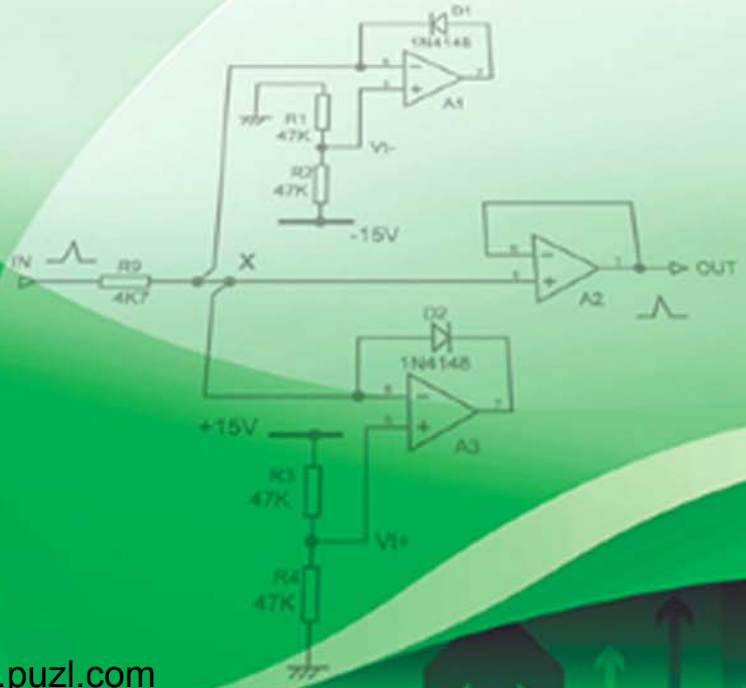
DOUGLAS SELF

SMALL SIGNAL AUDIO DESIGN



www.raguvaran.puzl.com

Focal Press

Small Signal Audio Design

Douglas Self

ELSEVIER

AMSTERDAM • BOSTON • HEIDELBERG • LONDON • NEW YORK • OXFORD
PARIS • SAN DIEGO • SAN FRANCISCO • SINGAPORE • SYDNEY • TOKYO

Focal Press is an imprint of Elsevier

www.raguvaran.puzl.com

Focal Press is an imprint of Elsevier
The Boulevard, Langford Lane, Kidlington, Oxford, OX5 1GB, UK
30 Corporate Drive, Suite 400, Burlington, MA 01803, USA

First published 2010

Copyright © 2010 Douglas Self. Published by Elsevier Ltd. All rights reserved.

The right of Douglas Self to be identified as the author of this work has been asserted in accordance with the Copyright, Designs and Patents Act 1988

No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage and retrieval system, without permission in writing from the publisher. Details on how to seek permission, further information about the Publisher's permissions policies and our arrangement with organizations such as the Copyright Clearance Center and the Copyright Licensing Agency, can be found at our website: www.elsevier.com/permissions

This book and the individual contributions contained in it are protected under copyright by the Publisher (other than as may be noted herein).

Notices
Knowledge and best practice in this field are constantly changing. As new research and experience broaden our understanding, changes in research methods, professional practices, or medical treatment may become necessary.

Practitioners and researchers must always rely on their own experience and knowledge in evaluating and using any information, methods, compounds, or experiments described herein. In using such information or methods they should be mindful of their own safety and the safety of others, including parties for whom they have a professional responsibility.

To the fullest extent of the law, neither the Publisher nor the authors, contributors, or editors, assume any liability for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions, or ideas contained in the material herein.

British Library Cataloguing in Publication Data
Self, Douglas.
Small signal audio design.
1. Acoustical engineering. 2. Signal processing.
I. Title
620.2-dc22

Library of Congress Number: 2009939461

ISBN: 978-0-240-52177-0

For information on all Focal Press publications visit our website at www.focalpress.com

Printed and bound in the United Kingdom

07 08 09 10 11 11 10 9 8 7 6 5 4 3 2 1

Working together to grow libraries in developing countries
www.elsevier.com | www.bookaid.org | www.sabre.org
ELSEVIER BOOK AID International Sabre Foundation

www.raguvaran.puzl.com

Contents

Preface ..... ***xiii***

About the Author ..... ***xvii***

Acronyms ..... ***xviii***

Chapter 1 The Basics ***1***

Signals 1
Amplifiers 2
Voltage Amplifiers 2
Transconductance Amplifiers 2
Current Amplifiers 3
Transimpedance Amplifiers 3
Negative Feedback 4
Nominal Signal Levels 5
Gain Structures 6
Amplification Then Attenuation 6
Attenuation Then Amplification 7
Raising the Input Signal to the Nominal Level 7
Active Gain Controls 8
Noise 8
Johnson Noise 9
Shot Noise 11
1/*f* Noise (Flicker Noise) 12
Popcorn Noise 13
Summing Noise Sources 13
Noise in Amplifiers 14
Noise in Bipolar Transistors 16
Bipolar Transistor Voltage Noise 16
Bipolar Transistor Current Noise 17
Noise in JFETs 20
Noise in Op-Amps 21
Low-Noise Op-Amp Circuitry 22
Noise Measurements 23
How to Attenuate Quietly 23

www.raguvaran.puzl.com

How to Amplify Quietly 25
How to Invert Quietly 26
How to Balance Quietly 28
Ultra-Low-Noise Design with Multipath Amplifiers 28
Ultra-Low-Noise Voltage Buffers 29
Ultra-Low-Noise Amplifiers 30
References 32

Chapter 2 Components 33
Conductors 33
Copper and Other Conductive Elements 34
The Metallurgy of Copper 35
Gold and Its Uses 35
Cable and Wiring Resistance 36
PCB Track Resistance 37
PCB Track-to-Track Crosstalk 39
Impedances and Crosstalk: A Case History 40
Resistors 42
Through-Hole Resistors 43
Surface-Mount Resistors 43
Resistor Imperfections 45
Resistor Noise 46
Resistor Non-Linearity 48
Capacitors 51
Capacitor Non-Linearity Examined 52
Non-Electrolytic Capacitor Non-Linearity 53
Electrolytic Capacitor Non-Linearity 58
Inductors 61
References 61

Chapter 3 Discrete Transistor Circuitry 63
Why Use Discrete Transistor Circuitry? 63
Bipolars and FETs 64
Bipolar Junction Transistors 64
The Transistor Equation 65
Beta 66
Unity-Gain Buffer Stages 66
The Simple Emitter-Follower 67
The Constant-Current Emitter-Follower 70
The Push-Pull Emitter-Follower 71
Emitter-Follower Stability 72
CFP Emitter-Followers 73
Improved Unity-Gain Buffers 75
Gain Stages 81

www.raguvaran.puzl.com

One-Transistor Shunt-Feedback Gain Stages 81
One-Transistor Series-Feedback Gain Stages 82
Two-Transistor Shunt-Feedback Gain Stages 82
Two-Transistor Series-Feedback Gain Stages 87
Discrete Op-Amp Design 88
High-Input-Impedance Bipolar Stages 92
References 94

Chapter 4 Op-Amps and Their Properties 95
A Very Brief History of Op-Amps 95
Op-Amp Properties: Noise 96
Op-Amp Properties: Slew Rate 97
Op-Amp Properties: Common-Mode Range 98
Op-Amp Properties: Input Offset Voltage 98
Op-Amp Properties: Bias Current 99
Op-Amp Properties: Cost 100
Op-Amp Properties: Distortion 101
Op-Amp Internal Distortion 101
Slew-Rate-Limiting Distortion 101
Distortion Due to Loading 103
Thermal Distortion 104
Common-Mode Distortion 104
Bipolar input op-amps 104
FET op-amps 108
Rail bootstrapping to reduce CM distortion 110
Simpler rail bootstrapping 113
Bootstrapping series-feedback JFET op-amp stages 114
Selecting the Right Op-Amp 115
Op-Amps Surveyed: BJT Input Types 117
The LM741 Op-Amp 118
The NE5532/5534 Op-Amp 118
Deconstructing the 5532 120
The LM4562 Op-Amp 121
The AD797 Op-Amp 124
The OP27 Op-Amp 124
The OP270 Op-Amp 125
The OP275 Op-Amp 126
Op-Amps Surveyed: JFET Input Types 128
The TL072 Op-Amp 128
The TL052 Op-Amp 129
The OPA2134 Op-Amp 131
The OPA604 Op-Amp 132
The OPA627 Op-Amp 135
References 136

www.raguvaran.puzl.com

Chapter 5 Filters *137*
Passive Filters 137
Active Filters 138
Low-Pass Filters 138
High-Pass Filters 138
Combined Low-Pass and High-Pass Filters 139
Bandpass Filters 139
Notch Filters 139
All-Pass Filters 139
Filter Characteristics 140
Sallen-and-Key Filters 140
Distortion in Sallen-and-Key Filters 145
Multiple-Feedback Bandpass Filters 147
Making Notch Filters 148
Differential Filters 150
References 151

Chapter 6 Preamplifier Architectures *153*
Passive Preamplifiers 153
Active Preamplifiers 155
Amplification and the gain-distribution problem 156
Active Gain Controls 158
Recording Facilities 158
Tone Controls 158
References 159

Chapter 7 Moving-Magnet Disc Inputs *161*
The Vinyl Medium 161
Spurious Signals 162
Maximum Signal Levels on Vinyl 163
Equalization and Its Discontents 166
The 'Neumann Pole' 168
Implementing RIAA Equalization 168
Passive and Semi-Passive RIAA Equalization 171
Calculating the RIAA Equalization Components 175
RIAA Equalization Component Tolerances 175
Simulating Inverse RIAA Equalization 178
Physical Inverse RIAA Equalization 179
Overload Margins 179
Cartridge Impedances 181
Cartridge Loading 181
Cartridge–Preamplifier Interaction 183
Cartridge DC and AC Coupling 183

www.raguvaran.puzl.com

Discrete MM Disc Input Stages 183
Op-Amp MM Disc Input Stages 187
Noise in RIAA Preamplifiers 191
Noise Results with A-Weighting 196
RIAA Noise Measurements 196
Electronic Cartridge Loading for Lower Noise 197
Subsonic Filters 201
Combining Subsonic and Ultrasonic Filters 204
References 204

Chapter 8 Moving-Coil Head Amplifiers 207
Moving-Coil Cartridge Characteristics 207
Amplification Strategies 208
Moving-Coil Transformers 209
Moving-Coil Input Amplifiers 210
A New Moving-Coil Amplifier Configuration 213
The Complete Circuit 216
Performance 216
References 218

Chapter 9 Volume and Balance Controls 219
Volume Controls 219
Volume Control Laws 220
Loaded Linear Pots 221
Dual-Action Volume Controls 223
Tapped Volume Controls 226
Slide Faders 229
Active Volume Controls 231
The Baxandall Active Volume Control 235
A Practical Baxandall Active Gain Stage 237
Potentiometers and DC 238
Motorized Potentiometers 239
Stepped Volume Controls 240
Switched Attenuator Volume Controls 241
Relay-Switched Volume Controls 249
Transformer-Tap Volume Controls 250
Integrated Circuit Volume Controls 250
Balance Controls: Passive 250
Active Balance Controls 252
Mono/Stereo Switches 254
Width Controls 254
References 255

www.raguvaran.puzl.com

Chapter 10 Tone Controls and Equalizers *257*
Passive Tone Controls 258
Baxandall Tone Controls 259
The Baxandall Two-Capacitor LF Control 260
The Baxandall Two-Capacitor Tone Control 263
The Baxandall Two-Capacitor HF Control 265
Switched-Frequency Baxandall Controls 266
Variable-Frequency LF and HF EQs 267
Tone-Balance Controls 269
Middle Controls 271
Fixed-Frequency Baxandall Middle Controls 271
Three-Band Baxandall EQ in One Stage 273
Wien Fixed Middle EQ 276
Sweep Middles 276
The Single-Gang Sweep Middle 278
Switched-Q Variable-Frequency Wien Middle EQs 280
Switchable Peak/Shelving LF/HF EQs 281
Parametric Middle EQs 283
Graphic Equalizers 286
References 289

Chapter 11 A Complete Preamplifier *291*
Architecture and Philosophy 291
The Line Inputs 293
The High-Impedance Input Buffer 293
The Tone-Control Stage 294
The Active Gain Stage 300
Performance 301
References 302

Chapter 12 Mixer Architectures *303*
Performance Factors 303
Mixer Internal Levels 304
Mixer Architecture 305
The Split Mixing Architecture 306
The In-Line Mixing Architecture 308
A Closer Look at Split-Format Modules 310
The Channel Module (Split Format) 310
Effects Return Modules 313
The Group Module 314
The Master Module 316
Talkback and Oscillator Systems 317
The In-Line Channel Module 319

www.raguvaran.puzl.com

Chapter 13 Microphone Preamplifiers *323*
Microphone Preamplifier Requirements 323
Transformer Microphone Inputs 324
The Simple Hybrid Microphone Preampilifier 325
The Balanced-Feedback Hybrid Microphone Preampilifier 326
Microphone and Line-Input Pads 329
The Padless Microphone Preampilifier 331
Capacitor Microphone Head Amplifiers 335

Chapter 14 Line Inputs *339*
External Signal Levels 339
Internal Signal Levels 339
Input Amplifier Functions 340
Unbalanced Inputs 340
Balanced Interconnections 344
The Advantages of Balanced Interconnections 345
The Disadvantages of Balanced Interconnections 345
Balanced Cables and Interference 346
Electrostatic Coupling 346
Magnetic Coupling 347
Ground Voltages 348
Balanced Connectors 348
Balanced Signal Levels 348
Electronic vs. Transformer Balanced Inputs 349
Common-Mode Rejection 349
The Basic Electronic Balanced Input 352
Common-Mode Rejection: The Basic Balanced Input and Op-Amp Effects 354
Op-Amp Frequency Response Effects 355
Op-Amp CMRR Effects 357
Amplifier Component Mismatch Effects 357
A Practical Balanced Input 359
Variations on the Balanced Input Stage 362
Combined Unbalanced and Balanced Inputs 362
The Superbal Input 363
Switched-Gain Balanced Inputs 364
Variable-Gain Balanced Inputs 365
High-Input-Impedance Balanced Inputs 366
The Inverting Two-Op-Amp Input 367
The Instrumentation Amplifier 368
Transformer Balanced Inputs 370
Input Overvoltage Protection 371
Low-Noise Balanced Inputs 372
Low-Noise Balanced Inputs in Real Life 377

www.raguvaran.puzl.com

Ultra-Low-Noise Balanced Inputs 378
References 380

Chapter 15 Line Outputs 381
Unbalanced Outputs 381
Zero-Impedance Outputs 382
Ground-Canceling Outputs 383
Balanced Outputs 386
Quasi-Floating Outputs 386
Transformer Balanced Outputs 388
Driving Heavy Loads 389
Reference 392

Chapter 16 Signal Switching 393
Mechanical Switches 393
Input-Select Switching 394
The Virtual Contact 395
Relay Switching 397
Electronic Switching 397
Switching with CMOS Analog Gates 397
CMOS Gates in Voltage Mode 399
CMOS Gates in Current Mode 404
Series-Shunt Current Mode 406
Control Voltage Feedthrough in CMOS Gates 407
Discrete JFET Switching 407
JFETs in Voltage Mode: The Series JFET Switch 408
The Shunt JFET Switch in Voltage Mode 413
JFETs in Current Mode 414
Reducing Distortion by Biasing 417
JFET Drive Circuitry 418
Physical Layout and Offness 419
Dealing with the DC Conditions 420
A Soft Changeover Circuit 420
Control-voltage feedthrough in JFETs 421
Reference 421

Chapter 17 Mixer Subsystems 423
Mixer Bus Systems 423
Input Arrangements 424
Equalization 424
Insert Points 424
How to Move a Circuit Block 426
Faders 428
Postfade Amplifiers 428

www.raguvaran.puzl.com

Direct Outputs 430
Panpots 431
Passive Panpots 432
The Active Panpot 438
LCR Panpots 439
Routing Systems 442
Auxiliary Sends 446
Group Module Circuit Blocks 446
Summing Systems 447
Voltage Summing 447
Virtual-Earth Summing 448
Balanced Summing Systems 449
Ground-Canceling Summing Systems 451
Distributed Summing Systems 452
Summing Amplifiers 456
Hybrid Summing Amplifiers 459
Balanced Hybrid Summing Amplifiers 462
PFL Systems 465
PFL Summing 466
PFL Switching 466
PFL Detection 468
Virtual-Earth PFL Detection 469
AFL Systems 471
Solo-in-Place Systems 473
Talkback Microphone Amplifiers 473
Line-Up Oscillators 474
Console Cooling and Component Lifetimes 477

Chapter 18 Level Indication and Metering 479
Signal-Present Indication 479
Peak Indication 479
Distributed Peak Detection 481
Combined LED Indicators 483
VU Meters 483
PPM Meters 485
LED Bar-Graph Metering 486
A More Efficient LED Bar-Graph Architecture 488
Vacuum Fluorescent Displays 491
LCD Meter Displays 491

Chapter 19 Level Control and Special Circuits 493
Gain-Control Elements 493
A Brief History 493
JFETs 493

www.raguvaran.puzl.com

Operational Transconductance Amplifiers 496
Voltage-Controlled Amplifiers 497
Compressors and Limiters 499
Attack Artefacts 503
Decay Artefacts 504
Subtractive VCA Control 505
Noise Gates 506
Clipping 508
Diode Clipping 508
Active Clipping with Transistors 510
Active Clipping with Op-Amps 512
Noise Generators 521
Pinkening Filters 523
References 524

Chapter 20 Power Supplies 525
Op-Amp Supply-Rail Voltages 525
Designing a ±15 V Supply 527
Designing a ±17 V Supply 529
Using Variable-Voltage Regulators 531
Improving Ripple Performance 532
Dual Supplies From a Single Winding 533
Power Supplies for Discrete Circuitry 534
Larger Power Supplies 534
Mutual Shutdown Circuitry 536
Very Large Power Supplies 536
Microcontroller and Relay Supplies 537
+48 V Phantom Power Supplies 538
Reference 540

Chapter 21 Interfacing with the Digital Domain 541
PCB Layout Considerations 541
Nominal Levels and ADCs 542
Some Typical ADCs 543
Interfacing with ADC Inputs 544
Some Typical DACs 546
Interfacing with DAC Outputs 547

Index 549

www.raguvaran.puzl.com

Preface

"Another damned thick book! Always scribble, scribble, scribble! Eh, Mr. Gibbon?"

Attributed to Prince William Henry, Duke of Gloucester, in 1781 upon receiving the second volume of *The History of the Decline and Fall of the Roman Empire* from its author.

This book deals with small-signal audio design: the amplification and control of audio in the analog domain, where the clever stuff is done with op-amps or discrete transistors, usually working at a nominal level of a volt or less. 'Small-signal design' is the opposite term to the large-signal technology used in power amplifiers, welding gear, and the electricity distribution grid.

Audio signals frequently need to be manipulated, by altering their level, frequency response, routing and so on. It is obviously much more economical to perform these functions on a small signal, rather than a large signal that has already been amplified up to loudspeaker level, and so has tens of volts and several amps to its name.

The total flexibility of digital signal processing may allow a greater scope of control – you might think how you would go about implementing a one-second delay in the analog domain, for example – but there are many times when greater quality or greater economy can be obtained by keeping the signal in the analog domain. Sometimes, of course, analog circuitry connects to the digital world, and so a complete chapter of this book deals with the subtleties of analog/digital interfacing.

I have devoted the first few chapters to the principles of high-quality small-signal design, moving on to look closely at, first, hi-fi preamplifiers, and then mixing consoles. These two genres were chosen partly because they are of wide interest in themselves, but mainly because they use a large number of different functional blocks, with very little overlap between them. They cover a wide range of circuit functions that will be useful for all kinds of audio systems. You will find out how to adapt or design these building blocks for audio, and how to put them together to form a system without bad things happening due to loading or interaction. You should then be able to design pretty much anything in this field.

www.raguvaran.puzl.com

It is my aim to provide information not available elsewhere. As an example, take the near-universal Baxandall bass and treble tone control. Many books give you the basic circuit, with a few conventional remarks, and that's it, but there are actually a huge number of useful variations on the basic circuit. To begin with, this book reveals that there are actually two different versions of the basic Baxandall control, with a quite different action at the low end of the audio spectrum. It goes on to show how to make a middle control using the Baxandall approach, and even how to fit three bands of tone control into one stage. From there we move on to a whole range of more sophisticated controls with variable frequency, Q, and so on.

You will also find material on how to make amplifiers with apparently impossibly low noise, how to design discrete transistor circuitry that can handle enormous signals with vanishingly low distortion, how to use humble low-gain transistors to make an amplifier with a startlingly high input impedance of more than 50 MΩ, how to transform the performance of low-cost op-amps, how to make filters with very low noise and distortion, how to make incredibly accurate volume controls, how to make magnetic cartridge preamplifiers that have noise so low it is limited by basic physics, and generally how to sum, switch, clip, compress, and route audio signals. Finally, the chapter on power supplies, in which I give some very practical ways to keep both the ripple and the cost down, shows how to power everything.

