* analog switch, *see* logic
- CMRR, 103  
     at high-frequency, preserving, 360  
     ground-loop isolation, 583  
     instrumentation amplifier, 364g  
     trim, 360  
 coax, *see* cable  
     digital signals through, 1116–1122  
     driving, 1116–1122  
 coil  
     driver  
         precision, 898, 897–899  
 Coldfire, *see* microcontroller  
 common-mode, *see* amplifier, *see* op-amp, *see also* comparator  
     pickup, 584, *see also* CMRR  
     torture test, 869s  
 common-mode rejection ratio, *see* CMRR  
 comparator, 24, 236, *see also* Schmitt trigger, 812t, 813t, 809–817  
 window discriminator, *see* AoE 2nd ed., p 669  
     cautions, 816  
     digital magnitude, 728  
     driving logic from, 806–808  
     hysteresis  
         internal, 815  
     input  
         common-mode range, 812  
     input current, 814, 814g  
     output current, 811, 811g  
     output swing, 810  
     response time, 816g  
     saturation voltage, 811g  
     schematic, LM393, 270  
     speed, 815  
     supply voltage, 815  
     voltage  
         differential, 815  
         offset, 812  
 compensation  
     active  
         of op-amp phase error, 314, 315g, 315t  
         dominant-pole, 282  
         of transimpedance amplifier (TIA), 537, 541g, 545  
         pole-zero, 284  
 complex  
     conjugate, 1098  
     numbers, 1097  
     plane, 1098  
 compliance, 8, *see also* current source  
 components  
     best avoided, 63p  
     confusing markings, 65p  
     itty-bitty, 65p, 269, 627p, 713, 821  
     passive, 56ff  
     where to buy, 1150  
 composite amplifier, *see* amplifier, composite computer, *see also* microcontroller  
     architecture, 990–992  
     Harvard, 991  
     von Neumann, 991  
     assembly language, 993, 1086  
     BASIC, 1086  
     bus-oriented, 991  
     C/C++, 1086  
     CPU, 990  
     data bus, 990, 1029t  
         1-wire, 1035–1036  
         ADDRESS, 992, 997  
         CAN bus, 1043–1045  
         control lines, 992  
         DATA, 992, 997  
         eSATA, 1037  
         Ethernet, 1045  
         FireWire, 1042  
         GPIB, 1031  
         I<sup>2</sup>C, 1034, 1034–1035  
         JTAG, 1036, 1036  
         parallel, 1028  
         PATA, 1031  
         PC104/ISA, 992  
         PCIe, 1037  
         PLD address decoder, 1000  
         printer port (Centronics), 1031  
         SAS, 1037  
         SATA, 1037  
         SCSI, 1031  
         serial, 1032–1037  
         signals, 997–1013, 1013t  
         signals, summary of, 1012  
         SPI, 1032, 1033, 1032–1034  
         STROBE, 997  
         USB, 1042  
     direct memory access (DMA), 1010–1012  
     flags, 990  
     instruction, 990  
     instruction decoder, 990  
     instruction set  
         addressing modes, 994–995  
         argument, 994  
         subroutine, 996  
         x86, 1008  
         x86, simplified, 994t, 993–994  
     interrupt, 1005–1010  
         acknowledgment, 1009  
         autovectored, 1009  
         handler, 1006–1008  
         mask, 1010  
         polling, 1008  
         shared, 1008  
         software, 1010  
     Java, 1088  
     keyboard bus interface, 1003, 1005  
     machine language, 993  
     memory, 991, *see also* memory  
         cache, 991

- computer (cont.)  
 number storage in memory, 1048  
 object code, 993  
 PC104, 997p, 1013, 1014, 1015p  
 port, 992  
 program  
   counter, 990  
   example, 996  
 programmed IO, **998–1005**  
   command register, 1004  
   data in, 1001  
   data out, 998, 999  
   read cycle, 1001  
   status bit, 1002  
   status register, 1002  
   write cycle, 998  
 progress in, 990p  
 Python, 1088  
 real-time IO, 992  
 register, 990, 994–995  
 SBC, 997p, 1013  
 software interrupt, 1010  
 stack, 995  
   pointer, 990  
 subroutine, 996  
 terminology, 989  
 UART, *see* UART  
 word, 990  
 xy vector display, 1000
- conductance, **6**
- conduction  
   angle, 635
- confusion central, 65p
- connector  
   avoid these, 63p  
   BNC, 56, 59, 62p  
   card-edge, 61  
   circular, 61p  
   DisplayPort, 1145  
   DVI, 1145  
   HDMI, 1144  
   multipin, 61  
   photographs of, 60–63p, 1144p  
   rectangular, 60p  
   RF and shielded, 62p  
   shielded-cable, 59  
   video, **1143ff**, 1144p
- converter, *see also* ADC; DAC; power  
 AD/DA  
   errors, 881  
   linearity, **899–900**  
   performance parameters, 879–880
- power  
   boost, 30  
   buck, 30  
   charge-pump, 183  
   flying-capacitor, 183
- correlated double sampling, 553, 570
- coulomb, 18, *see also* charge, integrator
- counter, *see also* logic  
   timing with, 465
- cPLD, *see* programmable logic device (PLD)
- creepage, 665
- crest factor, 568, 981
- crossover, input, *see* op-amp, rail-to-rail, input  
   crossover
- crossover distortion, *see* distortion
- crosstalk, 581, 588
- crowbar, *see also* voltage regulator  
   overvoltage, 598, **690ff**, 691
- crystal, *see* oscillator, crystal
- CUPL, *see* programmable logic device (PLD)
- current, **1**  
   class-A, 754  
   comparator input, 814, 814g  
   comparator output, 811, 811g  
   critical, in SMPS, *see* switchmode  
   feedback, *see* feedback, current  
   input  
     of digital logic, 795  
     of instrumentation amplifier, 362  
     of op-amp, 244  
   inrush, 661, 668  
   Kirchhoff's current law, 2  
   limiting, 597  
     foldback, 693, 694g, 824, 824g  
     of outbound pass transistor, 695  
   load-sensing, 277  
   loop, 1084  
   mirror, *see* current mirror  
   mode  
     PWM, in SMPS, 652, *see also* switchmode  
   noise, *see* noise  
   pulsation in SMPS, 649  
   reactive, 41  
   ripple, 634, 635s  
     in boost converter, 647  
     in bridge converter, 659  
     in buck converter, 644  
     in storage capacitor, 634  
     in zener bias, 596, 598  
   sensing, 944  
   shoot-through, 186, 660, 760, 760g, 856,  
     942  
   shunt, 277, 944  
   source, *see* current source  
   spy, 754  
   transformer, 944  
   transient, output-stage, 856
- current mirror, **101ff**  
   active load, 105, 153, 496  
   compliance, 101, 104  
   Wilson, 102, 146
- current source, **85ff**, **620ff**, 623g  
   3-terminal regulator as, 620  
   bandwidth of, 254  
   BJT, 85–88, 146, 534
- current source (cont.)  
   compliance, 8, 27, 28, 86, 895  
   deficiencies of, 87  
   depletion-mode  
     FET as, 211, 622, 623g, 696  
   difference amplifier as, 350–351  
   discrete, 622  
   floating, 895, 895–897  
   high voltage, 622, 696  
   ICs, 621  
   instrumentation amplifier as, 367  
   JFET, 142  
   JFET and BJT compared, 145  
   nanoamp  
     programmable wide-compliance, 896  
     pulsed, 555  
   nanoamp wide-compliance, 894–897  
   noise of, **487–489**  
   op-amp, **228–230**, 242, 254, 344, 367, 623,  
     895  
     Howland, 229, 230  
   oscillation in, 442  
   precision, 367, 898, **897–898**  
   resistor as, 85  
   slew rate of, 367  
   transistor, *see* current source,  
     BJT/JFET/depletion-mode  
   curve tracer, 1115

**D**

- DAC, *see also* converter, AD/DA, **881–900**, 1061, 1081  
 1-bit, *see* delta-sigma  
 application examples, **891–899**  
 audio, 939t  
 charge redistribution, 909, 909  
 choosing, 891  
 code-change glitch, 892  
 current-steering, **884**, **883–886**  
   generating voltage output from, **885–886**, 886  
 delta-sigma, *see* delta-sigma  
 exemplary, six, 887, 886–888, 889t  
 for offset correction, 948  
 for zero correction, 948  
 frequency-to-voltage, 890  
 introduction, 879–880  
 multiplying, **884–885**, 894t  
 PWM as, 888  
 PWM to, converter, 889  
 R–2R, **882–883**, 883  
 resistor string, **881–882**, 882  
 selected, 893t  
 tables, 889t, 893, 894t  
 Darlington, *see* BJT  
   photo, 842  
 data bus, *see* bus, *see* computer  
 data-acquisition system, *see* DAQ  
 datasheet, too big, 1063

dB, 15ff  
 dB per octave, 51–52  
 dc  
     blocking capacitor, 43  
     power switching, 202  
     supply splitter, 262  
 dc-dc converter, *see* power supply, *see also* switchmode  
 DDS, *see* oscillator, DDS  
 debugging  
     microcontroller, in-circuit, 1092  
 decibel, *see* dB  
 decoder, *see* logic  
 delay equalizer, 415  
 delay line  
     lumped element, 1126, 1128p, 1129  
         impedance of, 1127  
     transmission line, 1128p, 1130s  
 delta-sigma, 902, 915–916  
     20-bit ADC, 934  
     24-bit ADC, 934–936, 937  
 ADC, 935t, **922–940**  
     digital filter in, 924, 926  
 application examples, **932–940**  
 audio ADC, 937t  
 block diagram, 924  
 charge integrator, 933  
 DAC, 930–931, 938–939  
 demystifying, **923–931**  
 dynamic range, 925, 929  
     vs OSR and *m*, 929g  
 ENOB, 928, 929  
 fast multichannel IC, 955, 956  
 idle tones, 931–932  
     mitigating, 932  
     simulation, 933g  
 “industrial” ADC, 936, 937  
 magic in the modulator, 927  
 microcontroller, implemented with, 932–933  
 modulator, 925, 925–929  
     higher order, 928, 929  
 monotonicity, 931  
 noise shaping, 927g, **927–931**  
 powerline rejection, 954–955, 955  
 pro-audio ADC, 936–938, 940  
 pros and cons, **931–932**  
 simplest, 922  
 simulation, 929g, 930g, 931g, **928–931**  
 spectrum, 931g  
 suntan monitor, 922–923  
 the paradox, 923–924, 926  
 time delay, 931  
     vs the competition, **938–940**  
 derivative, 1099  
 design review, 1059  
 Designs by the masters  
     instrumentation amplifier, 361, 359–362  
 Keysight’s multislope converters, 919,  
     918–922

Designs by the masters (cont.)  
     nanoamp programmable current source, 896  
     paralleling MOSFETs, 214  
     powerline SMPS, 595p, 666, 667s, 665–671  
     precision DMM, 342–347  
     SR560 preamp, 514, 512–515  
 desoldering station  
     favorite, 1152  
 detector  
     optical, **840–843**  
 device programmer, 769  
 dielectric  
     absorption, 300–301, 301g  
     constant, 18  
 differential, *see* amplifier  
     cable, **864–874**  
     pseudo, 584  
 differentiation, 1099  
 differentiator  
     bridged, 415  
     op-amp, 260  
     RC, 25, 26s, 51  
 digital, *see also* logic  
     conquest over analog, history of, 703  
     filtering, 419  
     logic, *see also* gate  
     multimeter, 10  
     potentiometer, 63  
     signal processing (DSP), 418–422  
         sampling, 419  
 television  
     broadcast, 1138  
     cable, 1138  
     satellite, 1139  
     spectrum, 1136s  
     vs analog, 703  
 digital-to-analog, *see* DAC  
 diode, 31ff, 32  
     AND gate, 711  
     as log converter, 37  
     clamp, 36  
     current-regulator, 143g  
     drop compensation, 35  
     forward current, 294–295, 295g  
     forward drop, 31  
     gate, 36  
     *I* vs *V*, 92g  
     inductive clamp, 38  
     JFET as, 294  
     LED, *see* LED  
     limiter, 37  
     MOSFET body, 199  
     photo, 841p, 841–842  
         amplifier, 548  
     PIN, 837p  
     protection array, 805  
     representative, 32t  
     Schottky, 32, 614, 632, 641, 669, 790  
     tunnel, 1113  
 diode (cont.)  
     varactor, 64, 440, 960, 969, 971  
     capacitance vs voltage, 441g  
     zener, 12ff, 82, 674, *see also* voltage reference  
     biasing, 675  
     drift, 675  
     dynamic resistance of, 13  
     IC, 676  
     low voltage, 13g  
     noise, 674, 676  
     tempco, 674  
     the astonishing LTZ1000, 675  
 DIP (dual in-line package), *see* IC, packages  
 direct digital synthesis, *see* oscillator, DDS  
 discharge  
     curves, of battery, 687g  
     of high-voltage supply, 211  
     RC, 21  
 discontinued parts, 273 (box)  
 display, 837–840  
     LCD  
         interface timing, 1028  
         smart, 839–840  
         codes, 840  
         xy vector  
         interface, 1000  
 display devices, 62  
 DisplayPort, *see* connector  
 distortion  
     crossover, 309  
         effect of feedback, 235s  
         in push-pull follower, 107  
     from slew rate, 386g  
     of differential amplifier, 386  
     of op-amp, 329–331, 331, 332g  
     rail-to-rail, 318  
     ultra-low, oscillator, 437  
 divider  
     voltage, 7  
 DMM, 10  
     Agilent’s accurate, 342–347  
     favorites, 1152  
 Dobkin, Bob, 604, 679  
 dogs, 63p  
 dominant pole  
     compensation, 282  
 droop, *see* integrator  
 dual slope  
     conversion, *see* ADC  
 dual-tracking, *see* voltage regulator  
 duty-cycle  
     limiting, 465  
 DVI, *see* connector  
 dynamic range  
     of delta-sigma converter, 925, 929  
     of switched-capacitor filters, 418  
     vs precision, 292

- E**
- Early effect, 92
    - in current mirror, 101
  - Ebers-Moll model, 90ff
  - edge, *see also* jitter
    - detector, 26
    - leading, 25
    - trailing, 25
  - EEPOT, 63
  - egg
    - Easter, *see* if you can find all eight
  - 8b/10b code, 1041
  - 8051, *see* microcontroller
  - electrical engineer
    - view of biology, 805
  - electromagnetic interference, *see* EMI
  - electrometer, 297, 570–574
  - electromotive force, *see* EMF
  - electrostatic discharge, *see* ESD
  - EMF, 2
    - back, 28
    - thermal, 340, 675
  - EMI
    - in instrumentation amplifiers, 365
  - emitter
    - optical, 829–840
  - emitter follower, 79–85, *see also* BJT
    - as voltage regulator, 82–83
    - biasing, 83
    - cancelling offset, 85, 87
    - example, 84
    - JFET current-sink pulldown, 144
    - noise of, 487–489
    - output impedance of, 93
    - revisited (Ebers-Moll), 93
    - with split supplies, 84
  - $e_nC$ , *see* transimpedance amplifier
  - energy
    - density
      - battery vs capacitor, 690g
    - in capacitor, 19
    - in inductor, 28
    - storage
      - battery vs capacitor, 690t
  - ENOB (effective number of bits), 900, 905t, 928, 929, 932, 936, 985
  - equalization, 381, 864, 870, 871s
  - equivalent series inductance, *see* capacitor; resistor
  - equivalent series resistance, *see* capacitor error
    - budget, 293, 295–299
    - component, 299
  - eSATA, *see* computer, data bus
  - ESD, 200, *see also* human body model; spike input protection, 259, 362
  - ESL, *see* capacitor; resistor
  - ESR, *see* capacitor
  - Ethernet, 867, *see also* computer, data bus, 1045–1046, 1061, 1081
    - 8b/10b coding in, 1041
    - Manchester coding in, 1040
  - Euler, Leonhard, 1098
  - eye diagram, 1028
- F**
- fan, 627, *see also* AoE 2nd ed., p 858
    - proportional control, 608
  - farad, 18
  - feature size
    - scaling, 794
  - feedback, 115ff
    - bang-bang, 1070
    - BJT examples, 121
    - closed-loop gain, 116
    - current, 118, 150, 308, 378, 527
    - effect on amplifier circuits, 117ff
    - in PLL, 961
    - nonlinear, 276
    - nonlinear control, 1075
    - output-current sensing, 120
    - output-voltage sensing, 117
    - PID, 1069, 1074, 1077
      - circuit, 1075
      - convergence, 1076
    - split, 264, 285, 913
    - stability, 120
      - with inductive load, 899
  - FET, 131ff, *see also* JFET, MOSFET
    - analog switch, *see* analog switch
    - basic circuits, 140
    - depletion-mode
      - as current source, 622
    - dynamic gate current, 164
    - enhancement and depletion, 135
    - family tree, 136
    - $G_{\max}$ , 141t
    - interdigitation, 140
    - JFET and MOSFET, 134
    - leakage, circuit cure for, 259
    - linear region, 137
    - linear switch, *see* analog switch
    - manufacturing spread, 138–140
    - matching, 139
    - saturation current,  $I_{DSS}$ , 134, 136, 139g, 361, 847
    - saturation region, 137–142, 147, 158, 165ff
    - similarity to BJT, 132
    - square law, 137ff, 165
    - subthreshold region, 138, 139g, 166g, 168
    - switch, 132
    - tempco of  $I_D$ , 138g
    - threshold voltage ( $V_{th}$ ), 137ff, 138g, 147
    - transconductance ( $g_m$ ), 132, 141t, 146
    - transfer characteristics, 135g, 137g, 139g, 142g
    - $V-I$  curves, 132, 133g
  - FET (cont.)
    - $V_{GS}$  spread, 140g
  - FEVT (far-end crosstalk), 581
  - fiber optics, 853p
    - ST/SC, 855
    - TOSLINK (EIAJ), 852–853, 854s
    - Versatile Link, 853
  - field-effect transistor, *see* FET
  - filter
    - active, 396–399, 406–415
      - bandpass, 411–413
      - highpass, 408
      - lowpass, 241, 401–408
      - Sallen and Key, 242g
    - allpass, 415
    - analog FIR, 983
    - anti-alias, 395, 901, 902
    - attenuation
      - degraded by ESR, 393g
    - bandpass
      - noise bandwidth, 564t, 565g
    - Bessel, 403
      - VCVS, 408
    - brickwall, 391
    - bridged differentiator, 415
    - Butterworth, 401, 401g
      - LC, 1109ff, 1110t
      - VCVS, 408
    - Chebyshev, 401
      - phase and amplitude, 400g
      - VCVS, 408
        - vs  $RC$ , 396g
    - comparison, 402g
    - delay equalizer, 415
    - digital, 419–422
      - in delta-sigma, 924, 926
      - of PRBS, 979
      - response, 421g
    - eco-friendly, 1073
    - effect of component tolerances, 402g
    - effect of non-ideal capacitors, 392
    - electrically tunable, 412
    - elliptic, 403, 1068
    - equiripple, 401
    - finite-impulse-response (FIR), 419, 421g
    - implementation, 406–418
    - infinite-impulse-response (IIR), 420
      - LC, 54g, 54p, 55s, 393
        - multisection, 394
      - linearity, 422
      - lowpass
        - dc-accurate, 418
        - FIR, 420
        - IIR, 420
      - maximally flat, 401
      - multiple-feedback, 413
        - effect of op-amp  $Z_{out}$ , 414g
      - node of difference amplifier, 355
      - noise bandwidth, 564t, 565g

filter (cont.)  
 notch, 414  
 passive, **391–396**  
 performance parameters, 399  
 phase shift, 404g  
 phase-sequence, 455, 456g  
 power-supply, 32  
 powerline, 631  
*RC* highpass  
 approximate, 43  
 exact, 48  
*RC* lowpass  
 approximate, 42g  
 exact, 50  
 phase shift of, 50g  
 phasor diagram, 52  
 switchable, 182  
*RC*  
 loading by, 44  
 phase shift, 393g  
 response, 392g  
 worst-case impedance of, 44  
 Sallen and Key, 399, 409–410  
 state-variable, 410–413  
 bandpass, 411  
 step response, 405g, 406t  
 comparison, 406, 409g  
 switched-capacitor, **415–418**  
 sine from square, 435, 436s  
 Thomson, 403  
 time delay, 404g  
 twin-T, 414  
 types, 400–406  
 VCVS, 407, 408t  
 effect of op-amp  $Z_{out}$ , 414g  
 finite state machine (FSM), *see* logic  
 FIR, *see* filter, digital  
 FireWire, *see* computer, data bus  
 first-in, first-out, *see* logic  
 flash  
 conversion, *see* ADC  
 flip-flop, *see* logic  
 flyback, *see* switchmode  
 flying  
 amplifier, 367  
 capacitor, *see* capacitor, flying  
 foldback, *see* power supply; voltage regulator  
 follower, *see* emitter follower, source follower,  
 op-amp  
 forward  
 converter, *see* switchmode  
 FPGA, *see* programmable logic device (PLD)  
 frequency  
 analysis of reactive circuits, **41ff**  
 angular, 14  
 compensation, **280ff**  
 in linear regulator, 597  
 synthesis, 451

frequency (cont.)  
 synthesizer, 17, 1082  
 microcontroller, **1065–1069**  
 to voltage converter, 890  
 translation, 562  
 function generator, 17  
 arbitrary, 18  
 favorites, 1152  
 fuse, 630

**G**

GAL, *see* programmable logic device (PLD)  
 gate, *see also* logic  
 delay  
 vs power, 792g  
 vs voltage, 794g  
 logic  
 as pulse generator, 459  
 CMOS, 185  
 speed  
 vs power, 719g, 792g  
 vs voltage, 719g, 794g  
 universal, 715  
 generalized impedance converter, 398  
 op-amp bandwidth, effect on, 399g  
 Gilbert cell, 162  
 glitch, *see also* logic  
 code-change, in DAC, 892  
 glossary, 1166  
 GPIB, *see* computer, data bus  
 GPIO, 1084  
 Gray code, 708  
 ground, 2  
 bounce, 758–759, 856  
 gridded PCB layout, 758  
 loop, 582  
 noise, 856  
 grounding, 579ff  
 between instruments, 583  
 blunders, 582  
 guard, 360, 587  
 Gummel plot, 92g, 500  
 Gummel-Poon, 93  
 gyrator, 397ff

**H**

H-bridge, 655, 659, 889  
 Hamming, *see* window, sampling  
 Hanning, *see* window, sampling  
 HDL, *see* programmable logic device (PLD)  
 HDMI, *see* connector  
 heat, **623ff**  
 heatsink, **624ff**, 625p  
 insulating, 626  
 LED, 835p  
 PCB foil as, 626, 627g  
 sizing, 626g  
 sizzle test, 625  
 thermal resistance of, 626g

henry, *see* inductor  
 heterodyne, 562  
 Hewlett, William, 437  
 hexadecimal, *see* number  
 high voltage, *see also* voltage regulator  
 clearance distance, 663  
 creepage distance, 663, 664, 665p  
 pulse generator, 333, 917, 978  
 hour of power (design example), 466  
 human  
 body model (HBM), 200, 804ff  
 system interface (HI), 1059  
 humidity  
 and MOSFET damage, 200  
 effect on ICs, 684  
 hysteresis, 237, *see also* Schmitt trigger  
 in tunnel diode, 1114  
 thermal, 683  
 hysteretic  
 SMPS, *see* switchmode