This book is mainly aimed at the design of audio equipment of the highest possible quality. This does not, fortunately, mean spending money like water and making everything out of unbendium or unobtainium to achieve this end. The limitations on performance are usually set by the laws of physics – which are proverbially hard to break – rather than the cost of components. This is not always the case and sometimes you have to grapple with a price/performance trade-off. I have kept the lupine from the lobby for many years by designing things to sell at a profit, so I have been able to include quite a few non-obvious ways to save money. Only at the so-called 'high end' of hi-fi is cost almost wholly irrelevant.

An eye for economy is not merely a negative influence. Possibly the man who said 'cost is an interesting extra constraint' was just trying to convince himself he hadn't sold out to Big Business, but to my mind there is a lot of truth in the saying. If you could always get better performance just by laying out more cash – well, where's the fun in that?

I have tried to ensure that all the essential points of theory that bear on practical performance are explained, with the mathematics kept to an essential minimum, but I have no intention of trying to create a basic electronics textbook – there are plenty of good ones about already. I do advise, however, that you get the basic theory straight before you try to make use of the circuitry and concepts here.

Small-signal design is very often based on op-amp circuitry, the great advantage being that all the tricky details of low-noise and almost distortion-free amplification are confined within the small black carapace of a 5532. This op-amp is, for very good reasons, employed more than

www.raguvaran.puzl.com

any other in quality audio design, and I make great use of it here. In the course of compiling this work, it has been brought home to me just what a bargain the 5532 is.

There are, however, times when circuitry built with discrete transistors gives better results. You may need more voltage swing than is available between op-amp supply rails, or it may be inconvenient to provide such rails, or you may want to do it purely for marketing purposes. I have therefore included a chapter on the design of discrete transistor stages. This contains a wealth of little-known techniques, ranging from some of the lesser-known quirks of the simple emitter-follower to the design of complete op-amps from discrete devices.

It is sometimes highly advantageous to combine discrete devices and IC op-amps in the same feedback loop, to exploit the best qualities of each. When you need very low noise, say for a moving-coil phono amplifier, you have transistors followed by an op-amp. When you need high output current, say for driving headphones, you have an op-amp followed by transistors. Many examples of this approach are scattered through the book.

An important strand in this book is low-impedance design. By minimizing circuit resistances the contribution of Johnson noise is reduced, and conditions set for the best semiconductor noise performance. Op-amp common-mode distortion and crosstalk are also reduced. The notion is not exactly new (as some people would have you believe) but has been used explicitly in audio circuitry for at least 20 years. This approach relies on amplifiers that can drive heavy loads at very low distortion, and here again the 5532 is extremely useful. In this book you will find many examples of low-impedance design.

What you will not find here is any truck with the religious dogma of audio Subjectivism: the directional cables, the oxygen-free copper, the World War One vintage triodes and all the other depressing paraphernalia of pseudo- and anti-science. I have spent more time than I care to contemplate in double-blind listening tests – properly conducted ones, with rigorous statistical analysis – and every time the answer was that if you couldn't measure it you couldn't hear it. If you want to know more about my experiences and reasoning in this area, there is a full discussion in my earlier book, the *Audio Power Amplifier Design Handbook*.

I have included a few historical vignettes both where they are of interest in themselves, and where they also shed light on some general principles of design.

I have tried to avoid FGAs (Frequent Gratuitous Acronyms) but a few are used extensively so more useful text can be squeezed in. They are listed after this Preface. None of them, I think, will cause any great confusion. I also use the word 'offness' – which is not found in any spell-checker but is widely used in the pro audio sector to refer to the ratio (in dB) by which a controlled signal can be suppressed. I try to give value for money, so introductory material has been cut to a minimum, and many a rolling elegiac sentence has been ruthlessly culled to fit in another hard fact.

www.raguvaran.puzl.com

In many places I have given measured performance figures or test-gear plots. These are inevitably the result of relatively few measurements and are intended to be illustrative rather than defining a figure for all time. When op-amp properties were involved I have been careful to use high-quality parts from the well-known manufacturers rather than me-too parts of doubtful provenance. Your mileage may vary, but I think not by much.

The data here has been gathered over several years. Some of the plots are from the now obsolete Audio Precision System 1, while others are from the newer and quite remarkable AP SYS-2702. It would have been a Sisyphean task to repeat all the measurements, so some graphs look a little agricultural graphically, but they have at least a certain gritty realism.

I have a website (www.dself.dsl.pipex.com) where I will be adding supplementary material to this book. I really mean this; I have already written another thirty pages or so on both discrete and opamp design significantly extending the material in this book. From there you can contact me, so please let me know if you think something is wrong, or an important topic is missing, or if you have any other suggestions as to how to improve this book that do not involve combustion.

Further information, and PCBs, kits, and built circuit boards of some of the designs described here, such as phono input stages and complete preamplifiers, can be found at:

www.signaltransfer.freeuk.com

It is my hope you will both enjoy this book and find it useful.

Douglas Self

London, September 2009

www.raguvaran.puzl.com

About the Author

Douglas Self has a world-wide reputation as a leading authority on audio power amplifier design, but it is perhaps less well-known that he has devoted a good deal of study to preamplifiers and other small-signal circuitry, including many years as the Chief Design Engineer at one of the major mixing console manufacturers, where his achievements included winning a Design Council Award. His rigorous, skeptical, and thoroughly practical approach to design has been applied to the small-signal area as well, and some of the results are to be found in this book.

Douglas studied engineering at Cambridge University, and then Psychoacoustics at Sussex University. He has spent many years working at the top level of design in both the professional audio and hi-fi industries, and has taken out a number of patents in the field of audio technology.

Dedication

To Julie, with all my love

www.raguvaran.puzl.com

Acronyms

ADC	Analog-to-digital converter
AFL	After-fade listen
AGS	Active gain stage
BFMA	Balanced-feedback microphone amplifier
BJT	Bipolar junction transistor
CFA	Current feedback amplifier
CFP	Complementary feedback pair
CM	Common mode
CMOS	Complementary metal oxide semiconductor
CMRR	Common-mode rejection ratio
CRM	Control-room monitor
DAC	Digital-to-analog converter
EF	Emitter-follower
EIN	Equivalent input noise
EQ	Equalization
ESR	Equivalent series resistance
ETP	Electrolytic tough pitch
FET	Field-effect transistor
FS	Full scale
GC	Ground canceling
HF	High frequency
IC	Integrated circuit
IDC	Insulation-displacement connector
JFET	Junction field-effect transistor
LED	Light-emitting diode
LF	Low frequency
MC	Moving coil
MM	Moving magnet
MOSFET	Metal oxide semiconductor field-effect transistor
NF	Noise figure
NFB	Negative feedback
OTA	Operational transconductance amplifier
PA	Public address
PCB	Printed-circuit board
PGA	Programmable gain amplifier
PFL	Prefade listen
PPM	Peak program meter
PSRR	Power-supply rejection ratio
RF	Radio frequency
RTF	Return to flat
RIAA	Recording Industry Association of America
SIP	Solo in place
SM	Surface mount
TH	Through hole
THD	Total harmonic distortion
VAS	Voltage-amplifier stage
VCA	Voltage-controlled amplifier
VCVS	Voltage-controlled voltage source

www.raguvaran.puzl.com

CHAPTER 1

The Basics

Signals

An audio signal can be transmitted either as a voltage or as a current. The construction of the universe is such that almost always the voltage mode is more convenient; consider for a moment an output driving more than one input. Connecting a series of high-impedance inputs to a low-impedance output is simply a matter of connecting them in parallel, and if the ratio of the output and input impedances is high there will be negligible variations in level. To drive multiple inputs with a current output it is necessary to have a series of floating current-sensor circuits that can be connected in series. This can be done [1], as pretty much anything in electronics can be done, but it requires a lot of hardware, and probably introduces performance compromises. The voltage-mode connection is just a matter of wiring.

Obviously, if there's a current, there's a voltage, and vice versa. You can't have one without the other. The distinction is in the output impedance of the transmitting end (low for voltage mode, high for current mode) and in what the receiving end responds to. Typically, but not necessarily, a voltage input has a high impedance; if its input impedance was only 600 Ω, as used to be the case in very old audio distribution systems, it is still responding to voltage, with the current it draws doing so a side issue, so it is still a voltage amplifier. In the same way, a current input typically, but not necessarily, has a very low input impedance. Current outputs can also present problems when they are not connected to anything. With no terminating impedance, the voltage at the output will be very high, and probably clipping heavily; the distortion is likely to crosstalk into adjacent circuitry. An open-circuit voltage output has no analogous problem.

Current-mode connections are not common. One example is the Krell Current Audio Signal Transmission (CAST) technology, which uses current mode to interconnect units in the Krell product range. While it is not exactly audio, the 4–20 mA current loop format is widely used in instrumentation. The current-mode operation means that voltage drops over long cable runs are ignored, and the zero offset of the current (i.e. 4 mA = zero) makes cable failure easy to detect: if the current suddenly drops to zero, you have a broken cable.

Small Signal Audio Design; ISBN: 9780240521770
Copyright © 2010 Elsevier Ltd; All rights of reproduction, in any form, reserved.

www.raguvaran.puzl.com

The old DIN interconnection standard was a form of current-mode connection in that it had voltage output via a high output impedance, of 100 kΩ or more. The idea was presumably that you could scale the output to a convenient voltage by selecting a suitable input impedance. The drawback was that the high output impedance made the amount of power transferred very small, leading to a poor signal-to-noise ratio. The concept is now wholly obsolete.

Amplifiers

At the most basic level, there are four kinds of amplifier, because there are two kinds of signal (voltage and current) and two types of port (input and output). The handy word 'port' glosses over whether the input or output is differential or single-ended. Amplifiers with differential input are very common – such as all op-amps and most power amps – but differential outputs are rare and normally confined to specialized telecoms chips.

The four kinds of amplifier are summarized in Table 1.1.

Voltage Amplifiers

These are the vast majority of amplifiers. They take a voltage input at a high impedance and yield a voltage output at a low impedance. All conventional op-amps are voltage amplifiers in themselves, but they can be made to perform as any of the four kinds of amplifier by suitable feedback connections. Figure 1.1(a) shows a high-gain voltage amplifier with series voltage feedback. The closed-loop gain is (R1 + R2)/R2.

Transconductance Amplifiers

The name simply means that a voltage input (usually differential) is converted to a current output. It has a transfer ratio $A = I_{out}/V_{in}$, which has dimensions of I/V or conductance, so it is referred to as a transconductance or, less commonly, a transadmittance amplifier. It is possible to make a very simple, though not very linear, voltage-controlled amplifier with transconductance technology: differential-input operational transconductance amplifier (OTA) Integrated circuits (ICs) have an extra pin that gives voltage control of the transconductance, which when used with no negative feedback gives gain control (see Chapter

TABLE 1.1 The four types of amplifier

Amplifier type	Input	Output	Application
Voltage amplifier	Voltage	Voltage	General amplification
Transconductance amplifier	Voltage	Current	Voltage control of gain
Current amplifier	Current	Current	?
Transimpedance amplifier	Current	Voltage	Summing amplifiers, DAC interfacing

www.raguvaran.puzl.com

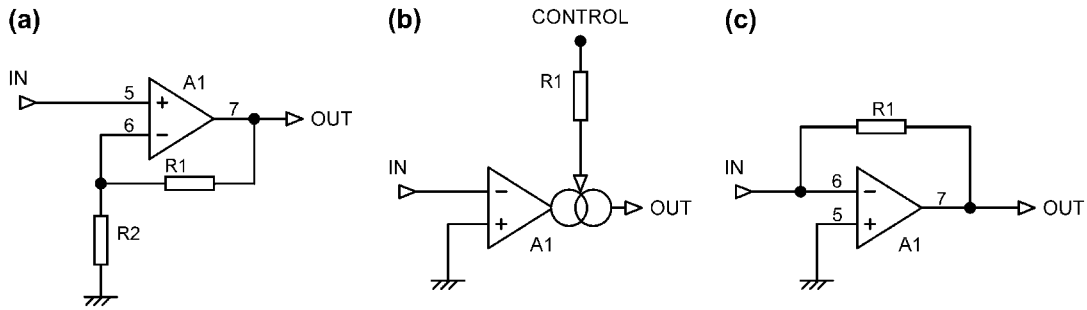


Figure 1.1: (a) A voltage amplifier. (b) A transconductance amplifier. (c) A transimpedance amplifier

19 for details). Performance falls well short of that required for quality hi-fi or professional audio. Figure 1.1(b) shows an OTA used without feedback; note the current-source symbol at the output.

Current Amplifiers

These accept a current in, and give a current out. Since, as we have already noted, current-mode operation is rare, there is not often a use for a true current amplifier in the audio business. They should not be confused with current feedback amplifiers (CFAs), which have a voltage output, the 'current' bit referring to the way the feedback is applied in current mode [2]. The bipolar transistor is sometimes described as a current amplifier, but it is nothing of the kind. Current may flow in the base circuit but this is just an unwanted side-effect. It is the *voltage* on the base that actually controls the transistor.

Transimpedance Amplifiers

A transimpedance amplifier accepts a current in (usually single-ended) and gives a voltage out. It is sometimes called an *I–V* converter. It has a transfer ratio $A = V_{out}/I_{in}$, which has dimensions of *V*/*I* or resistance. That is why it is referred to as a transimpedance or transresistance amplifier. Transimpedance amplifiers are usually made by applying shunt voltage feedback to a high-gain voltage amplifier. An important use is as virtual-earth summing amplifiers in mixing consoles (see Chapter 17). The voltage-amplifier stage (VAS) in most power amplifiers is a transimpedance amplifier. They are used for *I–V* conversion when interfacing to digital-to-analog converters (DACs) with current outputs (see Chapter 21). Transimpedance amplifiers are sometimes incorrectly described as 'current amplifiers'.

Figure 1.1(c) shows a high-gain voltage amplifier transformed into a transimpedance amplifier by adding the shunt voltage feedback resistor R1. The transimpedance gain is simply the value of R1, though it is normally expressed in V/mA rather than ohms.

www.raguvaran.puzl.com

Negative Feedback

Negative feedback is one of the most useful and omnipresent concepts in electronics. It can be used to control gain, to reduce distortion and improve frequency response, and to set input and output impedances, and one feedback connection can do all these things at the same time. Negative feedback comes in four basic modes, as in the four basic kinds of amplifier. It can be taken from the output in two different ways (voltage or current feedback) and applied to the amplifier input in two different ways (series or shunt). Hence there are four combinations.

However, unless you're making something exotic like an audio constant-current source, the feedback is always taken as a voltage from the output, leaving us with just two feedback types, series and shunt, both of which are extensively used in audio. When series feedback is applied to a high-gain voltage amplifier, as in Figure 1.1(a), the following statements are true:

- Negative feedback reduces voltage gain.
- Negative feedback increases gain stability.
- Negative feedback increases bandwidth.
- Negative feedback increases amplifier input impedance.
- Negative feedback reduces amplifier output impedance.
- Negative feedback reduces distortion.
- Negative feedback does not directly alter the signal-to-noise ratio.

If shunt feedback is applied to a voltage amplifier to make a transimpedance amplifier, as in Figure 1.1(c), all the above statements are still true, except since we have applied shunt rather than series negative feedback, the input impedance is reduced.

The basic feedback relationship is Equation 1.1, which is dealt with at length in any number of textbooks, but it is of such fundamental importance that I feel obliged to include it here. The open-loop gain of the amplifier is A, and β is the feedback fraction, such that if in Figure 1.1(a) R1 is 2 kΩ and R2 is 1 kΩ, β is $\frac{1}{3}$. If A is very high, you don't even need to know it; the 1 on the bottom becomes negligible, and the As on the top and bottom cancel out, leaving us with a gain of almost exactly 3.

$$\frac{V_{\text{out}}}{V_{\text{in}}} = \frac{A}{1 + A\beta} \qquad \text{(Equation 1.1)}$$

Negative feedback can, however, do much more than stabilizing gain. Anything unwanted occurring in the amplifier, be it distortion or DC drift, or any of the other ills that electronics is

www.raguvaran.puzl.com

prone to, is also reduced by the negative feedback factor (NFB factor for short). This is equal to:

$$\text{NFB factor} = \frac{1}{1 + A\beta} \qquad \text{(Equation 1.2)}$$

What negative feedback cannot do is improve the noise performance. When we apply feedback the gain drops, and the noise drops by the same factor, leaving the signal-to-noise ratio the same. Negative feedback and the way it reduces distortion is explained in much more detail in one of my other books [3].

Nominal Signal Levels

The absolute level of noise in a circuit is not of great significance in itself – what counts is how much greater the signal is than the noise – in other words the signal-to-noise ratio. An important step in any design is the determination of the optimal signal level at each point in the circuit. Obviously a real audio signal, as opposed to a test sine wave, continuously varies in amplitude, and the signal level chosen is purely a nominal level. One must steer a course between two evils:

- If the signal level is too low, it will be contaminated unduly by noise.
- If the signal level is too high, there is a risk it will clip and introduce severe distortion.

You will note that the first evil is a certainty, while the second is more of a statistical risk. The consequences of either must be considered when choosing a level, and the wider the gap between them the greater the dynamic range. If the best possible signal-to-noise is required in a studio recording, then the internal level must be high, and if there is an unexpected overload you can always do another take. In live situations it will often be preferable to sacrifice some noise performance to give less risk of clipping. The internal signal levels of mixing consoles are examined in detail in Chapter 12.

If you seek to increase the dynamic range, you can either increase the maximum signal level or lower the noise floor. The maximum signal levels in op-amp-based equipment are set by the voltage capabilities of the op-amps used, and this usually means a maximum signal level of about 10 Vrms or +22 dBu. Discrete transistor technology removes the absolute limit on supply voltage, and allows the voltage swing to be at least doubled before the supply rail voltages get inconveniently high. For example, ±40 V rails are quite practical for small-signal transistors and permit a theoretical voltage swing of 28 Vrms or +31 dBu. However, in view of the complications of designing your own discrete circuitry, and the greater space and power it requires, those nine extra decibels of headroom are dearly bought.

www.raguvaran.puzl.com

Gain Structures

There are some very basic rules for putting together an effective gain structure in a piece of equipment. Like many rules, they are subject to modification or compromise when you get into a tight corner. Breaking them reduces the dynamic range of the circuitry, either by worsening the noise or restricting the headroom; whether this is significant depends on the overall structure of the system and what level of performance you are aiming at. Three simple rules are:

1. Don't amplify then attenuate.
2. Don't attenuate then amplify.
3. The signal should be raised to the nominal internal level as soon as possible to minimize contamination with circuit noise.

Amplification Then Attenuation

Put baldly it sounds too silly to contemplate, but it is easy to thoughtlessly add a bit of gain to make up for a loss somewhere else, and immediately a few decibels of precious and irretrievable headroom are gone for good. This assumes that each stage has the same power rails and hence the same clipping point, which is usually the case in op-amp circuitry.

Figure 1.2(a) shows a system with a gain control designed to keep 10 dB of gain in hand. In other words, the expectation is that the control will spend most of its working life set

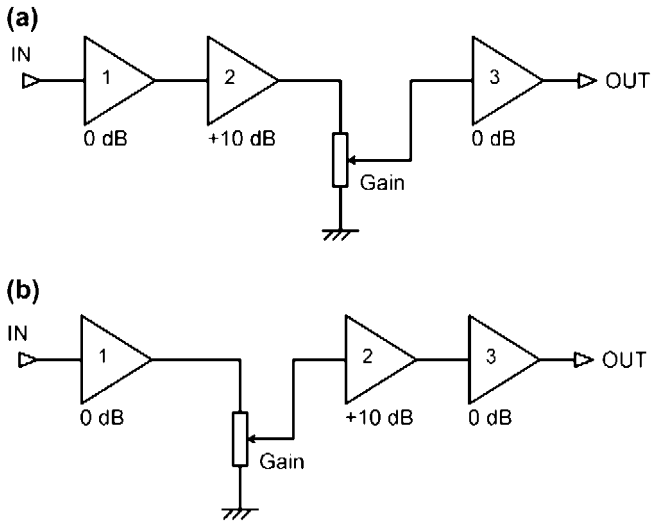


Figure 1.2: (a) Amplification then attenuation. Stage 2 will always clip first, reducing headroom. (b) Attenuation then amplification. The noise from Stage 2 degrades the signal-to-noise ratio. The lower the gain setting, the worse the effect

www.raguvaran.puzl.com

somewhere around its '0 dB' position where it introduces 10 dB of attenuation, as is typically the case for a fader on a mixer. To maintain the nominal signal level at 0 dBu we need 10 dB of gain, and a +10 dB amplifier (Stage 2) has been inserted just before the gain control. This is not a good decision. This amplifier will clip 10 dB before any other stage in the system, and introduces what one might call a headroom bottleneck.

There are exceptions. The moving-coil phono head-amp described in Chapter 8 appears to flagrantly break this rule, as it always works at maximum gain even when this is not required. But when considered in conjunction with the following RIAA stage, which also has considerable gain, it makes perfect sense, for the stage gains are configured so that the second stage always clips first, and there is actually no loss of headroom.

Attenuation Then Amplification

In Figure 1.2(b) the amplifier is now after the gain control, and noise performance rather than headroom suffers. If the signal is attenuated, any active device will inescapably add noise in restoring the level. Any conventional gain-control block has to address this issue. If we once more require a gain variable from +10 dB to off, i.e. $-\infty$ dB, as would be typical for a fader or volume control, then usually the potentiometer is placed before the gain stage as in Figure 1.2(b) because as a rule some loss in noise performance is more acceptable than a permanent 10 dB reduction in system headroom. If there are options for the amplifier stages in terms of a noise/cost trade-off (such as using the 5532 versus a TL072) and you can only afford one low-noise stage, then it should be Stage 2.