**I**

$I^2C$ , 887, *see* computer, data bus, 1061, 1081  
 in DAQ, 954  
 microcontroller peripherals, **1084–1086**

**IC**

discontinued, 273 (box)  
 packages  
 DIP, 3p, 57p, 65p, 73p, 224p, 225, 269,  
 340, 443p, 663, 715p, 720, 758, 765p,  
 792, 861, 1037, 1059  
 SMT, 73p, 224p, 225, 269, 340, 443p,  
 715p, 758, 765p  
 parallel interface, 1028, 1030, *see also*  
 computer, data bus  
 serial interface, **1032–1037**, *see also*  
 computer, data bus  
 IEC power entry, 629  
 IEEE 1394 (FireWire), *see* computer, data bus  
 IGBT (insulated-gate bipolar transistor), 207  
 desaturation, 847  
 impedance, **40ff**  
 characteristic, of transmission line, 1116  
 converter, generalized (GIC), 398  
 op-amp bandwidth, effect on, 399g  
 converter, negative (NIC), 397  
 in parallel, 46  
 in series, 46  
 input  
 of difference amplifier, 352  
 of differential amplifier, 376  
 of op-amp, 245  
 of *LC*, 52  
 of common-emitter amplifier, 88  
 of emitter follower, 80  
 of sources and loads, 79  
 output  
 of op-amp, 309  
 of RRO op-amp, 316

- output (cont.)
  - output, of op-amp, 246
  - small-signal, 81
  - transmission-line matching, 1122
- index, 1171
- inductance
  - leakage, **632–635**, 656, 668
    - energy in, 669–670
    - in powerline transformer, 633
  - magnetizing, 669–670
  - of coil, 28
  - of op-amp output, 250g
  - sensor, 1084
- inductive
  - kick, 38, *see also* spike
  - loads, 38
- inductor, **28ff**, 29p
  - flyback, transformer as, 656
  - reactance of, 44, 45, 49g
  - stored energy
    - in switchmode converter, 636
  - variable, 64p
  - volt-second balance, 30
- input
  - protection, 804–806, 949
- inrush current, *see* current, inrush
- instrumentation amplifier, 273, 297, 363t, **356–373**, 1071
  - as current source, 367
  - autozero, 370
  - CMRR, 363t, 364g, 364–365
  - CMRR trim, 366
  - CMRR vs  $R_{source}$  mismatch, 366g
  - configurations, 357, 359–360, 368
  - EMI, 365
  - in DAQ, 948–950
  - input current, 362
  - miscellany, 362
  - noise, 362
  - offset trim, 366
  - PGA, 370, 371t
  - roll-your-own, 359, 361
- integrated circuit, *see* IC
- integrating
  - conversion, *see* ADC
- integrator
  - droop, 257–260
  - in active filter, 410
  - in quadrature oscillator, 453
  - op-amp, 230, 231s, 257–260
  - $RC$ , 26, 51
  - switched-capacitor, 416
- interdigitation, 140
- interference, 478, 579ff
  - intersymbol, 870
- Internet
  - video streaming, 1140
- interrupt, *see* computer
- interrupter
  - optical, 851
- intersymbol interference, 870
- inverter
  - ac power converting, 673s
  - charge-pump, 183
  - flying-capacitor, 183
  - logic
    - CMOS, 185
    - op-amp, 225
    - optional, 232
  - inverting
    - switchmode converter, *see* switchmode
- IP, *see* programmable logic device (PLD)
- IrDA, 1061, 1081
- ISA, *see* computer, data bus
- isolation
  - barrier in offline SMPS, 663
- J**
- JFET (junction field-effect transistor), *see also* FET, 217t
  - 1/f noise, 510, 517g
  - amplifier, **146ff**
    - common-source, 149
    - gos gain error, 167
    - instrumentation, 512
    - low noise, **509–520**
    - maximum gain  $G_{max}$ , 167
    - with cascode, 148
  - as diode, 294
  - as variable resistor, **161ff**
    - linearizing, 161, 162g
  - capacitance, 170
  - care and feeding of, 346
  - characteristic curves, 143g
  - current noise, **511**
  - current source, 142
    - as BJT follower pulldown, 144
    - self-biased, 143g, 145g
  - deep subthreshold region, 166g
  - differential amplifier, 152
  - dynamic gate current, 164
  - gate current, 163g
    - impact ionization, 164g, 546
  - $G_{max}$ , 141t
  - hybrid amplifier, 152, 153g, 155g, 343–347
    - compensation, 154g
    - vs FET op-amp, 155
  - in Wien bridge oscillator, 437
  - leakage, *see* JFET gate current
  - linear circuits, **142ff**
  - low-noise, 516t
    - selecting, **515–517**
  - mini-table, 141t
  - noise, 509g, **509–520**
    - vs BJT, 517
  - op-amp
    - fast, 155t
- JTAG, 769, 1036, *see also* computer, data bus, 1061
  - bootloader, 1091
- K**
- Karnaugh map, *see* logic
- KCL, KVL, *see* Kirchhoff's laws
- Kelvin connection, 277
- keyboard
  - parallel bus interface, 1003, 1005
- Kirchhoff's laws, 2, 1107
- L**
- L network, 1123

- ladder  
 $R-2R$ , 234  
 replaced by charge redistribution, 909
- lamp, 62  
 laser diode, 830p, 835s, 834–836, 836s, 973–974  
 current, 835s  
 linearizing, 836s  
 laser trimming, 244, 305, 352, 357, 361, 416
- latch, *see* logic  
 latchup  
 in analog switch, 174
- LC*, *see also* filter, *LC*  
 Butterworth filter, 1109ff, 1110t  
 filter, 54g, 54p, 55s  
 impedance of, 52  
 notch, 52  
 tank, 52
- LCD, 1061, 1066, *see* optoelectronics  
 leakage, *see also* capacitor; JFET, *see also* inductance  
 spectral, 420
- Lebowksi, *see* egg  
 LED, 62, 830p, 829–834  
 angular illumination, 832g  
 backlight, 1084  
 BJT driver, 76  
 constant-current drive, 834  
 desklamp, 835p  
 display, 837  
 driver, 1086  
 forward drop, 76, 827g  
 lamp (design example), 833  
 multiplexed display, 751, 751–752  
 panel-mount, 832t  
 pulsed, 466, 834  
 synchronous detection of, 578  
 voltage budget, 752  
 wavelength, 76
- Lehrer, Tom, 1099  
 lightning detector (design example), 497  
 limiter  
 diode, 37
- line driver  
 difference amplifier, 351  
 linear  
 circuit, 14  
 linearity  
 converter, 881, 899  
 INL vs DNL, 899–900
- Little Logic, *see* mini-logic  
 LMOS, *see* mini-logic  
 load  
 capacitive  
 logic, 857  
 op-amp, 263s, 264  
 dump, 618  
 inductive, 38  
 load lines, 156g, 1112ff
- loading, 11ff, 11g, 79  
 lock-in, *see* bandwidth, narrowing  
 lockup, *see* logic, pathology  
 logic  
 0 and 1, 704  
 active-HIGH, active-LOW, 714  
 adder, 728  
 Verilog code for, 729  
 arithmetic logic unit (ALU), 728, 990  
 assertion-level, 713  
 asynchronous, 733  
 buffer, *see* buffer, logic  
 capacitive load, 857  
 clock skew, 757  
 CMOS, 714  
 characteristics, 718–720  
 circuit, 717  
 failure modes, 760  
 minimizing power, 754  
 open inputs with, 760  
 pathology, 760  
 shoot-through current, 760  
 threshold skew, 760  
 weak outputs, 827
- combinational, 708  
 combinatorial, 708  
 counter, 742t, 741–744  
 asynchronous, 733, 734s  
 modulo- $n$ , 748–751  
 modulo- $n$ , timing, 750g  
 ripple, 733, 734s  
 synchronous, 737, 737s  
 data bus, 720, 721  
 decoder, 726  
 Verilog code for, 727  
 delay  
 vs power, 792g  
 vs voltage, 794g  
 demultiplexer, 726  
 digital vs analog, 703  
 divide-by-2, 731  
 divide-by-3, 734–736, 743s  
 asynchronous glitch, 743  
 Verilog code for, 736  
 driving  
 external loads, 721, 817, 817–829  
 from comparator, 806–808  
 from op-amp, 808  
 driving ac loads from, 821, 822  
 driving negative loads from, 821, 822  
 ECL, 722  
 excluded states, 735  
 expansion, 726  
 families, 704, 706t  
 characteristics, 718–720  
 circuits, 791  
 CMOS, 791–794  
 ECL, 790  
 history, 790–794
- logic, families (cont.)  
 input and output, 718  
 life cycle, 715g  
 logic levels, 706 (box), 718  
 RTL, 790  
 speed and power, 719  
 speed vs power, 719g  
 speed vs voltage, 719g  
 supply voltage, 718  
 TTL, 790  
 fan-out, 711, 720, 794–797  
 none in transmission gate, 725
- FIFO, 747  
 flip-flop, 196, 729–733  
 clocked, 730  
 edge-triggered, 731  
 hold time, 732  
 master-slave, 732  
 RS, 729  
 setup time, 732  
 toggling, 732  
 type-D, 731  
 type-JK, 731  
 type-T, 731  
 gate, 708–724, *see also* buffer, logic; Single Gate  
 2-input select, 724  
 AND, 709  
 circuits, discrete, 711  
 circuits, integrated, 717  
 example, 712, 714  
 exclusive-OR, 710  
 interchangeability, 713  
 inverter, 709  
 NAND, 710, *see also* memory, nonvolatile  
 NOR, 710, *see also* memory, nonvolatile  
 NOT, 709  
 OR, 709  
 standard, 709  
 transmission, 172, 725ff  
 XOR, 710, 737, 753, 771ff, 782, 957, 975, 983  
 XOR realization, 723  
 glitch, 738, 743s  
 hardware description language, *see also* programmable logic device, *see* logic, HDL  
 HDL, 711, *see also* programmable logic device  
 HIGH and LOW, 704  
 hold time, 732  
 IC packages, 715p  
 identities, 722t  
 input characteristics, 795–796  
 input current, 795g, 795  
 input overdrive, 806  
 input protection, 804–806  
 interfacing, 790–874  
 between logic families, 799, 798–801

- logic (cont.)
  - Karnaugh map, 723
  - latch, 740
  - level, 17
  - levels, 706 (box)
  - life cycle, 715g
  - lockup, 755
  - magnitude comparator, 728
  - metastability, 732, 733s, 756
  - micropower design, 753
  - mini, 712, 715
  - minimization, 723
  - MOSFET, 184
  - multiplexed display, 751–752
  - multiplexer, 724
    - expansion, 726
  - multiplier-accumulator, 728
  - negative-true, 714
  - nMOS
    - input characteristics, 828
    - interfacing, 826–829
    - output characteristics, 827
  - open-collector, 721
  - open-drain, 721
  - output characteristics, **796–797**
    - voltage vs current, 798g
  - parity generator, 728
  - pathology, **755–760**, 760
    - unspecified rules, 757
  - positive-true, 714
  - power
    - dynamic, 754
    - minimizing, 754
  - priority encoder, 727
  - probing, 808, 809s, 810p
  - propagation time, 710s
  - pulse generator
    - design example, 739
    - with flip-flops and counters, 739
  - race, 724, 738, 756, *see also* runt pulse
  - register, 740, *see also* computer
    - power, 819t
    - shift, **744**
  - removal time, 758
  - reset/supervisor, 756t
  - robust link, 1120
  - RTL, 712
  - runt pulse, *see* runt pulse
  - sequential, **729–745**
    - miscellaneous, 746–748
    - timing example, 750
  - setup time, 732
  - shift register, 1081
    - feedback PRBS, 976
    - RAM as, 745
    - timing, 746g
  - speed, 710s
    - vs power, 792g
    - vs voltage, 794g- logic (cont.)
  - standard
    - combinational ICs, 724
    - gates, 716t
    - sequential ICs, 740
    - when to use, 782, 785
  - startup reset, 755
  - state machine, 734–737
    - Mealy, Moore, 736
  - states, 704, 797g
  - switch, 172, **725ff**, 916ff, 917t
  - synchronizer, 737
  - synchronous, 734
  - syntax, 711
  - three-state, 720
  - through cables, **856–874**, **1116–1122**
    - comparison, 874
    - toggling, 196
  - transfer function, 796g
  - transmission gate, 172, **725ff**, 917t, 1016
  - truth table, 708
  - TTL, 714
    - characteristics, **718–720**
    - circuit, 717
    - noise in, 758
    - output characteristics, 827
  - unspecified rules, 757
  - unused input, 718, 738, 754, 759, 806
    - LVDS trick, 800
  - wired-AND, 722, 1081
  - wired-OR, 721, 722
- loss
  - switching, 662
- loudspeaker
  - as microphone, 486
- low-voltage differential signaling, *see* LVDS
  - LVDS, 705, 864, **868–871**
    - levels, 873
    - power, 874g
    - signal conditioning, 870, 871s
    - torture test, 869s
  - LVPECL (low-voltage PECL), 874
- M**
  - magazines, 1153
  - man
    - transistor, 75
  - Manchester coding, **1039–1042**
  - Mark (RS-232), 871
  - math, **1097ff**
  - MDAC, *see* DAC, multiplying
  - memory, *see also* logic, *see also* computer, *see also* microcontroller, **1014–1027**
    - cache, 991
    - computer, 991
    - configuration, for FPGA, 769
    - dynamic, 991, **1018–1023**
      - asynchronous, 1018, 1020
      - synchronous, 1021
      - timing, 1021, 1022
  - memory (cont.)
    - flash, 991
    - HDD, 992
    - nonvolatile, **1021–1026**
      - EEPROM, **1023–1025**, 1053, 1082, 1084
      - EPROM, 1023
      - ferroelectric, 1022, **1025**
      - flash, 1022, **1024–1053**
      - floating gate, 1022
      - magnetoresistive, 1022, **1026**
      - mask ROM, 1023
      - NAND flash, 1024, *1025*
      - NOR flash, 1024, *1024*
      - phase change, 1022, **1026**
      - PROM, 1023
      - pseudo-static, 1017
    - RAM
      - as shift register, 745
    - ROM
      - in state machine, 735
    - SSD, 992
    - static, 991, **1015–1018**
      - asynchronous, 1016, *1016*
      - synchronous, 1018
      - timing, *1017*
    - static vs dynamic, 1015
    - volatile, 991
      - volatile/nonvolatile, 1014
      - wrapup, 1026–1027
    - metastability, *see* logic
    - meter, 61
    - mho, 6
    - microcomputer
      - terminology, 989
    - microcontroller, 747, 749, 887, **1053–1094**
      - 8051, 1060, *1067*
      - ac power control (design example),
        - 1062–1065**
        - circuit, 1063
        - pseudocode, 1064, *1065*
      - ARM, 1060
      - as timer, 469
      - assembly code
        - PRBS generator, 779, 783
        - AVR, 1060
      - block diagram, 1054
      - C-language code
        - PRBS generator, 782, 784
      - constraints, 1086
      - development environment, **1086–1092**
        - cost, 1092
      - development kit, 1090p
      - downloading code, 1059
      - frequency synthesizer (design example),
        - 1067p, **1065–1069**
        - block diagram, 1066
        - pseudocode, 1067, *1068*
      - how to select, 1094
      - implementing delta-sigma ADC, 932–933

- microcontroller (cont.)  
 in-circuit debugging, 1092  
 in-circuit programming, 1053  
 loading code on, **1089–1092**  
   JTAG bootloader, 1091  
   parallel bootloader, 1091  
   proprietary bootloader, 1091  
   SPI bootloader, 1091  
   UART bootloader, 1090  
   USB bootloader, 1091  
 MSP430, 1061  
 output characteristics, 798g  
 peripheral ICs, **1078–1086**  
   direct connection, *1080, 1079–1082*  
   I<sup>2</sup>C connection, *1085, 1084–1086*  
   SPI connection, *1083, 1082–1084*  
 peripherals  
   on-chip, **1061–1062**  
 PIC, 1060  
 popular families, **1059–1060**  
 PRBS generator, 777ff  
 Rabbit, 1060  
 realtime constraints, **1088–1089**  
 software, **1086–1088**  
   assembly code, 1086  
   BASIC, 1086  
   C/C++, 1086  
   Java, 1088  
   Python, 1088  
 stabilized platform (design example), 1077p,  
**1077–1078**, 1079p  
 block diagram, 1078  
 pseudocode, 1078  
 suntan monitor (design example),  
**1054–1059**  
   C-language code, *1058*  
   circuit, 1055  
   pseudocode, *1056, 1057*  
 terminology, 989  
 thermal controller (design example),  
**1069–1077**  
   block diagram, 1069  
   circuit, 1071  
   control loop, 1074, *1075*  
   pseudocode, *1075*  
 when to use, 785, 1093  
 wrapup, **1092–1094**  
 microconverter, 1070  
 microphone  
   loudspeaker as, 486  
   ribbon preamp, 505  
 micropower  
   digital design, 753  
 microstrip, 858, 1117–1119, *see also* stripline  
 Miller effect, **113ff**  
 Millman's theorem, 6, 1108  
 mini-logic, 715p, 715–719, 725, 772, 800, 916  
 MiniGate, *see* mini-logic  
 mixer, 562
- modified sinewave, 673s  
 modulation  
   amplitude  
    homodyne, 970  
    synchronous, 970  
   amplitude (AM), 55, 56s  
   BPSK, 970  
   delta, 910  
   frequency  
    demodulation with PLL, 969  
    Foster-Seeley detector, 970  
    quadrature demodulation, 969  
   ratio detector, 970  
   slope detector, 970  
   QAM, 970  
 modulator  
   delta-sigma, 925, 925–929  
   magic in, 927  
 monostable multivibrator, *see* one-shot  
 monotonicity  
   converter, 881  
 Monticelli output circuit, 318  
 MOSFET, *see also* FET  
   avalanche rating, 1073  
   construction, 134  
   depletion-mode, 209, 210t  
    as current source, 622  
    as input protection, 210  
   current source, 211  
   extending regulator  $V_{IN}$ , 211, 693  
   HV discharge, 211  
   dynamic gate current, 164  
   floating-gate reference, *see* voltage reference  
   gate capacitance, 197g, **197ff**  
   gate driver, 218t, 861  
   gate threshold, 136  
   handling precautions, 200  
    gate damage, 201p  
   high-voltage  
    creepage barrier, 665p  
   logic switch, 184–186  
   noise, 519  
   power, **187ff**, 187–191t  
    applications, **202ff**  
    body diode, 199  
    drain capacitance, 198  
    gate breakdown, 199  
    gate capacitance, 197g  
    gate charge, 197g, 198s  
    gate protection, 199  
    high-side switching, **202ff**, 204  
    lateral, 214g  
    light-at-night, 206  
    linear applications, 208  
    logic-level, 192, *193*  
    paralleling, 201, 212–215  
    piezo amplifier, 208  
    piezo driver, 207  
    power rating, 199
- MOSFET, power (cont.)  
 protected switch, 205  
 push-pull follower, 214  
 source ballasting, 213  
 switch, 206t  
 switch examples, 206  
 switching applications, *195, 196, 194–196*  
 switching cautions, 196  
 thermal runaway in, 213  
 thermal stability, 187  
   vs BJT and IGBT, 201, 202t, 208t  
 $R_{ON}$  vs temperature, 213g  
 schematic symbols, 134  
 series connection of, 697  
 switch, *see also* switchmode  
   high-side, 825, 826t  
   protected, 824, 825t  
   protecting, **823–825**  
   transfer characteristics, 212g  
   vs JFET, 170  
 motorboating, 286  
 MOV, *see* varistor  
 MPEG-2, 1137  
 MPEG-4, 1061, 1137  
 MSP430, *see* microcontroller  
 multimeter, 2, 10 (box)  
 multiplexer, *see also* logic  
   FET analog switch, 173  
   in DAQ, 948  
 multiply  
   how to, 762  
 multiplying DAC, *see* DAC, multiplying  
 multislope  
   conversion, *see* ADC
- N**  
 Nakamura, Shuji, 62  
 NAND, *see* logic, *see* memory, nonvolatile  
 negative-impedance converter, 397  
 negative-true, *see* logic  
 netlist, *see* programmable logic device (PLD)  
 NEXT (near-end crosstalk), 581  
 nice  
   diagram, *1102*  
 night light, 206  
 noise, 15, *see also* op-amp, **473–481**, *see also* amplifier  
   1/f, 476  
   corner frequency, 566  
   filtered bandwidth, 564t  
   forever?, 566, 567s  
   integrated, 566g  
   JFET, 510, 517g  
   op-amp, 324g  
 analog  
   filtered PRBS, 975, **977–983**  
   averaging time, 574  
   band-limited, 477

noise (cont.)  
 bandpass filters  
 response, 565g  
 bandwidth, 477, 561ff  
 multisection  $RC$ , 561  
 of  $LC$ , 562  
 of an average, 562  
 bistable, 477  
 BJT, *see* BJT, noise  
 test circuit, 557  
 burst, 477  
 capacitively coupled, 581  
 common-mode choke, 584  
 crest factor, 568, 981  
 current, 483–484  
 measuring, 569ff  
 of bias-canceled op-amp, 327, 527  
 ultra-low measurement, 570, 572  
 current source, of, 487–489  
 density, 474, 479  
 voltage, 481–483  
 differential amplifier, 520  
 digital  
 generation with PRBS, 974ff  
 emitter follower, 487–489  
 $e_n$ -C, 538, 540g, 542s  
 exceedance probability curve, 574g  
 excess, 475ff  
 figure, 479–480, 480g  
 contour plot, 490, 492  
 example, 492  
 filtered  
 over decades, 567s  
 filtering out-of-band, 911  
 flicker, *see* noise, 1/f  
 Gaussian, 475g, 979, 980  
 generator, 559  
 pink, 559  
 PRBS, 559  
 immunity, 705  
 instrumentation amplifier, 362  
 integrals, 564t  
 integrated  
 calculating, 563  
 of op-amps, 338g, 530, 531g, 531t  
 interference as, 478  
 JFET, *see* JFET, noise  
 Johnson, *see* Johnson noise  
 low-frequency specification, 564  
 magnetically coupled, 581  
 maxima, 574  
 measurement, 556ff  
 hot-cold, 557  
 transistor test circuit, 557  
 Nyquist, *see* Johnson noise  
 of switching power supplies, 649  
 of voltage regulator, 618  
 op-amp, 522t, 523t, 524t, 528g, 529g, 532s,  
**521–533**