If all stages have the same noise performance this configuration is 10 dB noisier than the previous version when gain is set to 0 dB.

Raising the Input Signal to the Nominal Level

Getting the incoming signal up to the nominal internal level in one jump is always preferable as it gives the best noise performance. Sometimes it has to be done in two amplifier stages; typical examples are microphone preamps with wide gain ranges and phono preamps that insist on performing the RIAA equalization in several goes. (These are explored in their respective chapters.) In these cases the noise contribution of the second stage may no longer be negligible.

Consider a signal path which has an input of −10 dBu and a nominal level of 0 dBu. The first version has an input amplifier with 10 dB of gain followed by two unity-gain circuit blocks, A and B. All circuit blocks are assumed to introduce noise at −100 dBu. The noise output for the first version is −89.2 dBu. Now take a second version of the signal path that has an input amplifier with 5 dB of gain, followed by block A, another amplifier with 5 dB of gain, then block B. The noise output is now −87.5 dB, 1.7 dB worse, due to the extra amplification

www.raguvaran.puzl.com

of the noise from block A. There is also more hardware, and the second version is clearly an inferior design.

Active Gain Controls

The previous section should not be taken to imply that noise performance must always be sacrificed when a gain control is included in the signal path. This is not so. If we move beyond the idea of a fixed-gain block, and recognize that the amount of gain present can be varied, then less gain when the maximum is not required will reduce the noise generated. For volume-control purposes it is essential that the gain can be reduced to near-zero, though it is not necessary for it to be as firmly 'off' as the faders or sends of a mixer.

An active volume-control stage gives lower noise at lower volume settings because there is less gain. The Baxandall active configuration also gives excellent channel balance as it depends solely on the mechanical alignment of a dual linear pot – all mismatches of its electrical characteristics are canceled out, and there are no quasi-log dual slopes to induce anxiety.

Active gain controls are looked at in depth in Chapter 9.

Noise

Noise here refers only to the random noise generated by resistances and active devices. The term is sometimes used to include mains hum, spurious signals from demodulated RF and other non-random sources, but this threatens confusion and I prefer to call the other unwanted signals 'interference'. In one case we strive to minimize the random variations arising in the circuit itself, in the other we are trying to keep extraneous signals out, and the techniques are wholly different.

When noise is referred to in electronics it means white noise unless it is specifically labeled as something else, because that is the form of noise that most electronic processes generate. There are two elemental noise mechanisms that make themselves felt in all circuits and active devices. These are Johnson noise and shot noise, which are both are forms of white noise. Both have Gaussian probability density functions. These two basic mechanisms generate the noise in both bipolar junction transistors (BJTs) and field-effect transistors (FETs), though in rather different ways.

There are other forms of noise that originate from less fundamental mechanisms such as device processing imperfections that do not have a white spectrum; examples are $1/f$ (flicker) noise and popcorn noise. These noise mechanisms are described later in this chapter.

Non-white noise is given a color that corresponds to the visible spectrum; thus, red noise has a larger low-frequency content than white noise, while pink is midway between the two.

www.raguvaran.puzl.com

White noise has equal power in equal absolute bandwidth, i.e. with the bandwidth measured in Hz. Thus, there is the same power between 100 and 200 Hz as there is between 1100 and 1200 Hz. It is the type produced by most electronic noise mechanisms.

Pink noise has equal power in equal ratios of bandwidth, so there is the same power between 100 and 200 Hz as there is between 200 and 400 Hz. The energy per Hz falls at 3 dB per octave as frequency increases. Pink noise is widely used for acoustic applications like room equalization and loudspeaker measurement as it gives a flat response when viewed on a third-octave or other constant-percentage-bandwidth spectrum analyzer.

Red noise has energy per Hz falling at 6 dB per octave rather than 3. It is important in the study of stochastic processes and climate models, but has little application in audio. The only place you are likely to encounter it is in the oscillator section of analog synthesizers. It is sometimes called Brownian noise as it can be produced by Brownian motion; hence its alternative name of random-walk noise. Brown here is a person and not a color.

Blue noise has energy per Hz rising at 3 dB per octave. Blue noise is used for dithering in image anti-aliasing, but has, as far as I am aware, no application to audio. The spectral density of blue noise (i.e. the power per Hz) is proportional to the frequency. It appears that the light-sensitive cells in the retina of the mammalian eye are arranged in a pattern that resembles blue noise [4]. Great stuff, this evolution.

Violet noise has energy per Hz rising at 6 dB per octave (I imagine you saw that one coming). It is also known as 'differentiated white noise' as a differentiator circuit has a frequency response rising at 6 dB per octave. It is sometimes called purple noise.

Gray noise is pink noise modified by a psychoacoustic equal loudness curve, such as the inverse of the A-weighting curve, to give the perception of equal loudness at all frequencies.

Green noise really does exist, though not in the audio domain. It is used for stochastic half-toning of images, and consists of binary dither patterns composed of homogeneously distributed minority pixel clusters. I think we had better leave it there.

Johnson Noise

Johnson noise is produced by all resistances, including those real resistances hiding inside transistors (such as r_{bb}, the base spreading resistance). It is not generated by the so-called intrinsic resistances, such as r_e, which are an expression of the V_{be}/I_c slope and not a physical resistance at all. Given that Johnson noise is present in every circuit, and often puts a limit on noise performance, it is a bit surprising that it was not discovered until 1928 by John B. Johnson at Bell Labs [5].

www.raguvaran.puzl.com

The rms amplitude of Johnson noise is easily calculated with the classic formula:

$$v_n = \sqrt{4kTRB} \qquad \text{(Equation 1.3)}$$

where v_n is the rms noise voltage, T is absolute temperature (°K), B is the bandwidth in Hz, k is Boltzmann's constant, and R is the resistance in ohms.

The only thing to be careful with here (apart from the usual problem of keeping the powers of 10 straight) is to make sure you use Boltzmann's constant (1.380662×10^{-23}), and *not* the Stefan–Boltzmann constant (5.67×10^{-8}), which relates to black-body radiation and will give some spectacularly wrong answers. Often the voltage noise is left in its squared form for ease of summing with other noise sources. Table 1.2 gives a feel for how resistance affects the magnitude of Johnson noise. The temperature is 25°C and the bandwidth is 22 kHz.

Johnson noise theoretically goes all the way to daylight, but in the real world is ultimately band-limited by the shunt capacitance of the resistor. Johnson noise is not produced by circuit reactances – i.e. pure capacitance and inductance. In the real world, however, reactive components are not pure, and the winding resistances of transformers can produce significant Johnson noise; this is an important factor in the design of moving-coil cartridge step-up

TABLE 1.2 Resistances and their Johnson noise

Resistance (Ω)	Noise voltage (μV)	Noise voltage (dBu)	Application
1	0.018	−152.2	Moving-coil cartridge impedance (low output)
3.3	0.035	−147.0	Moving-coil cartridge impedance (medium output)
10	0.060	−142.2	Moving-coil cartridge impedance (high output)
47	0.13	−135.5	Line output isolation resistor
100	0.19	−132.2	Output isolation or feedback network
150	0.23	−130.4	Dynamic microphone source impedance
200	0.27	−129.2	Dynamic microphone source impedance (older)
600	0.47	−124.4	The ancient matched-line impedance
1000	0.60	−122.2	A nice round number
2500	0.95	−118.2	Worst-case output impedance of 10 kΩ pot
5000	1.35	−115.2	Worst-case output impedance of 20 kΩ pot
12,500	2.13	−111.2	Worst-case output impedance of 50 kΩ pot
25,000	3.01	−100.2	Worst-case output impedance of 100 kΩ pot
1 meg (10^6)	19.0	−92.2	Another nice round number
1 giga (10^9)	190	−62.2	As used in capacitor microphone amplifiers
1 tera (10^{12})	1900	−32.2	Insulation testers read in tera-ohms
1 peta (10^{15})	19,000	−2.2	OK, it's getting silly now

www.raguvaran.puzl.com

transformers. Capacitors with their very high leakage resistances approach perfection much more closely, and the capacitance has a filtering effect. They usually have no detectable effect on noise performance, and in some circuitry it is possible to reduce noise by using a capacitative potential divider instead of a resistive one [6].

The noise voltage is of course inseparable from the resistance, so the equivalent circuit is of a voltage source in series with the resistance present. While Johnson noise is usually represented as a voltage, it can also be treated as a Johnson noise current, by means of the Thevenin–Norton transformation, which gives the alternative equivalent circuit of a current source in shunt with the resistance. The equation for the noise current is simply the Johnson voltage divided by the value of the resistor it comes from:

$$i_{\mathrm{n}} = v_{\mathrm{n}}/R$$

When it is first encountered, this ability of resistors to generate electricity from out of nowhere seems deeply mysterious. You wouldn't be the first person to think of connecting a small electric motor across the resistance and getting some useful work out – and you wouldn't be the first person to discover it doesn't work. If it did, then by the First Law of Thermodynamics (the law of conservation of energy), the resistor would have to get colder, and such a process is flatly forbidden by … the Second Law of Thermodynamics. The Second Law is no more negotiable than the First Law, and it says that energy cannot be extracted by simply cooling down one body. If you could it would be what thermodynamicists call a perpetual motion machine of the second kind, and they are no more buildable than the common sort of perpetual motion machine.

It is interesting to speculate what happens as the resistor is made larger. Does the Johnson voltage keep increasing, until there is a hazardous voltage across the resistor terminals? Obviously not, or picking up any piece of plastic would be a lethal experience. Johnson noise comes from a source impedance equal to the resistor generating it, and this alone would prevent any problems. Table 1.2 ends with a couple of silly values to see just how this works; the square root in the equation means that you need a peta-ohm resistor ($1 \times 10^{15}\ \Omega$) to reach even 600 mVrms of Johnson noise. Resistors are made up to at least 100 GΩ but peta-ohm resistors (PΩ?) would really be a minority interest.

Shot Noise

It is easy to forget that an electric current is not some sort of magic fluid, but is actually composed of a finite though usually very large number of electrons, so current is in effect quantized. Shot noise is so called because it allegedly sounds like a shower of lead shot being poured on to a drum, and the name emphasizes the discrete nature of the charge carriers. Despite the picturesque description the spectrum is still that of white noise, and the noise current amplitude for a given steady current is described by a surprisingly simple equation

www.raguvaran.puzl.com

TABLE 1.3 How shot noise varies with current

Current (DC)	Current noise (nA_{rms})	Fluctuation (%)	R (Ω)	Voltage noise (μV)	Voltage noise (dBu)
1 pA	0.000084	8.4	100	8.4×10^{-6}	−219.3
1 nA	0.0026	0.27	100	0.000265	−189.3
1 μA	0.084	0.0084	100	0.0084	−159.3
1 mA	2.65	0.00027	100	0.265	−129.3
1 A	84	0.000008	100	8.39	−99.3

(as Einstein said, the most incomprehensible thing about the universe is that it is comprehensible) that runs thus:

$$\text{Noise current } i_n = \sqrt{ZqI_{dc}B} \qquad \text{(Equation 1.4)}$$

where q is the charge on an electron (1.602×10^{-19} C), I_{dc} is the mean value of the current, and B is the bandwidth examined.

As with Johnson noise, often the shot noise is left in its squared form for ease of summing with other noise sources. Table 1.3 helps to give a feel for the reality of shot noise. As the current increases, the shot noise increases too, but more slowly as it depends on the square root of the DC current; therefore, the *percentage* fluctuation in the current becomes less. It is the small currents that are the noisiest.

The actual level of shot noise voltage generated if the current noise is assumed to flow through a 100 Ω resistor is rather low, as the last column shows. There are very few systems that will be embarrassed by an extra noise source of even −99 dBu unless it occurs right at the very input. To generate this level of shot noise requires 1 A to flow through 100 Ω, which naturally means a voltage drop of 100 V and 100 W of power dissipated. These are not often the sort of circuit conditions that exist in preamplifier circuitry. This does not mean that shot noise can be ignored completely, but it can usually be ignored unless it is happening in an active device where the noise is amplified.

1/*f* Noise (Flicker Noise)

This is so called because it rises in amplitude proportionally as the frequency examined falls. Unlike Johnson noise and shot noise, it is not a fundamental consequence of the way the universe is put together, but the result of imperfections in device construction. 1/*f* noise appears in all kinds of active semiconductors, and also in resistors. For a discussion of flicker noise in resistors, see Chapter 2.

www.raguvaran.puzl.com

Popcorn Noise

This form of noise is named after the sound of popcorn being cooked, rather than eaten. It is also called burst noise or bistable noise, and is a type of low-frequency noise that appears primarily in integrated circuits, appearing as low-level step changes in the output voltage, occurring at random intervals. Viewed on an oscilloscope this type of noise shows bursts of changes between two or more discrete levels. The amplitude stays level up to a corner frequency, at which point it falls at a rate of $1/f^2$. Different burst-noise mechanisms within the same device can exhibit different corner frequencies. The exact mechanism is poorly understood, but is known to be related to the presence of heavy-metal ion contamination, such as gold. This is very much a quality-control issue at the device fabrication stage, and the only measure that can be taken against it is to examine the reputation of each potential component supplier. Like $1/f$ noise, popcorn noise does not have a Gaussian amplitude distribution.

Summing Noise Sources

When random noise from different sources is summed, the components do not add in a $2 + 2 = 4$ manner. Since the noise components come from different sources, with different versions of the same physical processes going on, they are uncorrelated and will partially reinforce and partially cancel, so root-mean-square (rms) addition holds, as shown in Equation 1.5. If there are two noise sources with the same level, the increase is 3 dB rather than 6 dB. When we are dealing with two sources in one device, such as a bipolar transistor, the assumption of no correlation is slightly dubious, because some correlation is known to exist, but it does not seem to be enough to cause serious calculation errors.

$$V_{\text{ntot}} = \sqrt{(V_{\text{n1}}^2 + V_{\text{n2}}^2 + \cdots)} \qquad \text{(Equation 1.5)}$$

Any number of noise sources may be summed in the same way, by simply adding more squared terms inside the square root, as shown by the dotted lines. When dealing with noise in the design process, it is important to keep in mind the way that noise sources add together when they are not of equal amplitude. Table 1.4 shows how this works in decibels. Two equal voltage noise sources give a sum of +3 dB, as expected. What is notable is that when the two sources are of rather unequal amplitude, the smaller one makes very little contribution to the result.

If we have a circuit in which one source is twice the rms amplitude of the other, i.e. with a 6 dB difference, then it only increases the sum by 0.97 dB, a change that would be barely detectable on critical listening. If one source is 10 dB below the other, it only raises the result by 0.4 dB, which in most cases could be ignored. At 20 dB down, the extra noise contribution is lost in measurement error.

www.raguvaran.puzl.com

TABLE 1.4 The summation of uncorrelated noise sources

dB	dB	dB sum
0	0	+3.01
0	−1	+2.54
0	−2	+2.12
0	−3	+1.76
0	−4	+1.46
0	−5	+1.19
0	−6	+0.97
0	−10	+0.41
0	−15	+0.14
0	−20	+0.04

This mathematical property of uncorrelated noise sources is exceedingly convenient, because it means that in practical calculations we can neglect all except the most important noise sources with minimal error. Since all semiconductors have some variability in their noise performance, it is rarely worthwhile to make the calculations to great accuracy.

Noise in Amplifiers

There are basic principles of noise design that apply to all amplifiers, be they discrete or integrated, single-ended or differential. Practical circuits, even those consisting of an op-amp and two resistors, have multiple sources of noise. Typically one source of noise will dominate, but this cannot be taken for granted and it is essential to evaluate all the sources and the ways that they add together if a noise calculation is going to be reliable. Here I add the complications one stage at a time.

Figure 1.3 shows that most useful of circuit elements, the perfect noiseless amplifier. (These seem to be unaccountably hard to find in catalogs.) It is assumed to have a definite

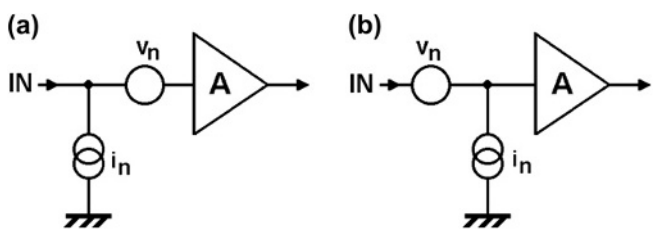


Figure 1.3: The noise sources of a perfect amplifier. The two circuits are exactly equivalent

www.raguvaran.puzl.com

gain without bothering about whether it is achieved by feedback or not, and an infinite input impedance. To emulate a real amplifier noise sources are concentrated at the input, combined into one voltage noise source and one current noise source. These can represent any number of actual noise sources inside the real amplifier. Figure 1.3 shows two ways of drawing the same situation.

It does not matter on which side of the voltage source the current source is placed; the 'perfect' amplifier has an infinite input impedance, and the voltage source a zero impedance, so either way all of the current noise flows through whatever is attached to the input.

Figure 1.4 shows the first step to a realistic situation, with a signal source now connected to the amplifier input. The signal source is modeled as a perfect zero-impedance voltage source with added series resistance R_s. Many signal sources are modeled accurately enough for noise calculations in this way. Examples are low-impedance dynamic microphones, moving-coil phono cartridges, and most electronic outputs. In others cases, such as moving-magnet phono cartridges and capacitor microphone capsules, there is a big reactive component that has a major effect on the noise behavior and cannot be ignored or treated as a resistor. The magnitude of the reactances tends to vary from one make to another, but fortunately the variations are not great enough for the circuit approach for optimal noise to vary greatly. It is pretty clear that a capacitor microphone will have a very high source impedance at audio frequencies, and will need a special high-impedance preamplifier to avoid low-frequency roll-off (see Chapter 13 for details). It is perhaps less obvious that the series inductance of a moving-magnet phono cartridge becomes the dominating factor at the higher end of the audio band, and designing for the lowest noise with the 600 Ω or so series resistance alone will give far from optimal results. This is dealt with in Chapter 7.

There are two sources of voltage noise in the circuit of Figure 1.4:

1. the amplifier voltage noise source v_n at the input;
2. the Johnson noise from the source resistance R_s.

These two voltage sources are in series and sum by rms addition as they are uncorrelated.

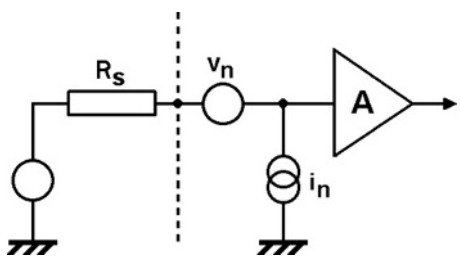


Figure 1.4: The perfect amplifier and noise sources with a signal source now connected

www.raguvaran.puzl.com

There is only one current noise component: the amplifier noise current source i_n across the input. This generates a noise voltage when its noise current flows through R_s. (It cannot flow into the amplifier input because we are assuming an infinite input impedance.) This third source of voltage noise is also added in by rms addition, and the total is amplified by the voltage gain A and appears at the output. The noise voltage at the input is the equivalent input noise (EIN). This is not easy to measure, so the noise at the amplifier output is divided by A to get the EIN. Having got this, we can compare it with the Johnson noise from the source resistance R_s; with a noiseless amplifier there would be no difference, but in real life the EIN will be higher by a number of decibels, which is called the noise figure (NF). This gives a concise way of assessing how noisy our amplifier is and if it is worth trying to improve it. Noise figures rarely appear in hi-fi literature, probably because most of them wouldn't look very good. For the application of noise figures to phono cartridge amplifiers, see Chapters 7 and 8.

Noise in Bipolar Transistors

An analysis of the noise behavior of discrete bipolar transistors can be found in many textbooks, so this is something of a quick summary of the vital points. Two important transistor parameters for understanding noise are r_{bb}, the base spreading resistance, and r_e, the intrinsic emitter resistance. r_{bb} is a real physical resistance – what is called an *extrinsic* resistance. The second parameter r_e is an expression of the V_{be}/I_c slope and not a physical resistance at all, so it is called an *intrinsic* resistance.

Noise in bipolar transistors, as in amplifiers in general, is best dealt with by assuming a noiseless transistor with a theoretical noise voltage source in series with the base and a theoretical noise current source connected from base to emitter. These sources are usually described simply as the 'voltage noise' and the 'current noise' of the transistor.

Bipolar Transistor Voltage Noise

The voltage noise v_n is made up of two components:

1. the Johnson noise generated in the base spreading resistance r_{bb};
2. the collector current (I_c) shot noise creating a noise voltage across r_e, the intrinsic emitter resistance.

These two components can be calculated from the equations given earlier, and rms-summed thus:

$$\text{Voltage noise density } v_n = \sqrt{4kTr_{bb} + 2(kT)^2/(qI_c)} \text{ (in V}/\sqrt{\text{Hz}}\text{; usually nV}/\sqrt{\text{Hz}}\text{)} \quad \text{(Equation 1.6)}$$

www.raguvaran.puzl.com

where k is Boltzmann's constant (1.380662×10^{-23}), q is the charge on an electron (1.602×10^{-19} C), T is absolute temperature (°K), I_c is the collector current, and r_{bb} is the base resistance in ohms.

The first part of this equation is the usual expression for Johnson noise, and is fixed for a given transistor type by the physical value of r_{bb}, so the lower this is the better. The absolute temperature is a factor; running your transistor at 25°C rather than 125°C reduces the Johnson noise from r_{bb} by 1.2 dB. Input devices usually run cool but this may not be the case with moving-coil preamplifiers where a large I_c is required, so it is not impossible that adding a heat-sink would give a measurable improvement in noise.

The second (shot noise) part of the equation decreases as collector current I_c increases; this is because as I_c increases, r_e decreases proportionally, following $r_e = 25/I_c$, where I_c is in mA. The shot noise, however, is only increasing as the square root of I_c, and the overall result is that the total v_n falls – though relatively slowly – as collector current increases, approaching asymptotically the level of noise set by the first part of the equation. The only way you can reduce this is by changing to another transistor type with a lower r_{bb}.

There is an extra voltage noise source resulting from flicker noise produced by the base current flowing through r_{bb}; this is only significant at high collector currents and low frequencies due to its $1/f$ nature, and is not usually included in design calculations unless low-frequency quietness is a special requirement.

Bipolar Transistor Current Noise

The current noise i_n is mainly produced by the shot noise of the steady current I_b flowing through the transistor base. This means it increases as the square root of I_b increases. Naturally I_b increases with I_c. Current noise is given by:

$$\text{Current noise density } i_n = \sqrt{2qI_b} \text{ (in A}/\sqrt{\text{Hz}}\text{; usual values are in pA)} \qquad \text{(Equation 1.7)}$$

where q is the charge on an electron and I_b is the base current.

So, for a fixed collector current, you get less current noise with high-beta transistors because there is less base current.

The existence of current noise as well as voltage noise means it is not possible to minimize transistor noise just by increasing the collector current to the maximum value the device can take. Increasing I_c reduces voltage noise, but it increases current noise, as in Figure 1.5. There is an optimum collector current for each value of source resistance, where the contributions are equal. Because both voltage and current noise are proportional to the square root of I_c, they change slowly as it alters, and the combined noise curve is rather flat

www.raguvaran.puzl.com

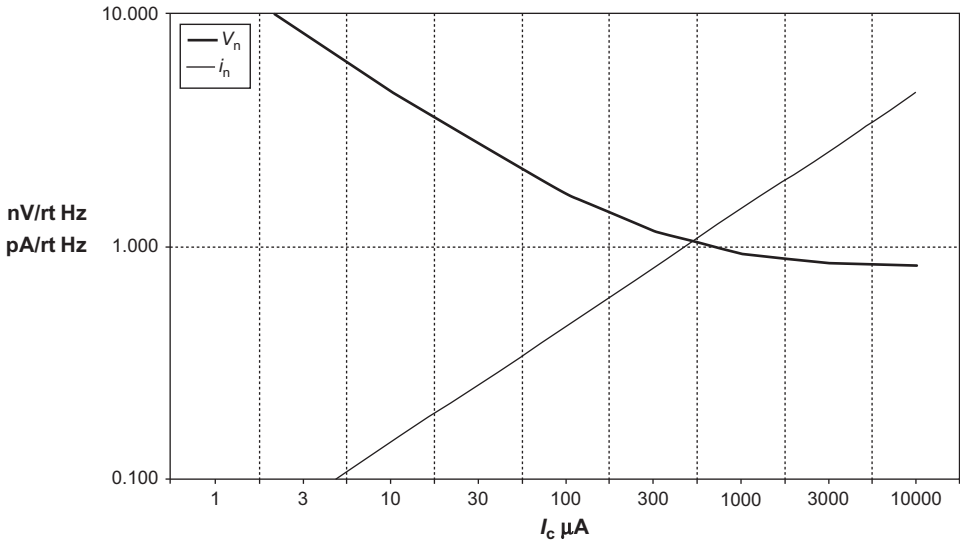


Figure 1.5: How voltage noise density v_n and current noise density i_n vary with collector current I_c in a generic transistor. As I_c increases, voltage noise asymptotes to a limit while current noise continuously increases

at the bottom. There is no need to control collector current with great accuracy to obtain optimum noise performance.

I must emphasize that this is a simplified noise model. In practice both voltage and current noise densities vary with frequency. I have also ignored 1/*f* noise. However, it gives the essential insight into what is happening and leads to the right design decisions so we will put our heads down and press on.

A quick example shows how this works. In a voltage amplifier we want the source impedances seen by the input transistors to be as low as possible, to minimize Johnson noise, and to minimize the effects of current noise. If we are lucky it may be as low as 100 Ω. How do we minimize the noise from a single transistor faced with a 100 Ω source resistance?

We assume the temperature is 25°C, the bandwidth is 22 kHz, and the r_{bb} of our transistor is 40 Ω. (Why don't they put this on spec sheets anymore?) The h_{fe} (beta) is 150. Set I_c to 1 mA, which is plausible for an amplifier input stage, step the source resistance from 1 to 100,000 Ω in decades, and we get Table 1.5.

Column 1 shows the source resistance and column 2 the Johnson noise density it generates by itself. Factor in the bandwidth actual noise voltage, and you get columns 3 and 4, which show the voltage in nV and dBu respectively. Column 5 is the noise density from the transistor, the rms sum of the voltage noise and the voltage generated by the current noise flowing in the source resistance. Column 6 gives total noise density when we sum the source resistance noise

www.raguvaran.puzl.com

TABLE 1.5 The summation of Johnson noise from the source resistance with transistor noise

1 R_s (Ω)	2 R_s Johnson (nV/√Hz)	3 R_s Johnson BW (nV)	4 R_s Johnson BW (dBu)	5 Transistor noise incl. in R_s (nV/√Hz)	6 Transistor noise plus R_s Johnson (nV/√Hz)	7 Noise in BW (nV)	8 Noise in BW (dBu)	9 Noise figure (dB)
1	0.128	19.0	−152.2	0.93	0.94	139.7	−134.9	17.3
10	0.406	60.2	−142.2	0.93	1.02	150.9	−134.2	8.0
100	1.283	190.3	−132.2	0.94	1.59	236.3	−130.3	1.9
1000	4.057	601.8	−122.2	1.73	4.41	654.4	−121.5	0.7
10,000	12.830	1903.0	−112.2	14.64	19.46	2886.9	−108.6	3.6
100,000	40.573	6017.9	−102.2	146.06	151.59	22484.8	−90.7	11.4

density with the transistor noise density. Factor in the bandwidth again, and the resultant noise voltage is given in columns 7 and 8. The final column 9 gives the noise figure (NF), which is the amount by which the combination of transistor and source resistance is noisier than the source resistance alone. In other words, it tells how close we have got to perfection, which would be a noise figure of 0 dB. The results for the 100 Ω source show that the transistor noise is less than the source resistance Johnson noise; there is little scope for improving things by changing transistor type or operating conditions.

The results for the other source resistances are worth looking at. The lowest noise output (−134.9 dBu) is achieved by the lowest source resistance of 1 Ω, as you would expect, but the NF is very poor at 17.3 dB, because the r_{bb} at 40 Ω is generating a lot more noise than the 1 Ω source. This gives you some idea why it is hard to design quiet moving-coil head amplifiers. The best noise figure, and the closest approach to theoretical perfection, is with a 1000 Ω source, attained with a *greater* noise output than 100 Ω. As source resistance increases further, NF worsens again; a transistor with $I_c = 1$ mA has relatively high current noise and performs poorly with high source resistances.

Since I_c is about the only thing we have control over here, let's try altering it. If we increase I_c to 3 mA we find that for 100 Ω source resistance, our amplifier is only a marginal 0.2 dB quieter (see Table 1.6, which skips the intermediate calculations and just gives the output noise and NF).

At 3 mA the noise with a 1 Ω source is 0.7 dB better, due to slightly lower voltage noise, but with 100 kΩ noise is higher by no less than 9.8 dB as the current noise is much increased.

If we increase I_c to 10 mA, this makes the 100 Ω noise worse again, and we have lost that slender 0.2 dB improvement.

At 1 Ω the noise is 0.3 dB better, which is not exactly a breakthrough, and for the higher source resistances things are worse again, the 100 kΩ noise increasing by another 5.2 dB. It therefore

www.raguvaran.puzl.com

TABLE 1.6 How input device collector current affects noise output and noise figure

	I_c = 3 mA		I_c = 10 mA		I_c = 10 mA, 2SB737		I_c = 100 μA	
R_s (Ω)	Noise (dBu)	NF (dB)	Noise (dBu)	NF (dB)	Noise (dBu)	NF (dB)	Noise (dBu)	NF (dB)
1	−135.6	16.6	−135.9	16.3	−145.9	6.3	−129.9	22.3
10	−134.8	7.4	−135.1	7.1	−140.9	1.3	−129.7	12.5
100	−130.5	1.7	−130.3	1.9	−131.5	0.7	−127.9	4.3
1000	−120.6	1.6	−118.5	3.7	−118.6	3.6	−121.5	0.7
10k	−105.3	6.9	−100.7	11.4	−100.7	11.4	−111.6	0.6
100k	−86.2	16.0	−81.0	21.2	−81.0	21.2	−98.6	3.6

appears that a collector current of 3 mA is actually pretty much optimal for noise with our 100 Ω source resistance.

If we now pluck out our 'ordinary' transistor and replace it with a specialized low-r_{bb} part like the much-lamented 2SB737, with its superbly low r_{bb} of 2 Ω, the noise output at 1 Ω plummets by 10 dB, showing just how important low r_{bb} is for moving-coil head amplifiers. The improvement for the 100 Ω source resistance is much less at 1.0 dB.

If we go back to the ordinary transistor and reduce I_c to 100 μA, we get the last two columns in Table 1.6. Compared with $I_c = 3$ mA, noise with the 1 Ω source worsens by 5.7 dB, and with the 100 Ω source by 2.6 dB, but with the 100 kΩ source there is a hefty 12.4 dB improvement, due to reduced current noise. Quiet BJT inputs for high impedances can be made by using low collector currents, but junction field-effect transistors (JFETs) usually give better noise performance under these conditions.

The transistor will probably be the major source of noise in the circuit, but other sources may need to be considered. The transistor may have a collector resistor of high value, to optimize the stage gain, and this naturally introduces its own Johnson noise. Most discrete transistor amplifiers have multiple stages, to get enough open-loop gain for linearization by negative feedback, and an important consideration in discrete noise design is that the gain of the first stage should be high enough to make the noise contribution of the second stage negligible. This can complicate matters considerably. Precisely the same situation prevails in an op-amp, but here someone else has done the worrying about second-stage noise for you; if you're not happy with it, all you can do is pick another type of op-amp.

Noise in JFETs

JFETs operate completely differently from bipolar transistors, and noise arises in different ways. The voltage noise in JFETs arises from the Johnson noise produced by the channel

www.raguvaran.puzl.com

resistance, the effective value of which is the inverse of the transconductance (g_m) of the JFET at the operating point we are looking at. An approximate but widely accepted equation for this noise is:

$$\text{Noise density } e_n = \sqrt{4kT\frac{2}{3g_m}} \text{ (in V}/\sqrt{\text{Hz}}\text{; usually nV}/\sqrt{\text{Hz}}) \qquad \text{(Equation 1.8)}$$

where k is Boltzmann's constant (1.380662×10^{-23}) and T is absolute temperature (°K).

FET transconductance goes up proportionally as the square root of drain current I_d. When the transconductance is inserted into the equation above, it is again square-rooted, so the voltage noise is proportional to the fourth root of drain current, and varies with it very slowly. There is thus little point in using high drain currents.

The only current noise source in a JFET is the shot noise associated with the gate leakage current. Because the leakage current is normally extremely low, the current noise is very low, which is why JFETs give a good noise performance with high source resistances. However, don't let the JFET get hot, because gate leakage doubles with each 10°C rise in temperature; this is why JFETs can actually show *increased* noise if the drain current is increased to the point where they heat up.

The g_m of JFETs is rather variable, but at $I_d = 1$ mA ranges over about 0.5–3 mA/V (or mMho) so the voltage noise density varies from 4.7 to 1.9 nV/$\sqrt{\text{Hz}}$. Comparing this with column 5 in Table 1.5, we can see that the BJTs are much quieter except at high source impedances, where their current noise makes them noisier than JFETs.

Noise in Op-Amps

The noise behavior of an op-amp is very similar to that of a single input amplifier, the difference being that there are now two inputs to consider, and usually more associated resistors.

An op-amp is driven by the voltage difference between its two inputs, and so the voltage noise can be treated as one voltage v_n connected between them (see Figure 1.6, which shows a differential amplifier).

Op-amp current noise is represented by two separate current generators i_n+ and i_n-, one in parallel with each input. These are assumed to be equal in amplitude and not correlated with each other. It is also assumed that the voltage and current noise sources are likewise uncorrelated, so that rms addition of their noise components is valid. In reality things are not quite so simple and there is some correlation, and the noise produced can be slightly higher than calculated. In practice the difference is small compared with natural variations in noise performance.

www.raguvaran.puzl.com

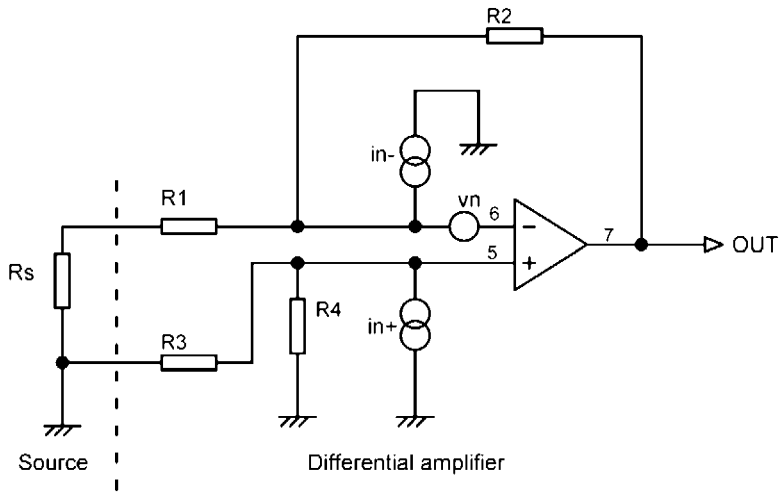


Figure 1.6: The noise sources in an op-amp differential amplifier circuit

Calculating the noise is somewhat more complex than for the simple amplifier of Figure 1.4. You must do the following:

1. Calculate the voltage noise from the voltage noise density.
2. Calculate the two extra noise voltages resulting from the noise currents flowing through their associated components.
3. Calculate the Johnson noise produced by each resistor.
4. Allow for the noise gain of the circuit when assessing how much each noise source contributes to the output.
5. Add the lot together by rms addition.

There is no space to go through a complete calculation, but here is a quick example. Suppose you have an inverting amplifier like that in Figure 1.9(a) below. This is simpler because the non-inverting input is grounded, so the effect of i_n+ disappears, as it has no resistance to flow through and cannot give rise to a noise voltage. This shunt-feedback stage has a 'noise gain' that is greater than the signal gain. The input signal is amplified by -1, but the voltage noise source in the op-amp is amplified by 2 times, because the voltage noise generator is amplified as if the circuit was a series-feedback gain stage.

Low-Noise Op-Amp Circuitry

The rest of this chapter deals with designing low-noise op-amp circuitry, dealing with op-amp selection and the minimization of circuit impedances. It also shows how adding more stages

www.raguvaran.puzl.com

can actually make the circuitry quieter. This sounds somewhat counter-intuitive, but as you will see, it is so.

When you are designing for low noise, it is obviously important to select the right op-amp, the great divide being between bipolar and JFET inputs. This chapter concentrates mainly on using the 5532, as it is not only a low-noise op-amp with superbly low distortion, but also a low-cost op-amp, due to its large production quantities. There are op-amps with lower noise, such as the AD797 and the LT1028, but these are specialized items and the cost penalties are high. The LT1028 has a bias-cancellation system that increases noise unless the impedances seen at each input are equal, and since audio does not need the resulting DC precision, it is not useful. The new LM4562 is a dual op-amp with somewhat lower noise than the 5532, but at present it also is much more expensive.

The AD797 runs its bipolar input transistors at high collector currents (about 1 mA), which reduces voltage noise but increases current noise. The AD797 will therefore only give lower noise for rather low source resistances; these need to be below 1 kΩ to yield benefit for the money spent. There is much more on op-amp selection in Chapters 4 and 14.

Noise Measurements

There are difficulties in measuring the low noise levels we are dealing with here. The Audio Precision System 1 test system has a noise floor of −116.4 dBu when its input is terminated with a 47 Ω resistor. When it is terminated in a short-circuit, the noise reading only drops to −117.0 dBu, demonstrating that almost all the noise is internal to the AP and the Johnson noise of the 47 Ω resistor is much lower. The significance of 47 Ω is that it is the lowest value of output resistor that will guarantee stability when driving the capacitance of a reasonable length of screened cable. This resistor value will keep cropping up in the noise situations we are about to examine.

To delve below this noise floor, we can subtract this figure from the noise we measure (on the usual rms basis) and estimate the noise actually coming from the circuit under test. This process is not very accurate when circuit noise is much below that of the test system, because of the subtraction involved, and any figure below −120 dBu should be regarded with caution. Cross-checking against the theoretical calculations and SPICE results is always wise; in this case it is essential.

We will now look at a number of common circuit scenarios and see how low-noise design can be applied to them.

How to Attenuate Quietly

Attenuating a signal by 6 dB sounds like the easiest electronic task in the world. Two equal-value resistors to make up a potential divider, and *voila*! This knotty problem is solved. Or is it?

www.raguvaran.puzl.com

To begin with, let us consider the signal going into our divider. Wherever it comes from, the source impedance is not likely to be less than 50 Ω. This is also the lowest output impedance setting for most high-quality signal generators (though it's 40 Ω on my AP SYS-2702). The Johnson noise from 50 Ω is −135.2 dBu, which immediately puts a limit – albeit a very low one – on the performance we can achieve. The maximum signal handling capability of op-amps is about +22 dBu, so we know at once our dynamic range cannot exceed 135 + 22 = 157 dB. This comfortably exceeds the dynamic range of human hearing, which is about 130 dB if you are happy to accept 'instantaneous ear damage' as the upper limit.

In the scenario we are examining, there is only one variable – the ohmic value of the two equal resistors. This cannot be too low, or the divider will load the previous stage excessively, increasing distortion and possibly reducing headroom. On the other hand, the higher the value, the greater the Johnson noise voltage generated by the divider resistances that will be added to the signal, and the greater the susceptibility of the circuit to capacitative crosstalk and general interference pickup. In Table 1.7 the trade-off is examined.

What happens when our signal with its −135.2 dBu noise level encounters our 6 dB attenuator? If it is made up of two 1 kΩ resistors, the noise level at once jumps up to −125.2 dBu, as the effective source resistance from two 1 kΩ resistors effectively in parallel is 500 Ω. Ten decibels of signal-to-noise ratio are irretrievably gone already, and we have only deployed two passive components. There will no doubt be more active and passive circuitry downstream, so things can only get worse.

However, a potential divider made from two 1 kΩ resistors in series presents an input impedance of only 2 kΩ, which is too low for most applications; 10 kΩ is normally considered the minimum input impedance for a piece of audio equipment in general use, which means we must use two 5 kΩ resistors, and so we get an effective source resistance of 2.5 kΩ. This produces Johnson noise at −118.2 dBu, so the signal-to-noise ratio has been degraded by another 7 dB simply by making the input impedance reasonably high.

TABLE 1.7 Johnson noise from 6 dB resistive divider with different resistor values (bandwidth 22 kHz, temperature 25°C)

Divider *R* value	Divider R_{eff}	Johnson noise (dBu)	Relative noise (dB)
100 Ω	50 Ω	−135.2	−27.0
500 Ω	250 Ω	−128.2	−20.0
1 kΩ	500 Ω	−125.2	−17.0
5 kΩ	2.5 kΩ	−118.2	−10.0
10 kΩ	5 kΩ	−115.2	−7.0
50 kΩ	25 kΩ	−108.2	0 reference
100 kΩ	50 kΩ	−105.2	+3.0

www.raguvaran.puzl.com

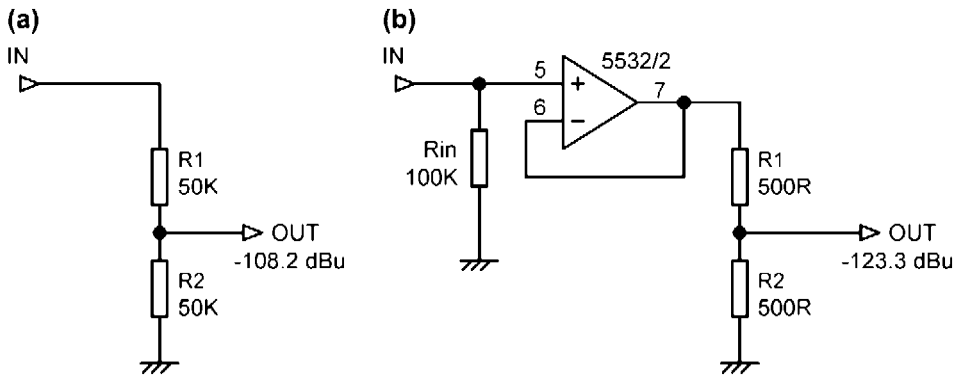


Figure 1.7: Two 6 dB attenuators with a 100 kΩ input impedance. (a) Simple attenuator with high resistor values. (b) Buffered attenuator with low resistor values. Despite the extra noise from the 5532 voltage-follower, this version is 15 dB quieter

In some cases 10 kΩ is not high enough, and a 100 kΩ input impedance is sought. Now the two resistors have to be 50 kΩ, and the noise is 10 dB higher again, at −108.2 dBu. That is a worrying 27 dB worse than our signal when it arrived.