noise, op-amp (cont.)  
 autozero, 334ff, 334s  
 choosing, 525–533  
 vs  $C_{in}$ , 532g  
 peak vs rms, 532, 564, 568, 569, 935  
 popcorn, 477, 478  
 potpourri, 574  
 power supply, 578ff, 580g, 580t, 913  
 capacitance multiplier, 508, 557, 578,  
 579g  
 pseudorandom, *see* PRBS  
 quantization, 574, 927–931  
 gain, 927  
 model, 927  
 resistance, 494  
 rms vs magnitude, 574  
 shaping, *see also* delta-sigma, 927–931  
 shot, 327, 476t, 475–476  
 source  
 PRBS circuit, 980  
 sources, 558ff  
 switchmode power supply, 636g  
 telegraph, 477  
 temperature, 480  
 threshold-crossing rate, 574  
 true random, 982  
 circuit, 983  
 types of, 474, 477  
 units of, 474  
 voltage  
 from filtered PRBS, 978  
 voltage reference, 682, *see also* voltage  
 reference  
 white, 474–476  
 zener, 674, 676  
 nonlinearity  
 converter error, 881, 899  
 nonmonotonicity  
 converter error, 881  
 NOR, *see* logic, *see* memory, nonvolatile  
 Norton equivalent circuit, 66, 1108  
 NRZ, 1041  
 NRZI, 1041  
 number  
 2's complement, 707, 1046  
 arithmetic in, 708  
 BCD, 706  
 codes, 705ff, 880  
 floating-point, 708, 1047–1048  
 formats, 1047  
 Gray-code, 708, 709  
 binary conversion, 711  
 hex, 705  
 notation, 705  
 integer  
 justification, 1046–1047  
 signed, 707t  
 life, the universe, and everything, 42  
 of the beast, the, 666

number (cont.)  
 offset binary, 707  
 sign-magnitude, 707  
 storage in memory, 1048  
 nybble, 990  
 Nyquist  
 noise, 474  
**O**  
 offline, *see* ac powered  
 offset binary, *see* number, codes  
 Ohm's law, 4  
 generalized, 46  
 not obeyed by diode, 31  
 OLED, *see also* optoelectronics, 839  
 one second per hour (design example), 467  
 1T1C cell, 1018  
 one-shot, 77, 242, 459, 462t, 463t, 460–465  
 555, 460  
 application examples, 465  
 cautions, 462–464  
 problems with, 462–464  
 protecting pulsed power devices with, 466  
 retriggerability, 462  
 timing coefficient, 464g  
 1-wire bus, *see* computer, data bus  
 op-amp, 223ff, 224p  
 1/f noise, *see* noise  
 absolute-value circuit, 257  
 ac amplifier, 226  
 biasing, 261  
 active clamp, 257  
 active peak detector, 254  
 active rectifier, 238, 257  
 as voltage regulator, 235  
 autozero, 272, 295, 335t, 333–342ff, 370  
 external, 341  
 integrated noise, 338g  
 noise, 336g, 336, 337s, 569t  
 noise current, 570g  
 selecting, 338ff  
 thermal EMFs and, 340  
 bandwidth, 247, 249, 308, 328, 329g  
 basic circuits, 225–231  
 bias canceled, 125  
 better than JFET, 303g  
 current noise of, 327  
 capacitive load  
 stability, 265g  
 stabilizing, 263s, 264  
 cautions, 231  
 chopper-stabilized, *see* op-amp, autozero  
 clamping, 913  
 CMRR, 305, 328  
 CMRR and PSRR, 249  
 common-mode input range, 245  
 of autozero, 340  
 of difference amplifier, 354  
 compensation, 280ff

- op-amp (cont.)  
 compensation pin, 320t  
 current feedback, 270  
 peaking vs gain-setting  $R$ , 379g  
 current source, **228–230**, 242, 254, 344, 367, 623, 895  
 Howland, 229, 230  
 current-to-voltage converter, 233  
 decompensated, 283, 283, 328, 542, 550  
 departure from ideal, 243–249  
 detailed look, 242–253  
 die (LT1028), 527p  
 difference amplifier, 227, **347–356**  
 CMRR trim, 356  
 filter node, 355  
 offset trim, 355  
 differential input range, 232, 246  
 differentiator, 260  
 distortion, 329–331, 331, 332g  
 slew-rate, 248s  
 vs frequency, 311g  
 driving logic from, 808  
 feedback at dc, 232  
 follower, 227  
 with bootstrap, 233  
 frequency compensation, 247, *see* op-amp, compensation  
 gain, 249  
 vs frequency, 281g  
 gain error, 312g  
 gain nonlinearity, 312–314  
 measuring, 313  
 golden rules, 225  
 high-power, 272t  
 high-speed, 310t  
 high-voltage, 272t  
 hybrid amplifier with JFET, 152, 153g, 155g, 343–347, 534, 545  
 ideal, 243  
 input  
 protection, 362  
 input current, 163, 244, 252, 302, 325  
 BJT better than JFET, 303g  
 common-mode voltage, variation with, 304, 304g, 305  
 effect on integrator, 258  
 measuring, 325–326  
 of autozero, 339  
 temperature, variation with, 303g  
 input impedance, 245, 250, 301  
 input offset current, 244, 252  
 input protection, 259, 362  
 instrumentation, *see* instrumentation amplifier  
 integrator, 230, 231s, 257–260  
 inverting amplifier, 225  
 JFET,  
 fast, 155t
- op-amp (cont.)  
 JFET hybrid, *see* op-amp, hybrid amplifier with JFET  
 limitations, effect on circuits, 249–254  
 linearity, 251  
 low input-current, 303t  
 low noise, 522–524t  
 choosing, **525–533**  
 hybrid, 534  
 micropower, 273  
 Monticelli output circuit, 318  
 noise, 528g, 529g, 532s, **521–533**  
 0.01–10Hz,  $1/f$ ,  $V_{n(pp)}$ , 522–523t, 530g, 532, 564  
 $1/f$ , 528, 529g  
 autozero, **334ff**, 334s  
 integrated, 338g, 530, 531g, 531t  
 low-frequency, 564  
 vs  $C_{in}$ , 532g  
 noise voltage and current, 249, 323, 324g,  
*see also* noise, 326  
 of autozero, 339  
 of bias-canceled, 327  
 vs supply current, 525g  
 noninverting amplifier, 226  
 nonlinear circuits, 236  
 offset voltage, 244, 251, 304, 323  
 effect on integrator, 258  
 of autozero, 339  
 test circuit, 323  
 offset voltage drift, 244  
 optional inverter, 232  
 oscillator  
 precision VCO, 267  
 output current, 251, 272t  
 output impedance, 246, 250, 309  
 effect on MFB filter, 414g  
 effect on VCVS filter, 414g  
 vs frequency, 250g, 311, 312g  
 output range, 231  
 output swing  
 vs frequency, 251g, 307g  
 vs load resistance, 246g  
 parameters, 245t  
 unspecified, 296  
 peak detector  
 resetting, 255  
 phase error, 314, 315g, 315t  
 active compensation, 314, 315g, 315t  
 phase reversal in, 275  
 phase shift, 247  
 photodiode amplifier, 234  
 photometer, 265  
 pinch-off voltage tester, 240  
 pinout, 225  
 power and high-voltage, 272t  
 power booster, 234  
 precision, **292ff**, 320–321t  
 choosing, 319
- precision (cont.)  
 representative, 302t  
 vs speed, 329  
 programmable pulse-width generator, 241  
 PSRR, 306, 328, 533  
 push-pull booster, 234  
 rail-to-rail, **315ff**  
 distortion, 317g, 318  
 gain, 318g  
 input crossover, 316, 317g  
 output impedance, 316  
 output saturation, 317  
 output swing, 247g  
 relaxation oscillator, *see* oscillator, relaxation  
 representative, 271t  
 sample-and-hold, 256  
 schematic  
 LF411, 243  
 TLC271, 266  
 sensitive millivoltmeter (design example), 253, 293, 296t  
 settling time, 308ff, 320–321t, 328ff  
 of autozero, 340  
 sign of feedback, 232  
 single-supply, 261, 265–270, 322  
 slew rate, 248g, 251, 328  
 distortion, 248s  
 of autozero, 340  
 vs input differential, 308g  
 smorgasbord, 232  
 split feedback, 264  
 stuck-node tracer, 276g  
 summing amplifier, 234  
 supply rail splitter, 262  
 stability, 262, 263s  
 supply voltage, 248, 322  
 of autozero, 338  
 supply-rail bypass, 232  
 tables of, 271–272t, 296t, 302–303t, 310t, 320–321t, 335t, 522–524t  
 track-and-hold, 256  
 transconductance, *see* OTA (operational transconductance amplifier)  
 transimpedance amplifier, *see* transimpedance amplifier  
 triangle oscillator, 239  
 uncompensated, 283, *see also* op-amp, decompensated  
 voltage gain, 247  
 vs frequency, 247g  
 zero-crossing detector, 269  
 optocoupler, 805, 837p, 844p, 845, 846, 847, **843–848**, 848–852  
 analog, 586, 847  
 optoelectronics, **829–856**  
 coupler, *see* optocoupler  
 detector, 837p, 841p

- optoelectronics (cont.)  
 display  
   LED, multiplexed, 751–752  
 displays, 830p, 837p  
 emitters, 830p  
 family tree, 829, 831 (box)  
 fiber optics, 853p, *see also* fiber optics  
 image sensor, 841p  
 interrupter, 578, 844p  
 LCD, 837p, 839, 840, 1143  
   interface timing, 1028  
 miscellaneous, 837p  
 OLED, 1143  
 optocoupler, 1081  
 photoconductive, 437  
 photomultiplier, 837p  
 PIN diode, 837p  
 plasma, 1143  
 television display, 1143  
 optointerrupter, *see* optoelectronics  
 optoisolator, *see* optocoupler  
 oscillation  
   parasitic, 426, 427s, 442  
   slow logic transition, from, 801s  
 oscillator, **425–457**  
   *LC*, 439  
     electrically tunable, 440  
   **555**, 430t, **428–432**  
     50% duty-cycle, 429  
     full-range duty-cycle, 430  
   atomic standard, 451  
   ceramic resonator, 450  
   CMOS, 426–427  
   crystal, 443p, **443–450**  
     a caution, 449  
     circuits, 447  
     micropower, 448, 449g  
     modules, 449  
     series and parallel modes, 444, 445g  
   DDS, 451, 453, 455, 1066, 1082  
   emitter-coupled, 440  
   JFET, 155  
     low-noise, 441  
     spectrum, 442g  
   jitter, 457  
   local (LO), 562, 971  
   low-distortion, 437  
   not needed in  $\mu$ C example, 1056  
   op-amp  
     triangle wave, 239  
   parasitic oscillations in, 426, 427s  
   phase-shift, 438  
   PLL, 452  
   precision VCO, 267  
   programmable, 267  
   quadrature, 454s, 455s, **453–456**  
   quartz crystal, **443–450**  
   ratiometric, 431
- oscillator (cont.)  
 relaxation, 267, **425–435**  
   CMOS logic, 426, 427s  
   op-amp, 425  
   unijunction, 427  
 resistor-programmed, 432g  
 SAW, 450  
 sawtooth, 430, 917–918, 918  
 silicon, 432, 432t  
 sinewave, **435–443**  
   Wien bridge, 436–438  
 TCXO, OCXO, 450  
 triangle, 239, 267, 431  
 types, 452t  
 VCXO, **446**, 960, 964, 969, 971  
 voltage-controlled, 267, **434ff**, 435, 959  
   in PLL, **959–960**
- oscilloscope, 2, **1158ff**  
   analog, **1159**  
   digital, 1152, **1162ff**, **1163**  
     aliasing, 1164  
     dead time, 1164  
     hints, 1160  
     probe, 67, 1160, **1161**  
     skew, 1164
- OTA (operational transconductance amplifier), 100
- output  
   comparator swing, 810  
   current, comparator, 811, 811g  
   current, op-amp, 251, 272t  
   output current  
     critical, in SMPS, *see* switchmode  
   oversampling, *see* sampling  
   overtemperature  
     sensor, 13, 203, 602–604, 604 (box), 671  
     absent, 695
- P**
- PAL, *see* programmable logic device (PLD)  
 PALASM, *see* programmable logic device (PLD)
- parallel  
   bootloader, 1091  
   conversion, *see* ADC  
   impedances, 46  
   MOSFETs, 201  
   resistors, 5  
   resonant crystal mode, 444, 445g
- parameters  
   unspecified, 296
- parity generator, 728
- PATA, *see* computer, data bus  
 PC104, *see* computer, data bus  
 PCIe, *see* computer, data bus  
 PCMCIA, 1061
- peak detector  
   active, 254  
   passive, 254
- PECL (positive ECL logic), 874
- permeability, 29
- PGA, *see* amplifier, programmable gain
- phase  
   detector, 575, *see also* PLL  
   equalizer, 415  
   error  
     of op-amp, 314, 315g, 315t  
     sequence filter, 455, 456g  
   splitter, 105  
     with discrete BJTs, 88
- phase shift  
   and feedback stability, **280–285**, *see also*  
     Bode plot  
     filter comparison, 404g  
     oscillator, 438  
   RC lowpass, 50g
- phase shifter  
   constant-amplitude, 88, 89
- phase-locked loop, *see* PLL
- phasor, 51
- photodiode, 841–842, *see also* diode, photo
- photometer  
   single-supply, 265
- photomultiplier, 842, 842–843  
   amplifier, 843
- photoresistor, 206, 844, 845p
- phototransistor, 841–842
- pi network, 1109, 1124
- PIC, *see* microcontroller
- pickup, 442  
   common-mode, 584
- PicoGate, *see* mini-logic
- PID loop, *see* feedback
- piezo, 24
- piezo driver, 207, 332
- piezoelectric effect  
   capacitor, in, 682
- PLA, *see* programmable logic device (PLD)
- platinum RTD, *see* sensor, temperature
- PLD, *see* programmable logic device (PLD)  
   noise generator, 981  
   output characteristics, 798g
- PLL, 452, 972t, **955–974**, 1071  
   AM demodulation, 970  
   applications, **966–974**  
   block diagram, 956  
   BPSK demodulation, 971  
   capture and lock, **964–965**  
   clock generation, 972  
   “cut and try”, 963  
   design, **960–964**  
   FM demodulation, 969  
   frequency multiplier (design example),  
     **961–964**  
     circuit, 962  
   frequency synthesis, **966–969**  
     fractional-*n*, 966, **966–967**  
     rational-approximation, 969, 968–969