If we insist on an input impedance of 100 kΩ, how can we improve on our noise level of −108.2 dBu? The answer is by buffering the divider from the outside world. The output noise of a 5532 voltage-follower is about −119 dBu with a 50 Ω input termination. If this is used to drive our attenuator, the two resistors in it can be as low as the op-amp can drive. The 5532 has a most convenient combination of low noise and good load-driving ability, and the divider resistors can be reduced to 500 Ω each, giving a load of 1 kΩ and a generous safety margin of drive capability (pushing the 5532 to its specified limit of a 500 Ω load tends to degrade its superb linearity by a small but measurable amount) (see Figure 1.7).

The noise from the resistive divider itself has now been lowered to −128.2 dBu, but there is of course the extra −119 dBu of noise from the voltage-follower that drives it. This, however, is halved by the divider just as the signal is, so the noise at the output will be the rms sum of −125 and −128.2 dBu, which is −123.3 dBu. A 6 dB attenuator is actually the worst case, as it has the highest possible source impedance for a given total divider resistance. Either more or less attenuation will mean less noise from the divider itself.

So, despite adding active circuitry that intrudes its own noise, the final noise level has been reduced from −108.2 to −123.3 dBu, an improvement of 15.1 dB.

How to Amplify Quietly

OK, we need a low-noise amplifier. Let's assume we have a reasonably low source impedance of 300 Ω, and we need a gain of four times (+12 dB). Figure 1.8(a) shows a very ordinary

www.raguvaran.puzl.com

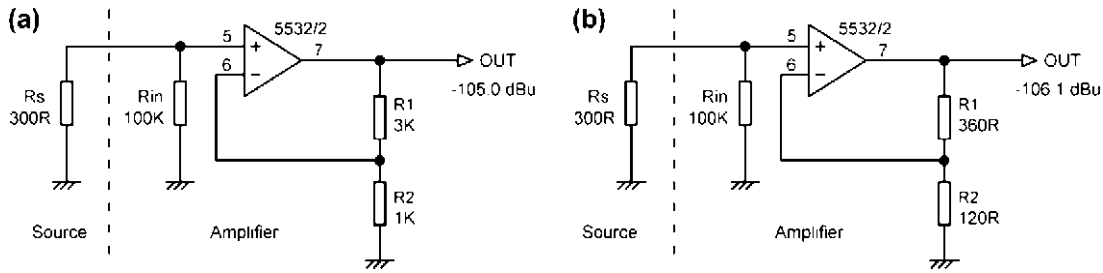


Figure 1.8: A 4× amplifier. (a) With 'normal' feedback resistances. (b) With low-impedance feedback arm resistances. Noise is only reduced by 1.1 dB

circuit using half a 5532 with typical values of 3 and 1 kΩ in the feedback network, and the noise output measures as −105.0 dBu. The Johnson noise generated by the 300 Ω source resistance is −127.4 dBu, and amplifying that by a gain of 4 gives −115.4 dBu. Compare this with the actual −105.0 dBu we get, and the noise figure is 10.4 dB – in other words, the noise from the amplifier is three times the inescapable noise from the source resistance, making the latter essentially negligible. This amplifier stage is clearly somewhat short of noise-free perfection, despite using one of the quieter op-amps around.

We need to make things quieter. The obvious thing to do is to reduce the value of the feedback resistances; this will reduce their Johnson noise and also reduce the noise produced in them by the op-amp current noise generators. Figure 1.8(b) shows the feedback network altered to 360 and 120 Ω, adding up to a load of 480 Ω, pushing the limits of the lowest resistance the op-amp can drive satisfactorily. This assumes of course that the next stage presents a relatively light load so that almost all of the driving capability can be used to drive the negative-feedback network; keeping tiny signals free from noise can involve throwing some serious current about. The noise output is reduced to −106.1 dBu, which is only an improvement of 1.1 dB, and only brings the noise figure down to 9.3 dB, leaving us still a long way from what is theoretically attainable. However, at least it cost us nothing in extra components.

If we need to make things quieter yet, what can be done? The feedback resistances cannot be reduced further, unless the op-amp drive capability is increased in some way. An output stage made of discrete transistors could be added, but it would almost certainly compromise the low distortion we get from a 5532 alone. For one answer see the next section on ultra-low-noise design.

How to Invert Quietly

Inverting a signal always requires the use of active electronics (OK, you *could* use a transformer). Assume that an input impedance of 47 kΩ is required, along with a unity-gain inversion. A straightforward inverting stage as shown in Figure 1.9(a) will give this input

www.raguvaran.puzl.com

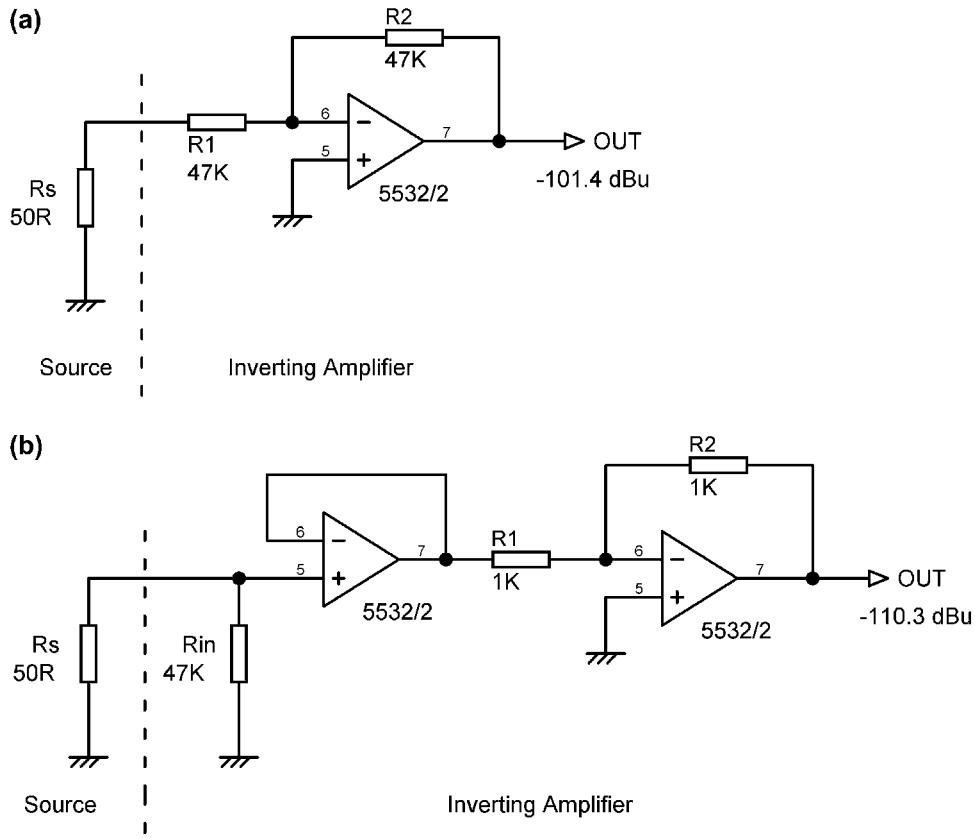


Figure 1.9: The noise from an inverter with 47 kΩ input impedance. (a) Unbuffered. (b) Buffered and with low-value resistors

impedance only if both resistors are 47 kΩ. These relatively high-value resistors contribute to a noise output of −101.4 dBu, which is also due to the fact that the op-amp is working at a noise gain of 2 times.

The only way to improve this noise level is to add another active stage. It sounds paradoxical – adding more non-silent circuitry to reduce noise – but that's the way the universe works. If a voltage-follower is added to the circuit given in Figure 1.9(b), then the resistors around the inverting op-amp can be greatly reduced in value without reducing the input impedance, which can now be pretty much as high as we like. The 'Noise buffered' column in Table 1.8 shows that if *R* is reduced to 2.2 kΩ the total noise output is lowered by 8.2 dB, which is a very useful improvement. If *R* is further reduced to 1 kΩ, which is perfectly practical with a 5532's drive capability, the total noise is reduced by 9.0 dB compared with the 47 kΩ case. The 'Noise unbuffered' column gives the noise output with specified *R* value but without the buffer, demonstrating that adding the buffer does degrade

www.raguvaran.puzl.com

TABLE 1.8 Measured noise from simple inverter and buffered inverter (5532)

R value (Ω)	Noise unbuffered (dBu)	Noise buffered (dBu)	Noise reduction (dB) with ref to 47k case
1 k	−111.0	−110.3	9.0
2k2	−110.1	−109.5	8.2
4k7	−108.9	−108.4	7.1
10k	−106.9	−106.6	5.3
22k	−104.3	−104.3	3.0
47k	−101.4	−101.3	0 reference

the noise slightly, but the overall result is still far quieter than the unbuffered version with 47 kΩ resistors. In each case the circuit input is terminated to ground via 50 Ω.

How to Balance Quietly

The design of low- and ultra-low-noise balanced amplifiers is thoroughly examined in Chapter 14.

Ultra-Low-Noise Design with Multipath Amplifiers

Are the above circuit structures the ultimate? Is this as low as noise gets? No. In the search for low noise, a powerful technique is the use of parallel amplifiers with their outputs summed. This is especially useful where source impedances are low and therefore generate little noise compared with the associated electronics.

If there are two amplifiers connected, the signal gain increases by 6 dB due to the summation. The noise from the two amplifiers is also summed, but since the two noise sources are completely uncorrelated (coming from physically different components) they partially cancel and the noise level only increases by 3 dB. Thus there is an improvement in signal-to-noise ratio of 3 dB. This strategy can be repeated by using four amplifiers, in which case the signal-to-noise improvement is 6 dB. Table 1.9 shows how this works for increasing numbers of amplifiers.

In practice the increased signal gain is not useful, and an active summing amplifier would compromise the noise improvement, so the output signals are averaged rather than summed, as shown in Figure 1.10. The amplifier outputs are simply connected together with low-value resistors; the gain is unchanged but the noise output falls. The amplifier outputs are nominally identical, so very little current should flow from one op-amp to another. The combining resistor values are so low that their Johnson noise can be ignored.

www.raguvaran.puzl.com

TABLE 1.9 How noise performance improves with multiple amplifiers

No. of amplifiers	Noise reduction (dB)
1	0 reference
2	−3.01
3	−4.77
4	−6.02
5	−6.99
6	−7.78
7	−8.45
8	−9.03
12	−10.79
16	−12.04
256	−24.58

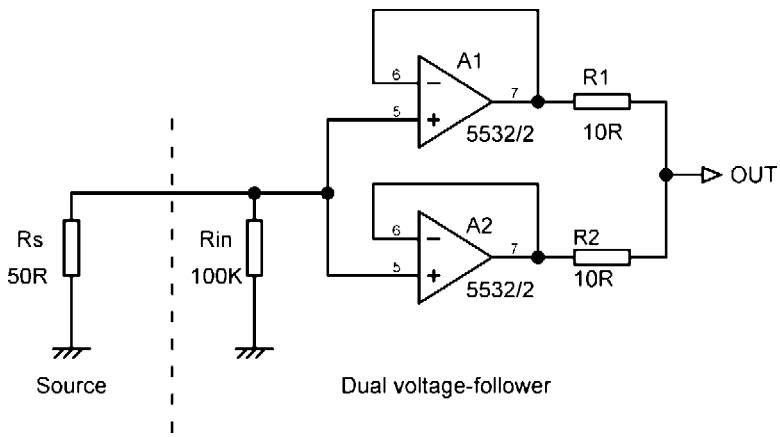


Figure 1.10: A double voltage-follower or buffer. The noise of this simple circuit is below that of the leading test equipment available. Noise output is approximately -123 dBu

Obviously there are economic limits on how far you can take this sort of thing. Unless you're measuring gravity waves or something equally important, 256 parallel amplifiers is probably not a viable choice.

Ultra-Low-Noise Voltage Buffers

The multiple-path philosophy works well even with a minimally simple circuit such as a unity-gain voltage buffer. Table 1.10 gives calculated results for 5532 sections (the noise output is

www.raguvaran.puzl.com

TABLE 1.10 Noise from parallel-array buffers using 5532 sections

No. of op-amps	Calculated noise out (dBu)
1	−120.4
2	−123.4
3	−125.2
4	−126.4

too low to measure reliably even with the best testgear) and shows how the noise output falls as more op-amps are added. The distortion performance is not affected.

The 10 Ω output resistors combine the op-amp outputs, and limit the currents that would flow from output to output as a result of DC offset errors. AC gain errors here will be very small indeed as the op-amps have 100% feedback. If the output resistors were raised to 47 Ω they would as usual give HF stability when driving screened cables or other capacitances, but the total output impedance is usefully halved to 23.5 Ω. Another interesting bonus of this technique is that we have doubled the output drive capability; this stage can easily drive 300 Ω.

Ultra-Low-Noise Amplifiers

We now return to the problem studied earlier; how to make a really quiet amplifier with a gain of 4 times. We saw that the minimum noise output using a single 5532 section and a 300 Ω source resistance was −106.1 dBu, with a not particularly impressive noise figure of 9.3 dB. Since almost all the noise is being generated in the amplifier rather than the source resistance, the multiple-path technique should work well here. And it does.

There is, however, a potential snag that needs to be considered. In the previous section, we were combining the outputs of voltage-followers, which have gains very close indeed to unity because they have 100% negative feedback and no resistors are involved in setting the gain. We could be confident that the output signals would be near-identical and unwanted currents flowing from one op-amp to the other would be small despite the low value of the combining resistors.

The situation here is different; the amplifiers have a gain of 4 times, so there is a smaller negative feedback factor to stabilize the gain, and there are two resistors with tolerances that set the closed-loop gain for each stage. We need to keep the combining resistors low to minimize their Johnson noise, so things might get awkward. It seems reasonable to assume that the feedback resistors will be 1% components. Considering the two-amplifier configuration in Figure 1.11, the worst case would be to have R1A 1% high and R2A 1% low in one amplifier, while the other had the opposite condition of R1B 1% low and R2B 1% high. This highly

www.raguvaran.puzl.com

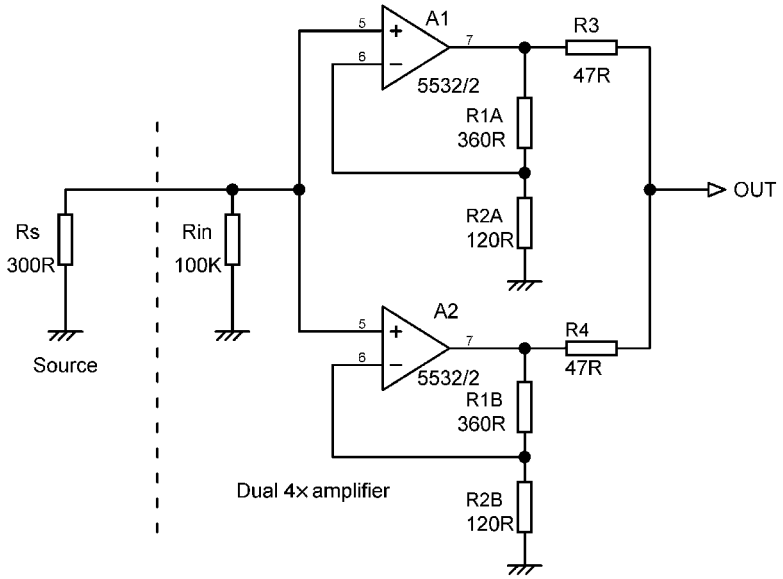


Figure 1.11: A 4× amplifier using two op-amps to reduce noise by approaching 3 dB

TABLE 1.11 Noise from multiple amplifiers with 4× gain

R_s (Ω)	No. of op-amps	Noise out (dBu)	Improvement (dB)
300	1	−106.1	0 reference
300	2	−108.2	2.2
300	3	−109.0	2.9
300	4	−109.6	3.5
200	1	−106.2	0 reference
200	2	−108.4	2.2
200	3	−109.3	3.1
200	4	−110.0	3.8
100	1	−106.3	0 reference
100	2	−108.7	2.4
100	3	−109.8	3.5
100	4	−110.4	3.9

www.raguvaran.puzl.com

unlikely state of affairs gives a gain of 4.06 times in the first amplifier and 3.94 times in the second. Making the further assumption of a 10 Vrms maximum voltage swing, we get 10.15 Vrms at the first output and 9.85 Vrms at the second, both applied to the combining resistors, which here are set at 47 Ω. The maximum possible current flowing from one amplifier output into the other is therefore 0.3 V/(47 Ω + 47 Ω), which is 3.2 mA; in practice it will be much smaller. There are no problems with linearity or headroom, and distortion performance is indistinguishable from that of a single op-amp.

Having reassured ourselves on this point, we can examine the circuit of Figure 1.11, with two amplifiers combining their outputs. This reduces the noise at the output by 2.2 dB. This falls short of the 3 dB improvement we might hope for because of a significant Johnson noise contribution from source resistance, and doubling the number of amplifier stages again only achieves another 1.3 dB improvement. The improvement is greater with lower source resistances; the measured results with one, two, three, and four op-amps for three different source resistances are summarized in Table 1.11.

The results for 200 and 100 Ω show that the improvement with multiple amplifiers is greater for lower source resistances, as these resistances generate less Johnson noise of their own.

References

[1] J. Smith, Modern Operational Circuit Design, Wiley-Interscience, 1971, p. 129.
[2] Intersil Application NoteAN9420.1, Current Feedback Amplifier Theory and Applications (April 1995).
[3] D. Self, Audio Power Amplifier Design Handbook, fifth ed. Focal Press, 2009 (Chapter 2).
[4] J.I. Yellott Jr., Spectral consequences of photoreceptor sampling in the rhesus retina, Science 221 (1983), pp. 382–385.
[5] J. Johnson, Thermal agitation of electricity in conductors, Phys. Rev. 32 (1928) 97.
[6] Sansen Chang, Low-Noise Wideband Amplifiers in Bipolar and CMOS Technologies, Kluwer, 1991, p. 106.

www.raguvaran.puzl.com

CHAPTER 2

Components

Conductors

It is easy to assume, when wrestling with electronic design, that the active devices will cause most of the trouble. This, like so much in electronics, is subject to Gershwin's Law: it ain't necessarily so. Passive components cannot be assumed to be perfect, and while their shortcomings are rarely discussed, they are all too real. In this chapter I have tried to avoid repeating basic stuff that can be found in many places, to allow room for information that goes deeper.

Normal metallic conductors, such as copper wire, show perfect linearity for our purposes, and as far as I am aware, for everybody's purposes. Ohm's Law was founded on metallic conductors after all, not resistors, which did not exist as we know them at that time. Georg Simon Ohm published a pamphlet in 1827 entitled 'The Galvanic Circuit Investigated Mathematically' while he was a professor of mathematics in Cologne. His work was not warmly received, except by a perceptive few; the Prussian minister of education pronounced that 'a professor who preached such heresies was unworthy to teach science'. This is the sort of thing that happens when politicians try to involve themselves in science, and in that respect we have progressed little since then.

Although the linearity is generally effectively ideal, metallic conductors will not be perfectly linear in some circumstances. Poorly made connections between oxidized or otherwise contaminated metal parts are capable of generating harmonic distortion at the level of several percent, but this is a property of the contact interface rather than the bulk material, and usually means that the connection is about to fail altogether. A more subtle danger is that of magnetic conductors – the soft iron in relay frames causes easily detectable distortion at power-amplifier current levels.

From time to time some of the dimmer audio commentators speculate that metallic conductors are actually a kind of 'sea of micro-diodes', and that non-linearity can be found if the test signal levels are made small enough. This is both categorically untrue and physically impossible. There is no threshold effect for metallic conduction.

I have myself added to the mountain of evidence on this, by measuring distortion at very low signal levels [1].

Small Signal Audio Design; ISBN: 9780240521770
Copyright © 2010 Elsevier Ltd; All rights of reproduction, in any form, reserved.

www.raguvaran.puzl.com

Copper and Other Conductive Elements

Copper is the preferred metal for conducting electricity in almost all circumstances. It has the lowest resistance of any metal but silver, is reasonably resistant to corrosion, and can be made mechanically strong. Being a heavy metal, it is unfortunately not that common in the earth's crust, and so is expensive compared with iron and steel. It is, however, cheap compared with silver. The price of both metals varies all the time due to changing economic and political factors, but at the time of writing silver was 100 times more expensive by weight. Given the same cross-section of conductor, the use of silver would only reduce the resistance of a circuit by 5%. Despite this, silver connection wire has been used in some very expensive hi-fi amplifiers; output impedance-matching transformers wound with silver wire are not unknown in valve amplifiers. Since the technical advantages are negligible such equipment is marketed on the basis of indefinable subjective improvements.

Table 2.1 gives the resistivity of the commonly used conductors, plus some insulators to give it perspective. The difference between copper and quartz is of the order of 10^{25}, an enormous range that is not found in many other physical properties.

TABLE 2.1 Properties of conductors and non-conductors

Material	Resistivity ρ (Ω-m)	Temperature coefficient per °C	Electrical usage
Silver	1.59×10^{-8}	0.0061	Conductors
Copper	1.72×10^{-8}	0.0068	Conductors
Gold	2.2×10^{-8}	0.0041	Inert coatings
Aluminum	2.65×10^{-8}	0.00429	Conductors
Tungsten	5.6×10^{-8}	0.0045	Lamp filaments
Iron	9.71×10^{-8}	0.00651	Barreters*
Platinum	10.6×10^{-8}	0.003927	Electrodes
Lead	22×10^{-8}		Storage batteries
Manganin (Cu, Mn, Ni)†	48.2×10^{-8}	0.000002	Resistances
Constantan (Cu, Ni)†	$49\text{–}52 \times 10^{-8}$	± 0.00002	Resistances
Mercury	98×10^{-8}	0.0009	Relays
Nichrome (Ni, Fe, Cr alloy)	100×10^{-8}	0.0004	Heating elements
Carbon (as graphite)	$3\text{–}60 \times 10^{-5}$	-0.0005	Brushes
Glass	$1\text{–}10{,}000 \times 10^{9}$		Insulators
Fused quartz	7.5×10^{8}		Insulators

*A barreter is an incredibly obsolete device consisting of thin iron wire in an evacuated glass envelope. It was typically used for current regulation of the heaters of RF oscillator valves, to improve frequency stability.

†Constantan and manganin are resistance alloys with moderate resistivity and a low temperature coefficient. Constantan is preferred as it has a flatter resistance–temperature curve and its corrosion resistance is better.

www.raguvaran.puzl.com

There are several reasonably conductive metals that are lighter than copper, but their higher resistivity means they require a larger cross-section to carry the same current, so copper is always used when space is limited, as in electric motors, solenoids, etc. However, when size is not the primary constraint, the economics work out differently. The largest use of non-copper conductors is probably in the transmission line cables that are strung between pylons. Here minimal weight is more important than minimal diameter and the cables have a central steel core for strength, surrounded by aluminum conductors.

It is clear that simply spending more money does not automatically bring you a better conductor; gold is a somewhat poorer conductor than copper, and platinum, which is even more expensive, is worse by a factor of 6. Another interesting feature of this table is the relatively high resistance of mercury, nearly 60 times that of copper. This often comes as a surprise; people seem to assume that a metal of such high density must be very conductive, but it is not so. There are many reasons for not using mercury-filled hoses as loudspeaker cables, and their conductive inefficiency is just one. The cost and the insidiously poisonous nature of the metal are two more. Nonetheless, it is reported that the Hitachi Cable company has experimented with speaker cables made from polythene tubes filled with mercury. There appear to have been no plans to put such a product on the market. RoHS compliance might be a problem.

The Metallurgy of Copper

Copper is a good conductor because the outermost electrons of its atoms have a large mean free path between collisions. The electrical resistivity of a metal is inversely related to this electron mean free path, which in the case of copper is approximately 100 atomic spacings.

Copper is normally used as a very dilute alloy known as electrolytic tough pitch (ETP) copper, which consists of very high purity metal alloyed with oxygen in the range of 100–650 ppm. In view of the wide exposure that the concept of oxygen-free copper has had in the audio business, it is worth underlining that the oxygen is deliberately alloyed with the copper to act as a scavenger for dissolved hydrogen and sulfur, which become water and sulfur dioxide. Microscopic bubbles form in the mass of metal but are completely eliminated during hot rolling.

The main use of oxygen-free copper is in conductors exposed to a hydrogen atmosphere at high temperatures. ETP copper is susceptible to hydrogen embrittlement in these circumstances, which arise in the hydrogen-cooled alternators in power stations.

Gold and Its Uses

As stated above, gold has a higher resistivity than copper, and there is no incentive to use it as the bulk metal of conductors, not least because of its high cost. However, it is very useful as a thin coating on contacts because it is almost immune to corrosion, though it is chemically

www.raguvaran.puzl.com

attacked by fluorine and chlorine. (If there is a significant amount of either gas in the air then your medical problems will be more pressing than your electrical ones.) Other electrical components are sometimes gold-plated simply because the appearance is attractive. A carat (or karat) is a $\frac{1}{24}$ part, so 24-carat gold is the pure element, while 18-carat gold contains only 75% of the pure metal. Eighteen-carat gold is the sort usually used for jewellery as it retains the chemical inertness of pure gold but is much harder and more durable; the usual alloying elements are copper and silver.

There are a couple of issues to bear in mind when dealing with gold in electronics; think before you solder gold-plated parts. Gold forms brittle intermetallics with the components of tin–lead solder, giving a weak high-resistance joint that is likely to cause horrible problems in high-current circuits. I have not so far been able to track down any reliable data on interaction between gold and lead-free solder, but I wouldn't risk it if I were you.

Eighteen-carat gold is widely used in jewellery and does not tarnish, so it is initially puzzling to find that some electronic parts plated with it have a protective transparent coating which the manufacturer claims to be essential to prevent blackening. The answer is that if gold is plated directly on to copper, the copper diffuses through the gold and tarnishes on its surface. The standard way of preventing this is to plate a layer of nickel on to the copper to prevent diffusion, then plate on the gold. I have examined some transparent-coated gold-plated parts and found no nickel layer; presumably the manufacturer finds the transparent coating is cheaper than another plating process to deposit the nickel.

Cable and Wiring Resistance

Electrical cable is very often specified by its cross-sectional area and current-carrying capacity, and the resistance per meter is seldom quoted. This can, however, be a very important parameter for assessing permissible voltage drops and for predicting the crosstalk that will be introduced between two signals when they share a common ground conductor. Given the resistivity of copper from Table 2.1, the resistance R of L meters of cable is simply:

$$R = \frac{\text{Resistivity} \cdot L}{\text{Area}} \qquad \text{(Equation 2.1)}$$

Note that the area, which is usually quoted in catalogs in square millimeters, must be expressed here in square meters to match up with the units of resistivity and length. Thus, 5 meters of cable with a cross-sectional area of 1.5 mm^2 will have a resistance of:

$$(1.72 \times 10^{-8}) \times 5/(0.0000015) = \ 0.057 \text{ ohms}$$

This gives the resistance of our stretch of cable, and it is then simple to treat this as part of a potential divider to calculate the voltage drop down its length.

www.raguvaran.puzl.com

PCB Track Resistance

It is also useful to be able to calculate the resistance of a printed-circuit board (PCB) track for the same reasons. This is slightly less straightforward to do; given the smorgasbord of units that are in use in PCB technology, determining the cross-sectional area of the track can present some difficulty.

In the USA and the UK, and probably elsewhere, there is inevitably a mix of metric and imperial units on PCBs, as many important components come in dual-in-line packages that are derived from an inch grid; track widths and lengths are therefore very often in thousandths of an inch, universally (in the UK at least) referred to as 'thou'. Conversely, the PCB dimensions and fixing-hole locations will almost certainly be metric as they interface with a world of metal fabrication and mechanical CAD that (once again, in the UK at least) went metric many years ago. Add to this the UK practice of quoting copper thickness in ounces (the weight of a square foot of copper foil) and all the ingredients for dimensional confusion are in place.

The most common type of PCB laminate is called 'one-ounce copper', which means a one-ounce weight of copper foil over a square foot. Less often employed, but still easily obtained, is 'two-ounce copper', which naturally has twice the thickness of copper cladding; it is somewhat more expensive. It is most often employed to increase track current-carrying capacity in power amplifiers, switch mode power supplies, and the like. 'Four-ounce copper' can be had but it is rarely used and is correspondingly costly.

Given the copper thickness, multiplying by track width and length gives the cross-sectional area. Since resistivity is always in metric units, it is best to convert to metric at this point, so Table 2.2 gives area in square millimeters. This is then multiplied by the resistivity, not forgetting to convert the area to meters for consistency. This gives the 'resistance' column in the table, and it is then simple to treat this as part of a potential divider to calculate the usually unwanted voltage across the track.

TABLE 2.2 Thickness of copper cladding and the calculation of track resistance

Weight (oz)	Thickness (thou)	Thickness (micron)	Width (thou)	Length (inch)	Area (mm^2)	Resistance (Ω)	Atten. ref. 8 Ω (dB)
1	1.38	35	12	3	0.0107	0.123	−36.4
1	1.38	35	50	3	0.0444	0.029	−48.7
2	2.76	70	12	3	0.0213	0.061	−42.4
2	2.76	70	50	3	0.0889	0.015	−54.7
4	5.52	140	50	3	0.178	0.0074	−60.7

www.raguvaran.puzl.com

For example, if the track in question is the ground return from an 8 Ω speaker load, this is the top half of a potential divider while the track is the bottom half (I am of course ignoring here the fact loudspeakers are not purely resistive loads), and a quick calculation gives the fraction of the input voltage found along the track. This is expressed in the last column of Table 2.2 as attenuation in dB. This shows clearly that loudspeaker outputs should not have common return tracks or the interchannel crosstalk will be dire.

It is very clear from this table that relying on thicker copper on your PCB as a means of reducing path resistance is not very effective. In some situations it may be the only recourse, but in many cases a path of much lower resistance can be made by using 32/02 cable soldered between the two relevant points on the PCB.

PCB tracks have a limited current capability because excessive resistive heating will break down the adhesive holding the copper to the board substrate, and ultimately melt the copper. This is normally only a problem in power amplifiers and power supplies. It is useful to assess if you are likely to have problems before committing to a PCB design, and Table 2.3, based on MIL-standard 275, gives some guidance.

Note that Table 2.3 applies to tracks on the PCB surface only. Internal tracks in a multilayer PCB experience much less cooling, and need to be about three times as thick for the same temperature rise. This factor depends on laminate thickness and so on, and you need to consult your PCB vendor.

Traditionally, overheated tracks could be detected visually because the solder mask on top of them would discolor to brown. I am not sure if this still applies with modern solder mask materials, as in recent years I have been quite successful in avoiding overheated tracking.

TABLE 2.3 PCB track current capacity in Amps for a permitted temperature rise

Track temp. rise (°C):	10		20		30	
Copper weight (oz):	1	2	1	2	1	2
Track width (thou)						
10	1.0	1.4	1.2	1.6	1.5	2.2
15	1.2	1.6	1.3	2.4	1.6	3.0
20	1.3	2.1	1.7	3.0	2.4	3.6
25	1.7	2.5	2.2	3.3	2.8	4.0
30	1.9	3.0	2.5	4.0	3.2	5.0
50	2.6	4.0	3.6	6.0	4.4	7.3
75	3.5	5.7	4.5	7.8	6.0	10.0
100	4.2	6.9	6.0	9.9	7.5	12.5
200	7.0	11.5	10.0	16.0	13.0	20.5
250	8.3	12.3	12.3	20.0	15.0	24.5

www.raguvaran.puzl.com

PCB Track-to-Track Crosstalk

The previous section described how to evaluate the amount of crosstalk that can arise because of shared track resistances. Another crosstalk mechanism is caused by capacitance between PCB tracks. This is not very susceptible to calculation, so I did the following experiment to put some figures to the problem.

Figure 2.1 shows the set-up; four parallel conductors 1.9 inches long on a standard piece of 0.1-inch pitch prototype board were used as test tracks. These are perhaps rather wider than the average PCB track, but one must start somewhere. The test signal was applied to track A, and track C was connected to a virtual-earth summing amplifier A1.

The tracks B and D were initially left floating. The results are shown as trace 1 in Figure 2.2; the coupling at 10 kHz is −65 dB, which is worryingly high for two tracks 0.2 inch apart. Note that the crosstalk increases steadily at 6 dB per octave, as it results from a very small capacitance driving into what is effectively a short-circuit.

It has often been said that running a grounded screening track between two tracks that are susceptible to crosstalk has a beneficial effect, but how much good does it really do? Grounding track B, to place a screen between A and C, gives trace 2 and has only improved matters by 9 dB – not the dramatic effect that might be expected from screening. The reason, of course, is that electric fields are very much three-dimensional, and if you could see the electrostatic 'lines of force' that appear in physics textbooks you would notice they arch up and over any planar screening such as a grounded track. It is easy to forget this when staring at

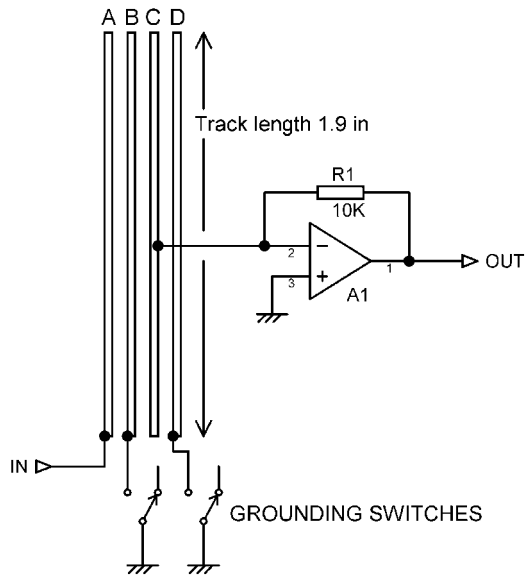


Figure 2.1: Test circuit for measuring track-to-track crosstalk on a PCB

www.raguvaran.puzl.com

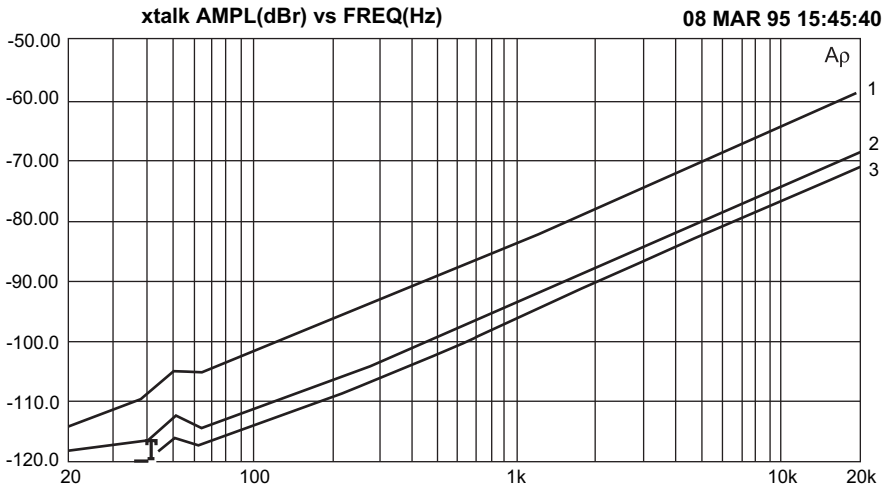


Figure 2.2: Results of PCB track-to-track crosstalk tests

a CAD display. There are of course two-layer and multilayer PCBs, but the visual effect on a screen is still of several slices of 2-D. As Mr Spock remarked in one of the Star Trek films: 'He's intelligent, but not experienced. His pattern indicates two-dimensional thinking.'

Grounding track D, beyond receiving track C, gives a further improvement of about 3 dB; this would clearly not happen if PCB crosstalk was simply a line-of-sight phenomenon.

To get more effective screening than this you must go into three dimensions too; with a double-sided PCB you can put one track on each side, with ground plane opposite. With a four-layer board it should be possible to sandwich critical tracks between two layers of ground plane, where they should be safe from pretty much anything. If you can't do this and things are really tough you may need to resort to a screened cable between two points on the PCB; this is of course expensive in assembly time. If components, such as electrolytics with their large surface area, are talking to each other you may need to use a vertical metal wall, but this costs money. A more cunning plan is to use electrolytics not carrying signal, such as rail decouplers, as screening items.

The internal crosstalk between the two halves of a dual op-amp is very low according to the manufacturer's specs. Nevertheless, avoid having different channels going through the same op-amp if you can because this will bring the surrounding components into close proximity, and will permit capacitative crosstalk.

Impedances and Crosstalk: A Case History

Capacitative crosstalk between two op-amp outputs can be surprisingly troublesome. The usual isolating resistor on an op-amp output is 47 Ω, and you might think that this impedance is

www.raguvaran.puzl.com

so low that the capacitative crosstalk between two of these outputs would be completely negligible, but you would be wrong.

A stereo power amplifier had balanced input amplifiers with 47 Ω output isolating resistors included to prevent any possibility of instability, although the op-amps were driving only a few centimeters of PCB track rather than screened cables with their significant capacitance. Just downstream of these op-amps was a switch to enable biamping by driving both left and right outputs with the left input. This switch and its associated tracking brought the left and right signals into close proximity, and the capacity between them was not negligible.

Crosstalk at low frequencies (below 1 kHz) was pleasingly low, being better than −129 dB up to 70 Hz, which was the difference between the noise floor and the maximum signal level. (The measured noise floor was unusually low at −114 dBu because each input amplifier was a quadruple noise-canceling type as described in Chapter 14, and that figure includes the noise from an AP System 1.) At higher frequencies things were rather less gratifying, being −96 dB at 10 kHz, as shown by the '47R' trace in Figure 2.3. In many applications this would be more than acceptable, but in this case the highest performance possible was being sought.

It was therefore decided to reduce the output isolating resistors to 10 Ω, so the inter-channel capacitance would have less effect. (Checks were done at the time and all through the prototyping and pre-production process to make sure that this would be enough resistance to ensure op-amp stability – it was.) This handily reduced the crosstalk to −109 dB at 10 kHz, an improvement of 13 dB at zero cost. This is the ratio between the two resistor values.

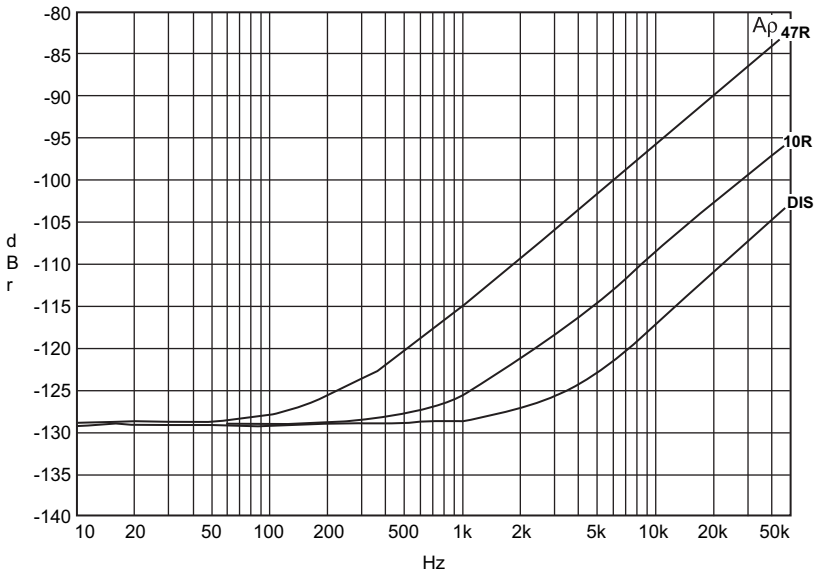


Figure 2.3: Crosstalk between op-amp outputs with 47 Ω and 10 Ω output isolating resistors

www.raguvaran.puzl.com

The third trace, marked 'DIS', shows the result of removing the isolating resistor from the speaking channel, so no signal reached the biamping switch. As usual, this reveals a further crosstalk mechanism, at about −117 dB, for reducing crosstalk is proverbially like peeling onions. There is layer after layer, and even strong men are reduced to tears.

Resistors

In the past there have been many types of resistor, including some interesting ones consisting of jars of liquid, but only a few kinds are likely to be met with now. These are usually classified by the kind of material used in the resistive element, as this has the most important influence on the fine details of performance. The major materials and types are shown in Table 2.4.

These values are illustrative only, and it would be easy to find exceptions. As always, the official data sheet for the component you have chosen is the essential reference. The sinister significance of the voltage coefficient is explained below.

It should be said that you are most unlikely to come across carbon composition resistors in modern signal circuitry, but they frequently appear in vintage valve equipment so they are included here. They also live on in specialized applications such as switch-mode snubbing circuits, where their ability to absorb a high peak power in a mass of material rather than a thin film is very useful.

Carbon-film resistors are currently still sometimes used in low-end consumer equipment, but elsewhere have been supplanted by the other three types. Note from Table 2.4 that they have a significant voltage coefficient.

Metal-film resistors are now the usual choice when any degree of precision or stability is required. These have no non-linearity problems at normal signal levels. The voltage coefficient is usually negligible.

Metal-oxide resistors are more problematic. Cermet resistors and resistor packages are metal oxide, and are made of the same material as thick-film surface-mount resistors. Thick-film

TABLE 2.4 Characteristics of resistor types

Type	Resistance tolerance (%)	Temperature coefficient	Voltage coefficient (ppm)
Carbon composition	±10	+400 to −900 ppm/°C	350
Carbon film	±5	−100 to −700 ppm/°C	100
Metal film	±1	+100 ppm/°C	1
Metal oxide	±5	+300 ppm/°C	Variable but too high
Wirewound	±5	±70% to ±250%	1

www.raguvaran.puzl.com

resistors can show significant non-linearity at op-amp-type signal levels, and should be kept out of high-quality signal paths.

Wirewound resistors are indispensable when serious power needs to be handled. The average wirewound resistor can withstand very large amounts of pulse power for short periods, but component manufacturers are often very reluctant to publish specifications on this capability, and endurance tests have to be done at the design stage. The voltage coefficient is usually negligible.