- PLL (cont.)  
 laser offset locking, 973, 973–974  
 loop  
   block diagram, 961  
 loop filter, 961  
 loop gain, 962, 963 (box)  
 loop stability, 961  
 phase detector, 957, 958s, **956–959**, 959s  
   backlash, 958–959  
   dead zone, 958–959  
 pulse synchronization, 971  
 rejection of noise and jitter, 974  
 sampling clock, 907  
 signal regeneration, 971  
 spur removal, 394  
 VCO, **959–960**  
   4046-type, 959, 959s  
   with JFET oscillator, 394
- PMT, *see* photomultiplier
- pod, 1091, 1093
- polar coordinates, 1097
- pole, 52
- popcorn  
   noise, 477, 478
- positive-true, *see* logic
- pot, *see* potentiometer
- pot core, 29
- potential difference, *see* voltage
- potentiometer, 7, 8p  
   digital, 63, 184, 412, 1082
- power  
   ac detect, 1081  
   ac monitor, 1082  
   an hour of (design example), 466  
   density  
     battery vs capacitor, 690g  
   design, **623ff**  
   dynamic, 754  
   factor, 47  
     correction in offline SMPS, 668  
     in offline SMPS, 661  
     in unregulated supply, 635  
   in reactive circuits, 47  
   in resistor, 6  
   micro, *see* micropower  
   minimizing, in CMOS design, 754  
   one minute of, 23  
   precision source, 285  
   semiconductor packages, 628–629p  
   spectrum  
     from PRBS, 979  
     of SMPS, 636  
   supervisor, 1079  
   switching  
     with MOSFETs, 192–200  
     with outboard driver, 237  
   tradeoffs  
     noise, 483g, 509g, 525g  
     speed, 792g
- power (cont.)  
   train, *see* switchmode (SMPS)  
   transfer, 11  
   transistor, 624  
   series connection, 698
- power supply, **594ff**, 595p, *see also* voltage regulator  
   clamp, 692  
   commercial modules, 684, 685p  
   current limit, 597  
   foldback, 693, 694g  
   of outboard pass transistor, 695  
   favorites, 1152  
   filtering, 32  
   high voltage, **695ff**, 696  
   linear  
     compared with switchmode, *see also* switchmode, 635–638  
     topology, compared with SMPS, 630  
   noise, *see* noise, power supply  
   offline, *see* switchmode, offline  
   overvoltage protection, 598, **691ff**  
   rectifier, 634  
   damping spikes, 634  
   regulated, 123  
   regulator, 34  
   resonant charging, 39  
   ripple voltage, 33  
   split, 34  
   storage capacitor, 633  
   ripple current in, 634
- switchmode, **636ff**, *see also* switchmode  
   ac-powered, *see* switchmode, offline  
   compared with linear, 630, 635–638  
   controller ICs, 658t  
   current-mode ICs, 654t  
   programmable, 1086  
   voltage-mode ICs, 653t  
   unregulated  
     bench vs SPICE, 634–635  
     from ac powerline, **628ff**  
     power factor in, 635  
   voltage multiplier, 34  
   zener as, 595
- powerline  
   capacitor, 631  
   discharge resistor, 631  
   entry, 629  
   filter, 631  
   monitor IC, 943, 943–944  
   rejection in integrating ADC, 915  
   switch, 631  
   transformer, 632  
   leakage inductance in, 633  
   rating, 633
- PRBS, **974–982**  
   analog noise from, 975, **977–983**  
   analog noise voltage, 978–981  
   chipping sequence, 975
- PRBS (cont.)  
   digital  
   filter, 979  
   eye diagram, 871s  
   feedback shift register, 976  
   properties, 976  
   taps, 976t  
   noise generator, 559, 771  
   spectrum, 560s  
   noise source  
   circuit, 980  
   power spectrum, 979  
   with cPLD, 772  
   with discrete logic, 559, 771–772  
   with ECL logic, 981, 983p  
   with microcontroller, **777ff**  
   assembly code, 779, 783  
   C-language code, 782, 784  
   with PLDs, 981  
   wrapup, 981
- pre-emphasis, 864, 870, 871s
- precision  
   vs dynamic range, 292
- prefixes, 4 (box)
- printed circuit  
   SMT vs through-hole, 268, 269
- printer port (Centronics), *see* computer, data bus
- priority encoder, *see* logic
- probe  
   for logic signals, 808, 810p
- probing  
   digital signals, 808, 809s, 810p
- programmable (with voltage)  
   current source, 228–230, 367, 896  
   oscillator, 267  
   power supply, 608, 610  
   pulse generator, 918  
   timer (one-shot), 242  
   voltage source, 195, 892
- programmable logic device (PLD), 728, 737, **745–746**, **764ff**, *see also* logic, 765p  
   ABEL, 764, 775, 776  
   address decoder, 1000  
   advice, **782ff**  
   behavioral coding, 777  
   cPLD, 764, 766  
   PRBS bit generator, **772–777**  
   CUPL, 764  
   FPGA, 764, **768–770**, 1031  
   configuration memory, 769, 770  
   logic element, 770  
   GAL, 764  
   HDL, 766, *see also* logic  
   HDL entry, **775ff**  
   mixed-signal, 769  
   netlist, 769  
   packages, 765p

- programmable logic device (cont.)  
 PAL, 764, **765–768**  
   circuit, 766–767  
   macrocell, 764, 768  
 PALASM, 764  
 PLA, 768  
 programmer, 769  
 schematic entry, **773ff**, 774  
 software, 769  
 structural coding, 777  
 Verilog, 766, 775, 778–781  
 VHDL, 766, 777  
   when to use, 782ff, 785  
 programmable oscillator, 267  
 pseudo-differential, 275, 584  
 pseudo-random bit sequence, *see* PRBS  
 PSOC, *see* microcontroller  
 PTAT, *see also* bandgap, 92, 104, 431, 606,  
   **680, 679–680**  
 pulse, 16  
   duty-cycle  
     limiting, 465  
   forming network, 1126  
   from step, 77, 459s, 458–468  
 generator, 17  
   high-voltage, 333, 917  
   n-pulse, 752–753, 753  
   programmable pulse-width, 241  
   single-pulse, 739  
   triggerable (design example), 739, 740  
   with discrete BJTs, 77–79  
     with flip-flops and counters, 739  
 limiting width of, 465  
 runt, *see* runt pulse, *see also* logic synchronization  
   with PLL, 971  
 width modulation, *see* PWM
- push-pull  
 BJT output stage, 106–108  
   biasing, 107  
   crossover distortion, 107  
   thermal stability of, 108  
 op-amp power booster, 234
- PWM, *see also* switchmode, 1061, 1071, 1077  
 as digital-to-analog converter, 888  
 in switchmode converters, 645ff  
 light dimming, 834  
 rate multiplier as, 890  
 thermal controller, **1069–1077**  
 to DAC converter, 889  
 torque-motor control, 890
- pyroelectric  
 detector, 841p
- Q**  
 Q (quality factor), 52  
 QAM, 456, 1138  
 quadrature  
   encoder, 852, 877
- quadrature 9cont.)  
   FM demodulation, 969  
   modulation, 577  
   oscillator, 454s, 455s, **453–456**
- quadrature-amplitude modulation,  
   *see* QAM
- quantization  
   noise, 927–931  
   gain, 927  
   model, 927
- R**
- R*–2*R*, *see* ladder
- Rabbit, *see* microcontroller
- radio  
   AM, 55, 56s  
   radiofrequency, *see* RF  
   rail splitter, 262, 277  
   rail-to-rail, *see* op-amp  
   ramp, 15  
     generator, 27  
   random walk, 981  
   RAS, 1019  
   raster, 1132  
   rate multiplier  
     as PWM, 890  
   *r*<sub>bb</sub>, *see* BJT (bipolar junction transistor)
- reactance, **40ff**  
   graph of, 49
- reactive circuits, **40ff**  
   power in, 47
- reading and references, **1154ff**
- rectification, **31ff**
- rectifier, **31ff**, 634  
   active, 238  
   bridge, 32  
   fast-recovery, 670  
   forward drop, 31  
   full-wave, 32  
     active, 257  
   full-wave center-tapped, 33  
   half-wave, 32  
   power supply  
     spikes, 634  
     signal, 35
- reference design, 671
- refresh, *see* memory, dynamic
- register, *see* logic
- regulation  
   secondary, in SMPS, 671
- regulator, *see* voltage regulator
- relay  
   driving from logic, 817–820  
   electromechanical, 59  
   optocoupled, 843–851  
   solid-state, *see* SSR
- reset  
   logic startup, 755  
   of core, in SMPS, 657
- resistance  
   dynamic, 12  
   internal, 11  
   leakage, 10  
   negative, 1113  
   of analog switch, 175  
   small-signal, 12  
   source, 11  
   thermal, *see* thermal resistance
- resistor, 3ff, 3p, 5 (box)  
 accuracy, 1106  
 chip  
   sizes, 4  
 color code, 1105t  
 digital, 1082  
 discharge, for powerline capacitor, 631  
 E96 series, 1104t  
 emitter-ballasting, 113  
 in parallel, 5  
 in series, 5  
 nonlinearity, 698, 1106  
 power in, 6  
 properties of, 300  
 standard values, 1104  
 tempco, 1106  
 types, **1104ff**, 1106t  
 variable, 63  
   JFET, 161  
 voltage coefficient, 698, 1106  
 wirewound, 3p, 8p, 63p, 591, 1070, 1105
- resonant  
   charging, 39  
   circuit, **52ff**
- RF  
   driving cable, 862
- RFI (radiofrequency interference), 582
- ribbon microphone, 505
- ringback, 759
- ringing  
   in SMPS, 668
- ripple, 635, 636s, 639g  
   current, 644  
   effect of capacitor ESR, 639  
   effect of switchmode frequency, 641  
   filtering, 619  
   in unregulated supply, 599, **632ff**  
   rejection, 605
- rise time, 16
- RLL, 1041
- Rogowski coil, 944
- RS-232, **871–874**, 1039p, **1038–1039**, 1081  
   and cable length, 872s  
   driver ICs, 872  
   levels, 873  
   power, 874g  
   pseudo-random byte generator, 771  
   rates, 873g  
   signals, 1039t  
   timing ambiguity in, 1038

- RS-422, 865s, 866s, **865–867**  
 levels, 873  
 power, 874g
- RS-485, 865s, 866s, **865–867, 1038–1039**  
 isolated, 868  
 levels, 873  
 power, 874g
- RTC, 1084
- RTD, *see* sensor, temperature
- rubidium, *see* oscillator, atomic
- runt pulse, 737s, 738, 743s, 748, 753, *see also* logic race, 757
- S**
- S/PDIF, 1061
- safe operating area (SOA), 192g, 216, 627
- Sallen and Key filter, 399, 409–410
- sample  
 PLL clock, 907  
 sample-and-hold, 183, *see also* op-amp  
 sampling, 419, 901, **901ff**  
 bandpass, 907  
 depth  
   and dynamic range, 900  
 direct IF to digital, 907  
 Nyquist criterion, 900, 901  
 over, *see also* delta-sigma, 901, 901, 923  
   ratio (OSR), 924  
 rate  
   and aliasing, 900, 901  
   and bandwidth, 900  
 rate and aliasing, 931  
 under, 901, **901ff**  
   band conversion, 907
- SAS, *see* computer, data bus
- SATA, *see* computer, data bus
- saturation  
 in BJT, 73  
 of magnetic core, 657, 659
- sawtooth, *see* oscillator
- schematic  
 diagram  
   how to draw, **1101ff**  
 entry  
   in SPICE, 1146
- Schmitt trigger, 237, 269, 459, 717, 754, 801s, 803s, 807–809s
- speedup capacitor, 808s  
 with discrete BJTs, 79
- SCR  
 latchup, 759  
 overvoltage crowbar, 598, **691ff**
- SCSI, *see* computer, data bus
- SD card, 1061, 1082
- second breakdown, 192, 216, 627
- Seebeck effect, *see* thermal EMF
- sensor, 1079–1086  
 ac power, 943–944  
 atom, 553ff
- sensor (cont.)  
 capacitive, 505  
 charge, 933  
 current, 278, 943–944  
 grasp, 372  
 light, 497, 539ff, 547, 549, 578, 837p, 841p, 840–843, 844p  
 lightning, 497  
 sound, 486, 505ff  
 strain, 297ff, 588, 937
- suntan, **278–280, 922, 1054–1059**
- temperature, **1070**  
 RTC, 1069ff, 1071  
 RTD, 347, 1070  
 semiconductor, 621, 680  
 thermistor, 123, 372, 608, 1070  
 thermocouple, *see* AoE 2nd ed., pp 988–992, 339, 341t, 367, 675, 937, 1070, 1084
- SERDES, 855, 869
- serial, *see* computer, data bus, *see also* logic  
 series  
 impedances, 46  
 resistors, 5  
 resonant crystal mode, 444, 445g  
 termination, **858ff**
- series-feedback pair, 122  
 with JFET, 150, 151g
- servo, 1069
- settling time, 309g  
 of op-amp, 308ff, 320t, 328ff
- setup time, *see* logic
- SH-4, *see* microcontroller
- shielding, 579ff
- shift register, *see* logic
- shoot-through, *see* current
- shootout  
 ADCs, SAR vs  $\Delta\Sigma$ , 939  
 at the ADC corral, 939t  
 JFET vs BJT, 517  
 low-power ADC, 941–942
- shunt  
 current, 277
- siemens, 6, 90, 147
- signal  
 averaging, 576  
 generator, 17  
 guard, 360, 587  
 mixed  
   PLD, 769
- regeneration  
 with PLL, 971
- sampling, 419
- slew rate of, 362
- to-noise ratio, *see* SNR
- signals, **14ff**
- silicon-controlled rectifier (SCR), *see* SCR
- SIM Card, 1061
- simulation, *see* SPICE
- sinewave, 14  
 from square wave, 394s, 435
- inverter, 673s  
 modified, 673s  
 oscillator, *see* oscillator, sinewave
- Single Gate, *see* mini-logic
- single pulse, 739
- single slope  
 conversion, *see* ADC
- 6T cell, 1016
- sizzle test, 625
- skin depth, 1122g
- slew rate, *see also* op-amp  
 and jitter, 457  
 causes interference, 581, 588, 705  
 filtered, 1073  
 in oscillators, 426  
 of composite amplifier, 920  
 of current source, 367  
 of differential amplifier, 385  
 of op-amp, 248, 308, 328, 330, 340, 515, 862  
 of signal, 362  
 reduced, 607, 650, 759, 820, 865, 871  
 too slow, 760, 808
- small-signal, 12, 45, **80, 90, 147, 583**  
 impedance, 81, 245  
 symbol for, 45, 80, 90, 1097
- transistor, 74t
- SMU, 554, 1115  
 favorites, 1152
- SNR, **478–481**
- snubber, 39, 630, 634, 656, 660, 669
- soft start, 606, 607, 650
- solar cell, 2
- soldering iron  
 favorite, 1152
- solenoid, 29
- source  
 floating, for bias, 215  
 sinewave, 41
- source follower  
**JFET, 156ff**  
 gain, 157  
 input impedance, 157  
 low distortion, 160, 161g  
 output impedance, 157  
 tutorial, 159  
   with active load, 158
- push-pull power MOSFET, 214
- source-measure unit, *see* SMU
- source-measure unit (SMU), 897
- Space (RS-232), 871
- SparkFun, 1082
- spectrum  
 analyzer, 567
- SPI, 887, *see also* computer, data bus, 1061, 1081, 1084, 1086
- ADC SAR example, 909

- SPI (cont.)  
 bootloader, 1091  
 in ADC, 936, 937, 943, 945  
 in DAQ, 946, 947, 951, 954  
 isolation in DAQ, 950–951  
 microcontroller peripherals, **1082–1084**
- SPICE  
 amplifier analysis, 314, *315*  
 comparison with bench, 150, 459s, 635s  
 crystal analysis, 443–446  
 filter analysis, 393, 399, 414, 456, 560  
 primer, **1146ff**  
 simulation  
     comparison with bench, 1148s  
 unregulated power supply  
     bench comparison, 634, 635g
- spike, *see also* charge injection, 17  
 differentiated edge, 25, 26s  
 from discharge, 200  
 from flyback, 39, 193  
 from leakage inductance, 669  
 from snapoff, 199, 634  
 ground-current, 856, 857  
 in frequency domain, 977  
 in power supply rectifier, 634  
 in SMPS driver, 669  
 logic driving capacitive load, 857
- spur  
 removal, 394
- square wave, 16  
 converting to sinewave, 394s  
 quadrature, 455
- SSR, 24, 59, 848, **848–850**, 850–851, 1062, 1081
- stabilized platform  
 microcontroller, 1077p, **1077–1078**
- standing-wave ratio, *see* VSWR
- state machine, *see* logic
- state-variable filter, 410–413
- step, 17  
 response  
     filter comparison, 406, 409g
- stepping motor, 1084
- STM (scanning tunneling microscope), **553ff**  
 image of Si crystal, 553p
- STP, 583
- stripline, 456, 858, 864, 871s, 1117, *see also* microstrip
- stuck node  
 tracer, 276g
- successive approximation, *see* ADC
- suntan monitor  
 delta-sigma, 922–923  
 microcontroller, **1054–1059**
- op-amp, current divider, 279
- op-amp, integrating, 278
- op-amp, two-step, 279
- with ADC, 922
- with microcontroller, **1054–1059**
- superbeta, *see* BJT
- surface acoustic wave, *see* oscillator, SAW
- surface-mount technology (SMT), 65, 66p, 627p
- susceptance, 90
- switch  
 3-way, 59  
 analog, *see* analog switch, 948  
     multiplexer, 948  
 as input device, 802–804  
 bounce, 802, 803s  
 break-before-make, 179  
 debouncer  
     latch as, 730  
     SR flip-flop as, 730  
 debouncing, 729, 802  
 DIP, 57p  
 dry-switching, 58  
 floating, 203  
 high-side, 821, 825, 826t  
 matrix, 1069  
 mechanical, 56, 57p  
 MOSFET, *see* switchmode; MOSFET  
 opto-interrupter, 851  
 PCB-mounting, 57p  
 powerline, 631  
 protected, 175, 824, 825t  
 protecting, **823–825**  
 pushbutton, 58  
 reed, 573  
 rotary, 58  
 simple logic with, 58  
 T-switch, 179  
 toggle, 57p  
 toggling, 196
- switch, analog, *see* analog switch
- switched-capacitor  
 filter, **415**  
 integrator, 416  
 resonator, 454
- switched-capacitor filter, **418**
- switching  
 amplifier, 109s, 673  
 loss, 662
- switchmode  
 zero-current switching (ZCS), 823  
 zero-voltage switching (ZVS), 823
- switchmode (SMPS), *see also* voltage regulator, **636ff**  
 boost, **647**  
 bridge converter, **659**  
 buck, **642ff**  
 buck, examples, 644ff  
 buck-boost, **651**  
 compared with linear, 630, 635–638  
 continuous-conduction mode, 644, 670  
 controller ICs, 658t  
 core, gapped, 656, 670  
 critical conduction mode, 643
- switchmode (cont.)  
 critical output current, 643  
 current pulsation, 649  
 current-mode ICs, 654t  
 design equations, 644, 647–648, 656, 657, 659  
 efficiency, 636  
 flyback, **655**  
     power train, 671  
     transformer in, 670  
 forward converter, **656**  
 H-bridge, 655, 659  
 hysteretic, 644  
 inductorless, **638ff**, 639g, 639t  
     regulated, 641  
     ripple, 639, 641  
 internally compensated, 645–646  
 inverting, **648**  
 isolated, **653ff**  
     flyback, 655  
     topologies, 655  
 loss, 662  
 low-noise, 649  
 noise, 649  
 noise spectrum, 636g  
 non-isolated, **641ff**  
     boost, **647**  
     buck, **642ff**  
     buck, examples, 644ff  
     buck-boost, 651  
     design equations, 644, 647–648  
     hysteretic, 644  
     internally compensated, 645–646  
     inverting, **648**  
     waveforms, 643, 647, 648, 652  
     wrapup, **649ff**
- offline, **660ff**  
 dual-voltage, 660  
 inrush current, 661, 668  
 power-factor correction (PFC), 661, 668  
 real-world example, **665ff**  
 when to use, 672  
 wrapup, 672
- point-of-load, 637, 686
- PWM modes, 651
- soft start, 650
- topologies, 638  
 compared with linear, 630
- voltage mode, 652
- voltage-mode ICs, 653t
- waveforms, 643, 647, 648, 652, 666, 667s
- wrapup, non-isolated, **649ff**
- zero-current switching (ZCS), 650, 662, 850, 1062
- zero-voltage switching (ZVS), 650, 659, 671, 850, 1062
- synchronous, *see also* logic, *see also* memory  
 synchronous detection, *see* bandwidth, narrowing