Resistors for general PCB use come in both through-hole and surface-mount types. Through-hole (TH) resistors can be any of the types tabled above; surface-mount (SM) resistors are always either metal film or metal oxide. There are also many specialized types; for example, high-power wirewound resistors are often constructed inside a metal case that can be bolted down to a heat-sink.

Through-Hole Resistors

These are too familiar to require much description. They are available in all the materials mentioned above: carbon film, metal film, metal oxide, and wirewound. There are a few other sorts, such as metal foil, but they are restricted to specialized applications. Conventional through-hole resistors are now almost always 250 mW 1% metal film. Carbon film used to be the standard resistor material, with the expensive metal-film resistors reserved for critical places in circuitry where low temperature coefficient and an absence of excess noise were really important, but as metal film got cheaper so it took over many applications.

TH resistors have the advantage that their power and voltage ratings greatly exceed those of surface-mount versions. They also have a very low voltage coefficient, which for our purposes is of the first importance. On the downside, the spiral construction of the resistance element means they have much greater parasitic inductance.

Surface-Mount Resistors

Surface-mount resistors come in two main formats, the common chip type and the rarer (and much more expensive) MELF format.

Chip surface-mount (SM) resistors come in a flat tombstone format, which varies over a wide size range (see Table 2.5).

MELF surface-mount resistors have a cylindrical body with metal endcaps, the resistive element is metal film, and the linearity is therefore as that of good as that of conventional resistors, with a voltage coefficient of less than 1 ppm. MELF is apparently an acronym for ‘Metal ELectrode Face-bonded’, though most people I know call them ‘Metal Ended Little Fellows’ or something quite close to that.

www.raguvaran.puzl.com

TABLE 2.5 The standard surface-mount resistor sizes with typical ratings

Size $L \times W$	Max. power dissipation	Max. voltage (V)
2512	1 W	200
1812	750 mW	200
1206	250 mW	200
0805	125 mW	150
0603	100 mW	75
0402	100 mW	50
0201	50 mW	25
01005	30 mW	15

Surface-mount resistors may have thin-film or thick-film resistive elements. The latter are cheaper and so more often encountered, but the price differential has been falling in recent years. Both thin-film and thick-film SM resistors use laser trimming to make fine adjustments of resistance value during the manufacturing process. There are important differences in their behavior.

Thin-film (metal-film) SM resistors use a nickel–chromium (Ni–Cr) film as the resistance material. A very thin Ni–Cr film of less than 1 μm thickness is deposited on the aluminum oxide substrate by sputtering under vacuum. Ni–Cr is then applied on to the substrate as conducting electrodes. The use of a metal film as the resistance material allows thin-film resistors to provide a very low temperature coefficient, much lower current noise and vanishingly small non-linearity. Thin-film resistors need only low laser power for trimming (one-third of that required for thick-film resistors) and contain no glass-based material. This prevents possible microcracking during laser trimming and maintains the stability of the thin-film resistor types.

Thick-film resistors normally use ruthenium oxide (RuO_2) as the resistance material, mixed with glass-based material to form a paste for printing on the substrate. The thickness of the printing material is usually 12 μm. The heat generated during laser trimming can cause microcracks on a thick-film resistor containing glass-based materials, which can adversely affect stability. Palladium/silver (PdAg) is used for the electrodes.

The most important thing about thick-film SM resistors from our point of view is that they do not obey Ohm's Law very well. This often comes as a shock to people who are used to TH resistors, which have been the highly linear metal film type for many years. They have much higher voltage coefficients than TH resistors, at between 30 and 100 ppm. The non-linearity is symmetrical about zero voltage and so gives rise to third-harmonic distortion. Some SM resistor manufacturers do not specify voltage coefficient, which usually means it can vary

www.raguvaran.puzl.com

disturbingly between different batches and different values of the same component, and this can have dire results on the repeatability of design performance.

Chip-type SM resistors come in standard formats with names based on size, such as 1206, 0805, 0603, and 0402. For example, 0805, which used to be something like the 'standard' size, is 0.08 inch by 0.05 inch (see Table 2.5). The smaller 0603 is now more common. Both 0805 and 0603 can be placed manually if you have a steady hand and a good magnifying glass. The 0402 size is so small that the resistors look rather like grains of pepper; manual placing is not really feasible. They are only used in equipment where small size is critical, such as mobile phones. They have very restricted voltage and power ratings, typically 50 V and 100 mW. The voltage rating of TH resistors can usually be ignored, as power dissipation is almost always the limiting factor, but with SM resistors it must be kept firmly in mind.

Recently, even smaller SM resistors have been introduced; for example, several vendors offer 0201, and Panasonic and Yageo offer 01005 resistors. The latter are truly tiny, being about 0.4 mm long; a thousand of them weigh less than a twentieth of a gram. They are intended for mobile phones, palmtops, and hearing aids; a full range of values is available from 10 Ω to 1 MΩ (jumper inclusive). Hand placing is really not an option.

Surface-mount resistors have a limited power-dissipation capability compared with their through-hole cousins, because of their small physical size. SM voltage ratings are also restricted, for the same reason. It is therefore sometimes necessary to use two SM resistors in series or parallel to meet these demands, as this is usually more economic than hand-fitting a through-hole component of adequate rating. If the voltage rating is the issue then the SM resistors will obviously have to be connected in series to gain any benefit.

Resistor Imperfections

It is well known that resistors have inductance and capacitance, and vary somewhat in resistance with temperature. Unfortunately there are other less obvious imperfections, such as excess noise and non-linearity; these can get forgotten because parameters describing how bad they are are often omitted from component manufacturers' data sheets.

Being components in the real world, resistors are not perfect examples of resistance and nothing else. Their length is not infinitely small and so they have series inductance; this is particularly true for the many kinds that use a spiral resistive element. Likewise, they exhibit stray capacitance between each end, and also between the various parts of the resistive element. Both effects can be significant at high frequencies, but can usually be ignored below 100 kHz unless you are using very high or low resistance values.

It is a sad fact that resistors change their value with temperature. Table 2.4 shows some typical temperature coefficients. This is not likely to be a problem in audio applications, where

www.raguvaran.puzl.com

extreme precision is not required unless you are designing measurement equipment. Carbon-film resistors are markedly inferior to metal film in this area.

Resistor Noise

All resistors, no matter what their resistive material or mode of construction, generate Johnson noise. This is white noise, its level being determined solely by the resistance value, the absolute temperature, and the bandwidth over which the noise is being measured. It is based on fundamental physics and is not subject to negotiation. In some cases it places the limit on how quiet a circuit can be, though the noise from active devices is often more significant. Johnson noise was covered in Chapter 1.

Excess resistor noise refers to the fact that a resistor, with a constant voltage drop across it, generates excess noise in addition to its inherent Johnson noise. According to classical physics, passing a current through a resistor should have no effect on its noise behavior; it should generate the same Johnson noise as a resistor with no steady current flow. In reality, some resistors do generate excess noise when they have a DC voltage across them. It is a very variable quantity, but is essentially proportional to the DC voltage across the component, a typical spec being '1 μV/V', and it has a $1/f$ frequency distribution. Typically it could be a problem in biasing networks at the input of amplifier stages. It is usually only of interest if you are using carbon- or thick-film resistors – metal-film and wirewound types should have very little excess noise. $1/f$ noise does not have a Gaussian amplitude distribution, which makes it difficult to assess reliably from a small set of data points. A rough guide to the likely specs is given in Table 2.6.

The level of excess resistor noise changes with resistor type, size, and value in ohms. The relevant factors are discussed below.

Thin-film resistors are markedly quieter than thick-film resistors; this is due to the homogeneous nature of thin-film resistive materials, which are metal alloys such as nickel–chromium deposited on a substrate. The thick-film resistive material is a mixture of fine metal (often ruthenium) oxides and glass particles; the glass is fused into a matrix for the metal particles by

TABLE 2.6 Resistor excess noise

Type	Noise (μV/V)
Carbon-film TH	0.2–3
Metal-oxide TH	0.1–1
Thin-film SM	0.05–0.4
Bulk metal-foil TH	0.01
Wirewound TH*	0

*Wirewound resistors are normally considered to be completely free of excess noise.

www.raguvaran.puzl.com

high-temperature firing. The higher excess noise levels associated with thick-film resistors are a consequence of their heterogeneous structure, due to the particulate nature of the resistive material. The same applies to carbon-film resistors where the resistive medium is finely divided carbon dispersed in a polymer binder.

A physically large resistor has lower excess noise than a small resistor. In the same resistor range, the highest wattage versions have the lowest noise (see Figure 2.4).

A low ohmic value resistor has lower excess noise than a high ohmic value resistor. Noise in μV per V rises approximately with the square root of resistance (see Figure 2.4 again).

A low value of excess noise is associated with uniform constriction-free current flow; this condition is not well met in composite thick-film materials. However, there are great variations among different thick-film resistors. The most readily apparent relationship is between noise level and the amount of conductive material present. Everything else being equal, compositions with lower resistivity have lower noise levels.

Higher resistance values give higher excess noise since it is a statistical phenomenon related to the total number of charge carriers available within the resistive element; the fewer the total number of carriers present, the greater will be the statistical fluctuation.

Traditionally at this point in the discussion of excess resistor noise, the reader is warned against using carbon composition resistors because of their very bad excess noise characteristics. Carbon composition resistors do still exist – their construction makes them good at handling pulse loads – but are not likely to be encountered in audio circuitry.

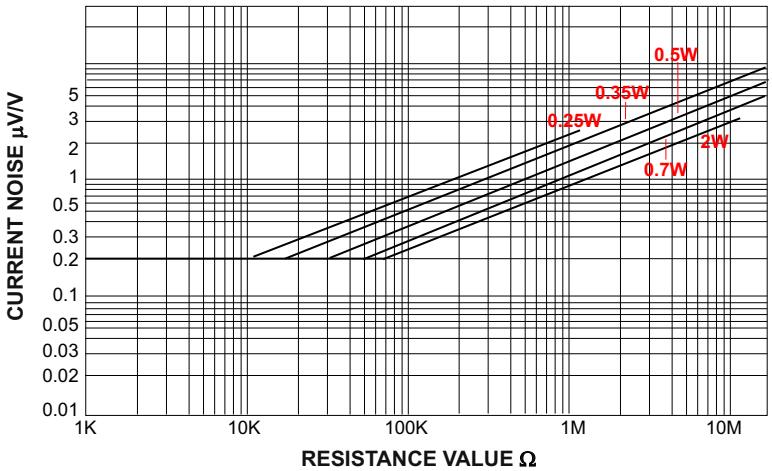


Figure 2.4: The typical variation of excess resistor noise with ohmic value and physical size; this is for a range of carbon-film resistors. The flat part of the plot represents the measurement floor, not a change in noise mechanism

www.raguvaran.puzl.com

One of the great benefits of op-amp circuitry is that it is noticeably free of resistors with large DC voltages across them. The offset voltages and bias currents are far too low to cause trouble with resistor excess noise. However, if you are getting into low-noise hybrid discrete/op-amp stages, such as the moving-coil head amplifier discussed in Chapter 8, you might have to consider it.

To get a feel for the magnitude of excess resistor noise, consider a 100 kΩ, 0.25 W carbon-film resistor with 10 V across it. This, from the graph in Figure 2.4, has a current noise parameter of about 0.7 μV/V and so the excess noise will be of the order of 7 μV, which is −101 dBu. This definitely could be a problem in a low-noise preamplifier stage.

Resistor Non-Linearity

Ohm's Law strictly is a statement about metallic conductors only. It is dangerous to assume that it also invariably applies to 'resistors' simply because they have a fixed value of resistance marked on them; in fact resistors – whose main *raison d'être* is packing a lot of controlled resistance in a small space – do not always adhere to Ohm's Law very closely. This is a distinct difficulty when trying to make low-distortion circuitry.

Resistor non-linearity is normally quoted by manufacturers as a voltage coefficient, usually the number of parts per million (ppm) that the resistor value changes when one volt is applied. The measurement standard for resistor non-linearity is IEC 6040.

Normal (i.e. lead-through-board) metal-film resistors show perfect linearity at the levels of performance considered here, as do wirewound types. The voltage coefficient is less than 1 ppm. Carbon-film resistors are quoted at less than 100 ppm; 100 ppm is, however, enough to completely dominate the distortion produced by active devices, if it is used in a critical part of the circuitry. Carbon composition resistors, probably of historical interest only, come in at about 350 ppm, a point that might be pondered by connoisseurs of antique equipment. The greatest area of concern over non-linearity is thick-film SM resistors, which have high and rather variable voltage coefficients (more on this below).

Table 2.7 (calculated with SPICE) gives the total harmonic distortion (THD) in the current flowing through the resistor for various voltage coefficients when a pure sine voltage is applied. If the voltage coefficient is significant this can be a serious source of non-linearity.

Distortion here is assumed to be second order, and so varies proportionally with level. Third-order distortion, which will be dominant if a resistor has no steady voltage across it, rises as the square of level.

My own test set-up is shown in Figure 2.5. The resistors are usually of equal value, to give 6 dB attenuation. A very-low-distortion oscillator that can give a large output voltage is necessary; the results in Figure 2.6 were taken at a 10 Vrms (+22 dBu) input level. Here thick-film SM and TH resistors are compared. The 'gen-mon' trace at the bottom is the record of the analyzer

www.raguvaran.puzl.com

TABLE 2.7 Resistor voltage coefficents and the resulting distortion at +15 and +20 dBu

Voltage coefficient (ppm)	THD at +15 dBu (%)	THD at +20 dBu (%)
1	0.00011	0.00019
3	0.00032	0.00056
10	0.0016	0.0019
30	0.0032	0.0056
100	0.011	0.019
320	0.034	0.060
1000	0.11	0.19
3000	0.32	0.58

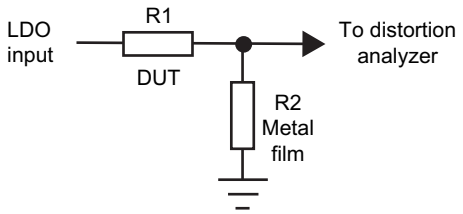


Figure 2.5: Test circuit for measuring resistor non-linearity. The not-under-test resistor R2 in the potential divider must be a metal-film type with negligible voltage coefficient

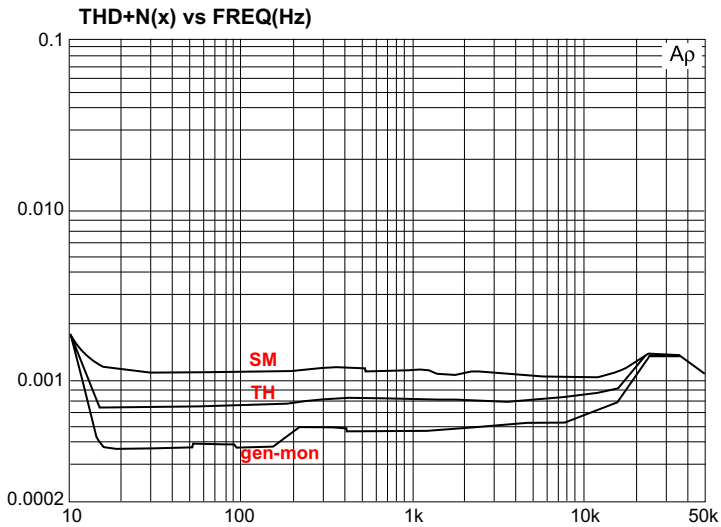


Figure 2.6: Surface-mount resistor distortion at 10 Vrms input, using 10 kΩ 0805 thick-film resistors

www.raguvaran.puzl.com

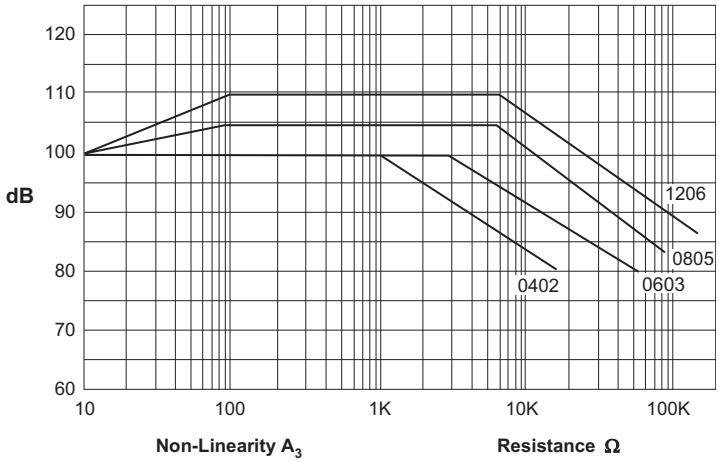


Figure 2.7: Non-linearity of thin-film surface-mount resistors of different sizes. THD is in dB rather than percent

reading the oscillator output and is the measurement floor of the AP System 1 used. The TH plot is higher than this floor, but this is not due to distortion. It simply reflects the extra Johnson noise generated by two 10 kΩ resistors. Their parallel combination is 5 kΩ, and so this noise is at −115.2 dBu. The SM plot, however, is higher again, and the difference is the distortion generated by the thick-film component.

For both thin-film and thick-film SM resistors non-linearity increases with resistor value, and also increases as the physical size (and hence power rating) of the resistor shrinks. The thin-film versions are much more linear (see Figures 2.7 and 2.8).

Sometimes it is appropriate to reduce the non-linearity by using multiple resistors in series. If one resistor is replaced by two with the same voltage coefficient in series, the THD in the

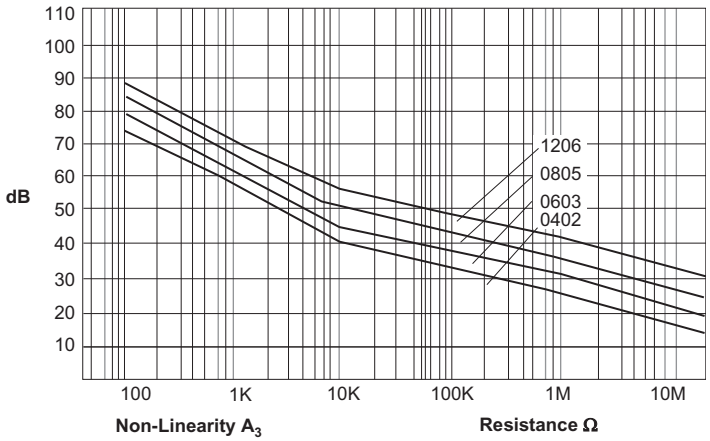


Figure 2.8: Non-linearity of thick-film surface-mount resistors of different sizes

www.raguvaran.puzl.com

current flowing is halved. Similarly, three resistors reduce THD to a third of the original value. There are obvious economic limits to this sort of thing, but it can be useful in specific cases, especially where the voltage rating of the resistor is a limitation.

Capacitors

Capacitors are diverse components. In the audio business their capacitance ranges from 10 pF to 100,000 μF, a ratio of 10^{10}. In this they handily outdo resistors, which usually vary from 0.1 Ω to 10 MΩ, a ratio of only 10^8. However, if you include the 10 GΩ bias resistors used in capacitor microphone head amplifiers, this range increases to 10^{11}. There is, however, a big gap between the 10 MΩ resistors, which are used in DC servos, and 10 GΩ microphone resistors; I am not aware of any audio applications for 1 GΩ resistors.

Capacitors also come in a wide variety of types of dielectric, the great divide being between electrolytic and non-electrolytic types. Electrolytics used to have *much* wider tolerances than most components, but things have recently improved and ±20% is now common. This is still wider than for typical non-electrolytics, which are usually ±10% or better.

This is not the place to reiterate the basic information about capacitor properties, which can be found from many sources. I will simply note that real capacitors fall short of the ideal circuit element in several ways, notably leakage, equivalent series resistance (ESR), dielectric absorption, and non-linearity.

Capacitor leakage is equivalent to a high value resistance across the capacitor terminals, and allows a trickle of current to flow when a DC voltage is applied. It is usually negligible for non-electrolytics, but is much greater for electrolytics.

ESR is a measure of how much the component deviates from a mathematically pure capacitance. The series resistance is partly due to the physical resistance of leads and foils, and partly due to losses in the dielectric. It can also be expressed as tan δ (tan-delta). Tan-delta is the tangent of the phase angle between the voltage across and the current flowing through the capacitor.

Dielectric absorption is a well-known phenomenon; take a large electrolytic, charge it up, and then make sure it is full discharged. You can use a screwdriver across the terminals if you're not too worried about either the screwdriver or the capacitor. Wait a few minutes, and the charge will partially reappear, as if from nowhere. This 'memory effect' also occurs in non-electrolytics to a lesser degree; it is a property of the dielectric, and is minimized by using polystyrene, polypropylene, NPO ceramic, or polytetrafluoroethylene dielectrics. Dielectric absorption is invariably simulated by a linear model composed of extra resistors and capacitances, and does not in itself appear to be a source of non-linearity.

Capacitor non-linearity is undoubtedly the least known of these shortcomings. A typical RC low-pass filter can be made with a series resistor and a shunt capacitor, and if you examine the

www.raguvaran.puzl.com

output with a distortion analyzer, you will find to your consternation that the circuit is not linear. If the capacitor is a non-electrolytic type with a dielectric such as polyester, then the distortion is relatively pure third harmonic, showing that the effect is symmetrical. For a 10 Vrms input, the THD level may be 0.001% or more. This may not sound like much but it is substantially greater than the mid-band distortion of a good op-amp. Capacitor non-linearity is dealt with at greater length below.