- synthesizer  
clock, 1084  
*Sziklai, see BJT*
- T**  
T network, 1109, 1124  
tables, *see Table of Contents*  
take-back-half, 1075  
convergence, 1076g  
television  
analog  
broadcast, 1134  
cable, 1134  
recording, 1135  
spectrum, 1136s  
cable  
switched broadcast, 1142  
digital, **1136ff**  
broadcast, 1138  
cable, 1138  
recording, 1142  
satellite, 1139  
spectrum, 1136s  
satellite  
uplink, 1140p  
tutorial, **1131ff**  
temperature  
coefficient, *see also referent, e.g., voltage reference, etc.*  
stabilized  
oscillator, 450  
voltage reference, 681  
temperature coefficient  
ways to specify, 683g  
temperature controller, *see thermal controller*  
temperature sensor, *see sensor*  
terminal nomenclature, 131  
termination  
of transmission line, **858ff**,  
1117  
*Thévenin equivalent circuit, 9ff*  
examples, 1107  
*Thévenin's theorem, 1107ff*  
generalized, 55  
thermal controller, 123  
microcontroller, **1069–1077**  
thermal EMF, 339–340, *see also op-amp, autozero, see EMF, 675*  
thermal hysteresis, 683  
thermal resistance, 216, **624ff**  
of heatsinks, 626g  
of PCB foil, 627g  
transient, 628, 1074  
thermal runaway, *see also second breakdown, 108ff, 112, 187, 212ff, 214*  
absent in MOSFET, 138, 215ff  
*thermistor, see sensor*  
*thermocouple, see sensors, temperature thyristor, 208*
- TIA, *see transimpedance amplifier*  
time  
delay  
filter comparison, 404g  
delay circuit, 23  
hold, *see logic*  
propagation, *see logic*  
removal, *see logic*  
setup, *see logic*  
time constant, **22**  
timer  
an hour of power, 466  
camera control (design example), 467  
ICs for, 467  
microcontroller as, 469  
one second per hour, 467  
programmable, 242  
timers, **457–469**  
555, 460  
*TinyLogic, see mini-logic*  
tolerance, 6, 300  
effect on filter response, 402  
tools  
favorites, 1152  
toroid  
common-mode choke, 585  
totem-pole, 695  
touchscreen, 944–945, 1082, 1084  
track-and-hold, *see also op-amp, 911*  
tracking ADC, 909  
transconductance, 89, *see also BJT, FET, JFET*  
transconductor, *see OTA (operational transconductance amplifier)*  
transducer, *see sensor*  
transfer function, 7  
transformer, **30**  
for wideband ADC input, 382  
in feedback loop, 285, 286g, 536  
isolated SMPS, **653ff**, *see also switchmode powerline, 632*  
leakage inductance in, 633  
rating, 633  
signal, **535–537**  
signal isolator, 586  
variable, 64p  
transient  
capture, in PLL, 964  
thermal resistance, 628  
transient response  
of voltage regulator, 616  
transimpedance amplifier, 233, **537ff**, 548  
bandwidth, 537–539  
bootstrapped cascode, 550  
bootstrapping, 547  
capacitive feedback, 552, 570  
cascode input, **548ff**  
commercial, 570  
composite, **543ff**  
current calibrator, 555
- transimpedance amplifier (cont.)  
noise, 497  
 $e_n \cdot C$ , 538, 540, 541g, 542s, 544s  
input, 538  
patch clamp, 552  
regulated cascode (RGC), 551  
stability, 537  
STM, 553  
world's best, 545
- transistor  
bias-resistor, 712  
bipolar, *see BJT*  
BJT switch, 73–79  
diamond, 100  
digital, 712  
field-effect, *see FET*  
man, 75  
packages, 73p  
pass, 594, 695  
protection, 695  
perfect, 99  
photo, 841p, 841–842  
pre-biased, 712  
series connection of, 697  
thermal resistance of, 624, *see also thermal resistance*  
unijunction, 427  
transmission gate, *see logic*  
transmission line, **1116ff**  
and sinewaves, **1120ff**  
characteristic impedance, 1116  
delay line, 1128p  
differential, 864–873  
LVDS, 868–873  
digital signals on, **858–873**  
power dissipation, 874g  
impedance matching, 1122  
broadband lossless, 1124  
minimum loss, 1123, 1124t  
narrowband lossless, 1125  
loss, 1121, 1122g  
matched  
capacitance disappears, 1116  
reflection, 1117, 1118s  
resistive attenuator, 1123, 1124t  
ringing, 1119s  
termination, 860–864, **1117ff**  
back, 1118  
series, 1118  
twisted pair, 1117  
unterminated, 859s, 860s, 858–860,  
1119  
transresistance amplifier, *see transimpedance amplifier*  
triangle wave, 15  
triboelectric effect, 573  
trigonometry, 1097  
truth table, *see logic*  
TTL, *see logic*

- TVS, 39  
 twin-T filter, 414  
 twisted pair  
     wideband analog over, 352  
 2's complement, *see* number, codes
- U**  
 UART, *see also* computer, 1038ff, 1061ff, 1071, 1089  
     bootloader, 1090  
 ugly  
     diagram, 1102  
 undersampling, *see also* ADC, *see* sampling  
 uninterruptible power supply (UPS), 673  
 unused input, *see also* logic  
     in ADC, 935  
 UPS, *see* uninterruptible power supply  
 USB, 1042, *see also* computer, data bus, 1061, 1081  
     bootloader, 1091  
     to RS-232 adapter, 871
- V**  
 V to f  
     conversion, *see* ADC  
 varactor, 64, 960, 969, 971  
     tuned oscillator, 440  
 Variac, 64p  
 varistor, 39  
 VCO, *see* oscillator, voltage-controlled  
 Verilog, *see also* logic, *see* programmable logic device (PLD)  
 VFD, 839  
 VHDL, *see* programmable logic device (PLD)  
 Victoreen, 3p, 698  
 video  
     analog, 1132  
     composite, 1133s  
     connectors, 1143ff, 1144p  
     digitizer, 1084  
     modulation, 1133  
     over Internet, 1140  
 volt, 1  
 voltage, 1  
     as complex quantity, 44ff  
     burden, 10  
     common-mode, *see* amplifier  
     comparator output swing, 810  
     compliance, 8  
     converter  
         flying capacitor, 183  
     divider, 7  
         adjustable, 7  
         equation, 7  
         generalized, 48  
     doubler, *see also* charge pump, 34, 639, 660, 872, 1063  
     drop  
         of diode, 31
- voltage (cont.)  
     dropout, *see* voltage regulator  
     high, *see* high voltage  
     Kirchhoff's voltage law, 2  
     mode  
         PWM, in SMPS, 652, *see also* switchmode  
     multiplier, 34  
     noise, *see* noise  
     offset, *see also* amplifier; op-amp  
         comparator, 812  
         correction with DAC, 948  
         op-amp, 244  
     offset in op-amp, *see* op-amp  
     power-supply ripple, 33  
     reference, *see* voltage reference  
     regulator, *see* voltage regulator  
     signal, 14, 17, 21–27, 35–37, 43–45, 49g  
     source, 8  
         8-channel, 893, 895  
         low-noise, 891, 892  
     source, general laboratory, 891, 892  
     spike, *see* spike  
     thermoelectric, 341t  
     threshold  
         of digital logic, *see also* logic, 796–797  
     triboelectric, 573  
 voltage reference, 674ff  
     2-terminal, 677t  
     3-terminal, 678t  
     accuracy, 683  
     bandgap, 104, 599, 602, 604 (box), 606, 611, 677, 678, 679  
         as temperature sensor, 680  
         noise in, 679, 681, 682  
     current source with, 895  
     drift, 683  
     humidity, 684  
     in DAQ, 949–950  
     JFET pinchoff, 680  
     MOSFET floating gate, 681  
     noise, 682g, 682  
     precision, 681  
     PTAT, *see* PTAT  
     self-heating, 684  
     series, 678t  
     shunt, 677t  
     tempco, 683  
     V<sub>BE</sub>, *see* voltage reference, bandgap  
     x-ray exposure, 684  
     zener, 674  
         biasing, 675  
         bootstrap biasing, 676  
         compared with bandgap, 679  
         drift, 675  
         dynamic resistance, 674  
         IC, 676, 677t  
         LTZ1000, the spectacular, 675  
         noise, 674, 676
- voltage reference, zener (cont.)  
     tempco, 674  
 voltage regulator  
     3-terminal  
         as current source, 620  
         floating HV, 698  
     317-style, 604 (box), 605t  
         ADJ-pin current, 605  
         as current source, 620  
         circuit, 606  
         dual voltage, 612  
         dual-tracking lab supply, 608  
         high-voltage supply, 609  
         hints, 604, 607  
         proportional fan control, 608  
         schematic, 603, 606  
     7800-style, 602t  
     723, 598ff  
         in defense of, 600  
 current limit, 693  
     foldback, 694  
 current-reference, 611  
 emitter follower as, 82–83  
 extending V<sub>IN</sub>, 211, 693  
 fault protection, 619  
 filtering, 619  
 fully integrated, 600ff  
 high voltage, 609, 610, 695ff, 696  
 linear  
     3-terminal adjustable, 602ff, *see also* 317-style  
     3-terminal fixed, 601, 602t  
     bypassing, 616  
     choosing, 613  
     crowbar, 598  
     current limit, 597  
     dropout voltage, 606g  
     error amplifier, 597  
     feedback stability, 597  
     ground-pin current, 616  
     high-voltage, 693, 698  
     idiosyncrasies, 613  
     LDO stability, 616  
     noise, 618  
     pinouts, 613  
     reverse polarity, 616  
     taxonomy, 601  
     transient response, 616, 618s  
     tutorial, 595ff, 596  
 load dump, 618  
 low dropout, 599  
 low-dropout (LDO), 610ff, 612g, 614–615t  
 LT3080-style, 611  
     as current source, 620  
 op-amp, 235  
 ramp-up control, 606  
 schematic, LM317, 603, 606  
 switching, 609

- voltmeter, 2  
 sensitive (design example), 253, 293  
 VOM, 10  
 VSWR, 1121
- W**  
 wall plug  
   blades and wires, 628ff  
 wall wart, 686  
 watt, 2  
 waveform  
   sampling, 419  
 Wheeler's formula, 28  
 wideband  
   analog link, 380  
 Widlar, Bob, 97, 101ff, 307, 598ff, 604, 679  
 Wien bridge, *see* oscillator  
 WiFi, 1062  
 Williams, Jim, 316, 330, 438, 1074, 1154  
 Willison, John, 968
- Wilson mirror, 102, 146  
 window  
   averaging, 562, 941g  
   discriminator, *see* AoE 2nd ed., p 669  
   quartz, 769, 1023  
   sampling, 420–421  
   switch, 573  
 wire  
   long  
     digital signals through, **856–874, 1116–1122**  
 wired-AND, -OR, *see* logic  
 wirewound resistor, *see* resistor, wirewound  
 world's best  
   DMM, 344  
   low  $e_n$  amplifier, 505  
   low distortion oscillator, 438  
   low-memory capacitor (Teflon), 301g  
   low-noise power supply, 580  
   multichannel fast  $\Delta\Sigma$  ADC, 956
- world's best (cont.)  
 op-amp noise plot, 531g  
 precision current source, 898  
 precision resistor, 227, 300, 1106  
 PWM DAC, 889  
 TIA, 547  
 voltage reference, 675, 677t, 681–682, 684, 892
- X**  
 XOR, *see* logic, gates
- Y**  
 yesterday, 273 (box)
- Z**  
 zener, *see* diode, zener  
   active, 692  
 zero-crossing detector, 269  
 ZigBee, 1061  
 ZVS, ZVC, *see* switchmode