Capacitors are used in audio circuitry for three main functions, where their possible non-linearity has varying consequences:

1. Coupling or DC-blocking capacitors. These are usually electrolytics, and if properly sized have a negligible signal voltage across them at the lowest frequencies of interest. The properties of the capacitor are pretty much unimportant unless current levels are high; power amplifier output capacitors can generate considerable mid-band distortion [2]. Much nonsense has been talked about mysterious coupling capacitor properties, but it *is* all nonsense. For small-signal use, as long as the signal voltage across the capacitor is kept low, non-linearity is not normally detectable. The capacitance value is non-critical, as it has to be, given the wide tolerances of electrolytics.

2. Supply filtering or decoupling capacitors. These are electrolytics if you are filtering out supply-rail ripple, etc., and non-electrolytics, usually around 100 nF, when the task is to keep the supply impedance low at high frequencies and so keep op-amps stable. The capacitance value is again non-critical.

3. Setting time-constants, for example the capacitors in the feedback network of an RIAA amplifier. This is a much more demanding application than the other two. Firstly, the actual value is now crucially important as it defines the accuracy of the frequency response. Secondly, there is by definition significant signal voltage across the capacitor and so non-linearity can be a serious problem. Non-electrolytics are normally used; sometimes an electrolytic is used to define the lower end of the bandwidth, but this is a bad practice likely to introduce distortion at the bottom of the frequency range. Small-value ceramic capacitors are used for compensation purposes.

In Subjectivist circles it is frequently asserted that it is essential for good sound quality to bypass all coupling electrolytics with small non-electrolytic capacitors. This is quite untrue, as a moment's thought shows. If there is no signal voltage across the main coupling capacitor, what effect could adding a much smaller capacitor in parallel have?

Capacitor Non-Linearity Examined

When attempting the design of linear circuitry, everyone knows that inductors and transformers with ferromagnetic core material can be a source of non-linearity. It is, however, less obvious that capacitors and even resistors can show non-linearity and generate some

www.raguvaran.puzl.com

unexpected and very unwelcome distortion. Resistor non-linearity has been dealt with earlier in this chapter; let us examine the shortcomings of capacitors.

The definitive work on capacitor distortion is a magnificent series of articles by Cyril Bateman in *Electronics World* [3]. The authority of this work is underpinned by Cyril's background in capacitor manufacture. (The series is long because it includes the development of a low-distortion THD test set in the first two parts.)

Capacitors generate distortion when they are actually implementing a time-constant – in other words, when there is a signal voltage across them. The normal coupling or DC-blocking capacitors have no significant signal voltage across them, as they are intended to pass all the information through, not to filter it or define the system bandwidth. Capacitors with no signal across them do not generally produce distortion at small-signal current levels. This was confirmed for all the capacitors tested below. However, electrolytic types may do so at power amplifier levels where the current through them is considerable, such as in the output coupling capacitor of a power amplifier [2].

Non-Electrolytic Capacitor Non-Linearity

It has often been assumed that non-electrolytic capacitors, which generally approach an ideal component more closely than electrolytics, and have dielectrics constructed in a totally different way, are free from distortion. It is not so. Some non-electrolytics show distortion at levels that are easily measured, and can exceed the distortion from the op-amps in the circuit. Non-electrolytic capacitor distortion is primarily third harmonic, because the non-polarized dielectric technology is basically symmetrical. The problem is serious, because non-electrolytic capacitors are commonly used to define time-constants and frequency responses (in RIAA equalization networks, for example) rather than simply for DC blocking.

Very small capacitances present no great problem. Simply make sure you are using the C0G (NP0) type, and so long as you choose a reputable supplier, there will be no distortion. I say 'reputable supplier' because I did once encounter some allegedly C0G capacitors from China that showed significant non-linearity [4].

Middle-range capacitors, from 1 nF to 1 μF, present more of a problem. Capacitors with a variety of dielectrics are available, including polyester, polystyrene, polypropylene, polycarbonate, and polyphenylene sulfide, of which the first three are the most common. (Note that what is commonly called 'polyester' is actually polyethylene terephthalate.)

Figure 2.9 shows a simple low-pass filter circuit which, in conjunction with a good THD analyzer, can be used to get some insight into the distortion problem; it is intended to be representative of a real bit of audio circuitry. The values shown give a pole frequency, or −3 dB roll-off point, at 710 Hz. Since it might be expected that different dielectrics give

www.raguvaran.puzl.com

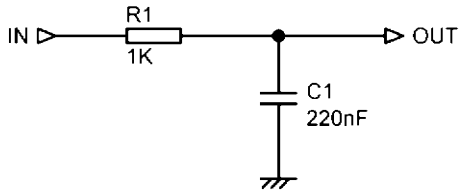


Figure 2.9: Simple low-pass test circuit for non-electrolytic capacitor distortion

different results (and they definitely do) we will start off with polyester, the smallest, most economical, and therefore the most common type for capacitors of this size.

The THD results for a microbox 220 nF, 100 V capacitor with a polyester dielectric are shown in Figure 2.10, for input voltages of 10, 15, and 20 Vrms. They are unsettling.

The distortion is all third harmonic and peaks at around 300–400 Hz, well below the pole frequency, and even with input limited to 10 Vrms will exceed the non-linearity introduced by op-amps such as the 5532 and the LM4562. Interestingly, the peak frequency changes with applied level. Below the peak, the voltage across the capacitor is constant but distortion falls as frequency is reduced, because the increasing impedance of the capacitor means it has less effect on a circuit node at 1 kΩ impedance. Above the peak, distortion falls with increasing frequency because the low-pass circuit action causes the voltage across the capacitor to fall.

The level of distortion varies with different samples of the same type of capacitor; six of the above type were measured and the THD at 10 Vrms and 400 Hz varied from 0.00128% to 0.00206%. This puts paid to any plans for reducing the distortion by some sort of cancellation method.

The distortion can be seen in Figure 2.10 to be a strong function of level, roughly tripling as the input level doubles. Third-harmonic distortion normally quadruples for doubled level, so there may well be an unanswered question here. It is, however, clear that reducing the voltage across

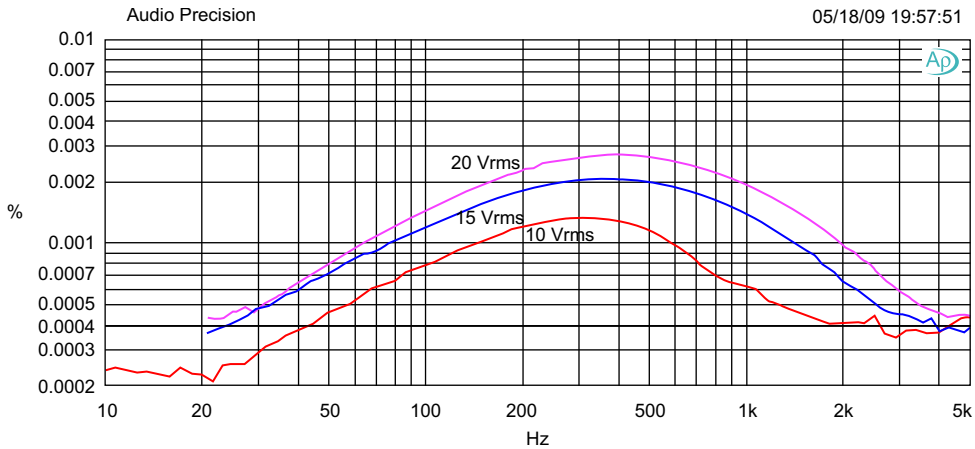


Figure 2.10: Third-harmonic distortion from a 220 nF, 100 V polyester capacitor, at 10, 15, and 20 Vrms input level, showing peaking around 400 Hz

www.raguvaran.puzl.com

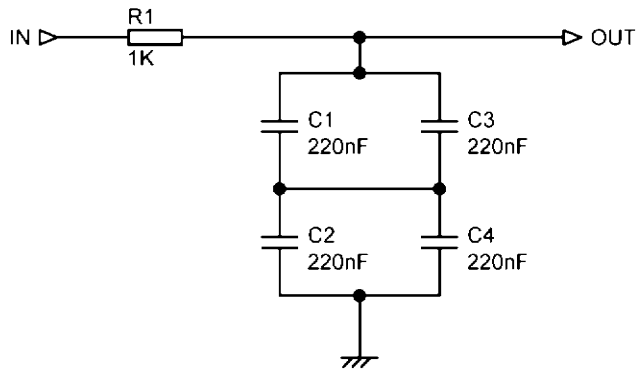


Figure 2.11: Reducing capacitor distortion by series-parallel connection

the capacitor reduces the distortion. This suggests that if cost is not the primary consideration, it might be useful to put two capacitors in series to halve the voltage, and the capacitance, and then double up this series combination to restore the original capacitance, giving the series-parallel arrangement in Figure 2.11. The results are shown in Table 2.8, and once more it can be seen that halving the level has reduced distortion by a factor of 3 rather than 4.

The series-parallel arrangement obviously has limitations in terms of cost and PCB area occupied, but might be useful in some cases.

Clearly polyester gives significant distortion, despite its extensive use in audio circuitry of all kinds. The next dielectric we will try is polystyrene. Capacitors with a polystyrene dielectric are extremely useful for some filtering and RIAA-equalization applications because they can be obtained at a 1% tolerance at up to 10 nF at a reasonable price. They can be obtained in larger sizes at an unreasonable, or at any rate much higher, price.

The distortion test results are shown in Figure 2.12 for a 4n7 2.5% capacitor; the series resistor R1 has been increased to 4.7 kΩ to keep the −3 dB point inside the audio band, and it is now at 7200 Hz. Note that the THD scale has been extended down to a subterranean 0.0001%, and if it was plotted on the same scale as in Figure 2.10 it would be bumping along the bottom of the graph. Figure 2.12 in fact shows no distortion at all, just the measurement noise floor, and the

TABLE 2.8 The reduction of polyester capacitor distortion by series-parallel connection

Input level (Vrms)	Single capacitor	Series-parallel capacitors
10	0.0016%	0.00048%
15	0.0023%	0.00098%
20	0.0034%	0.0013%

www.raguvaran.puzl.com

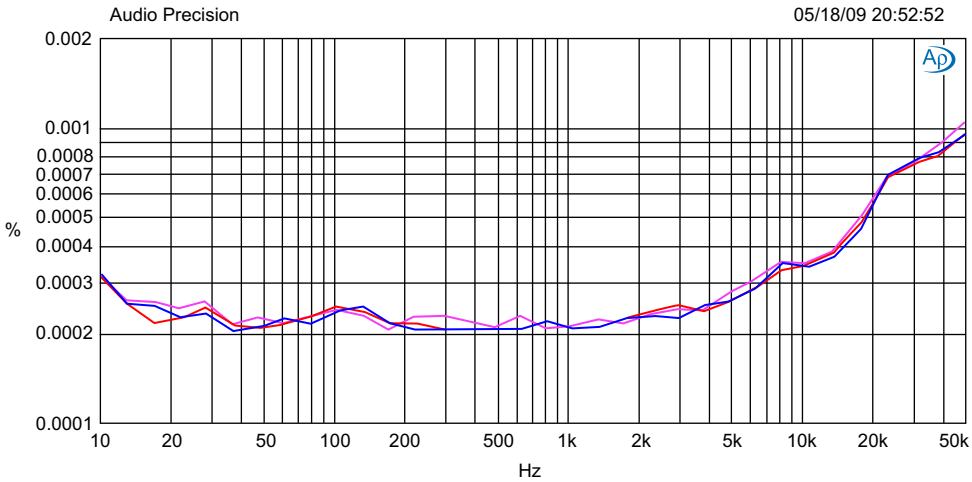


Figure 2.12: The THD plot with three samples of 4n7, 2.5% polystyrene capacitors, at 10 Vrms input level. The reading is entirely noise

apparent rise at the high-frequency (HF) end is simply due to the fact that the output level is decreasing, because of the low-pass action, and so the noise floor is relatively increasing. This is at an input level of 10 Vrms, which is about as high as might be expected to occur in normal op-amp circuitry. The test was repeated at 20 Vrms, which might be encountered in discrete circuitry, and the results were the same – no measurable distortion.

The tests were done with four samples of 10 nF 1% polystyrene from LCR at 10 and 20 Vrms, with the same results for each sample. This shows that polystyrene capacitors can be used with confidence; this finding is in complete agreement with Cyril Bateman's findings [5].

Having settled the issue of capacitor distortion below 10 nF, we need now to tackle its capacitor values greater than 10 nF. Polyester having proven unsatisfactory, the next most common dielectric is polypropylene, and I might as well say at once that it was with considerable relief that I found these were effectively distortion-free in values up to 220 nF. Figure 2.13 shows the results for four samples of a 220 nF 250 V 5% polypropylene capacitor from RIFA. Once more the plot shows no distortion at all, just the noise floor, the apparent rise at the HF end being increasing relative noise due to the low-pass roll-off. This is also in agreement with Cyril Bateman's findings [6]. Rerunning the tests at 20 Vrms gave the same result – no distortion. This is very pleasing, but there is a downside. Polypropylene capacitors of this value and voltage rating are much larger than the commonly used 63 or 100 V polyester capacitor, and more expensive.

It was therefore important to find out if the good distortion performance was a result of the 250 V rating, and so I tested a series of polypropylene capacitors with lower voltage ratings from different manufacturers. Axial 47 nF 160 V 5% polypropylene capacitors from Vishay

www.raguvaran.puzl.com

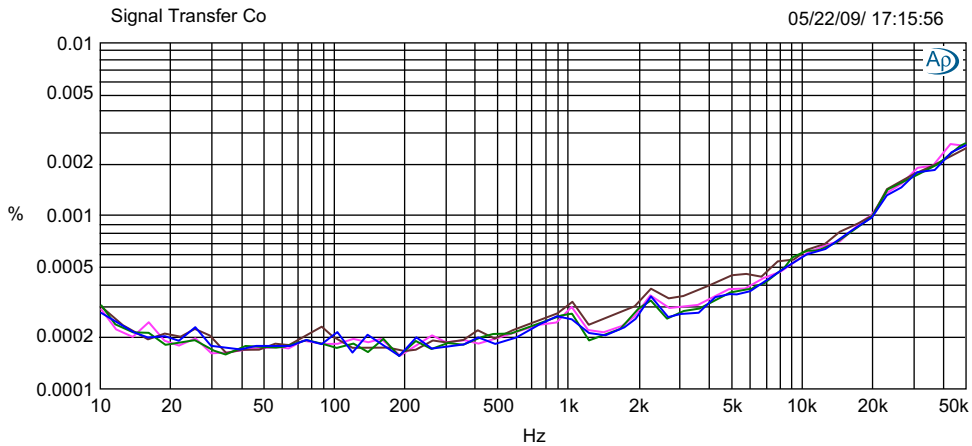


Figure 2.13: The THD plot with four samples of 220 nF, 250 V, 5% polypropylene capacitors, at 10 Vrms input level. The reading is again entirely noise

proved to be THD-free at both 10 and 20 Vrms. Likewise, microbox polypropylene capacitors from 10 to 47 nF, with ratings of 63 and 160 V from Vishay and Wima, proved to generate no measurable distortion, so the voltage rating appears not to be an issue. This finding is particularly important because the Vishay range has a 1% tolerance, making them very suitable for precision filters and equalization networks. The 1% tolerance is naturally reflected in the price.

The only remaining issue with polypropylene capacitors is that the higher values (above 100 nF) appear to be currently only available with 250 or 400 V ratings, and that means a physically big component. For example, the EPCOS 330 nF 400 V 5% part has a footprint of 26 mm by 6.5 mm, with a height of 15 mm. One way of dealing with this is to use a smaller capacitor in a capacitance multiplication configuration, so a 100 nF, 1% component could be made to emulate 470 nF. It has to be said that the circuitry for this is only straightforward if one end of the capacitor is connected to ground.

When I first started looking at capacitor distortion, I thought that the distortion would probably be lowest for the capacitors with the highest voltage rating. I therefore tested some RF-suppression X2 capacitors, rated at 275 Vrms, which translates into a peak or DC rating of 389 V. These are designed to be connected across the mains and therefore have a thick and tough dielectric layer. For some reason manufacturers seem to be very coy about saying exactly what the dielectric material is, normally describing them simply as 'film capacitors'. A problem that surfaced immediately is that the tolerance is 10% or 20%, not exactly ideal for precision filtering or equalization. A more serious problem, however, is that they are far from distortion-free. Four samples of a 470 nF X2 capacitor showed THD between 0.002% and 0.003% at 10 Vrms. Clearly a high voltage rating alone does not mean low distortion.

www.raguvaran.puzl.com

Electrolytic Capacitor Non-Linearity

Cyril Bateman's series in *Electronics World* [3] included two articles on electrolytic capacitor distortion. It proved to be a complex subject, and many long-held assumptions (such as 'DC biasing always reduces distortion') were shown to be quite wrong. Distortion was in general a good deal higher than for non-electrolytic capacitors.

My view is that electrolytics should never, ever, under any circumstances, be used to set time-constants in audio. There should be a time-constant early in the signal path, based on a non-electrolytic capacitor, that determines the lower limit of the bandwidth, and all the electrolytic-based time-constants should be much longer so that the electrolytic capacitors can never have significant signal voltages across them and so never generate measurable distortion. There is of course also the point that electrolytics have large tolerances, and cannot be used to set accurate time-constants anyway.

However, even if you obey this rule, you can still get into deep trouble. Figure 2.14 shows a simple high-pass test circuit designed to represent an electrolytic capacitor in use for coupling or DC blocking. The load of 1 kΩ is the sort of value that can easily be encountered if you are using low-impedance design principles. The calculated −3 dB roll-off point is 3.38 Hz, so the attenuation at 10 Hz, at the very bottom of the audio band, will be only 0.47 dB; at 20 Hz it will be only 0.12 dB, which is surely a negligible loss. As far as frequency response goes, we are doing fine. But examine Figure 2.15, which shows the measured distortion of this arrangement. Even if we limit ourselves to a 10 Vrms level, the distortion at 50 Hz is 0.001%, already above that of a good op-amp. At 20 Hz it has risen to 0.01%, and by 10 Hz a most unwelcome 0.05%. The THD is increasing by a ratio of 4.8 times for each octave fall in frequency – in other words, increasing faster than a square law. The distortion residual is visually a mixture of second and third harmonic, and the levels proved surprisingly consistent for a large number of 47 µF 25 V capacitors of different ages and from different manufacturers.

Figure 2.15 also shows that the distortion rises rapidly with level; at 50 Hz going from an input of 10 to 15 Vrms almost doubles the THD reading. To underline the point, consider Figure 2.16, which shows the measured frequency response of the circuit with 47 µF and 1 kΩ; note the effect of the capacitor tolerance on the real versus calculated figures. The roll-off that

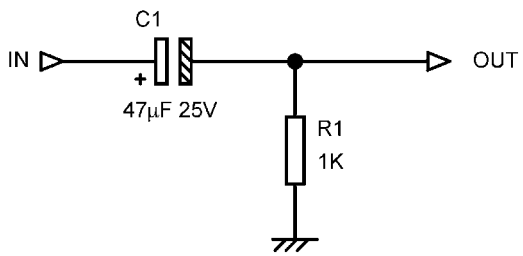


Figure 2.14: High-pass test circuit for examining electrolytic capacitor distortion

www.raguvaran.puzl.com

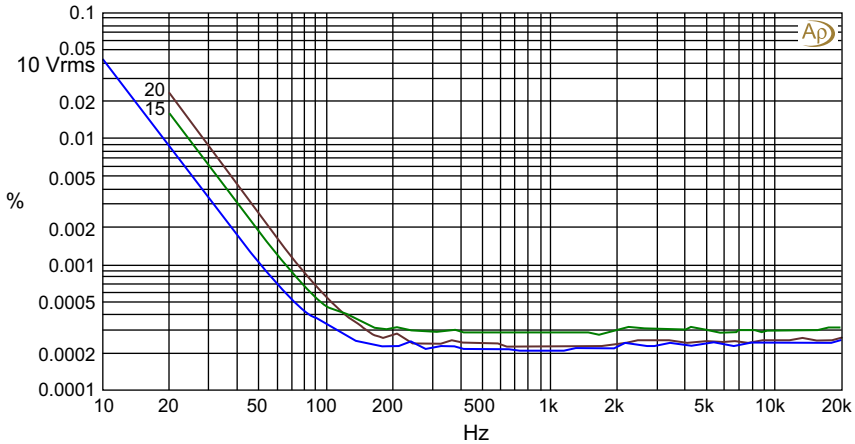


Figure 2.15: Electrolytic capacitor distortion from the circuit in Figure 2.13. Input level 10, 15, and 20 Vrms

does the damage, by allowing an AC voltage to exist across the capacitor, is very modest indeed, less than 0.2 dB at 20 Hz.

Having demonstrated how insidious this problem is, how do we fix it? As we have seen, changing capacitor manufacturer is no help. Using 47 µF capacitors of higher voltage does not work – tests showed there is very little difference in the amount of distortion generated. An exception was the sub-miniature style of electrolytic, which was markedly worse.

The answer is simple – just make the capacitor bigger in value. This reduces the voltage across it in the audio band, and since we have shown that the distortion is a strong function of the

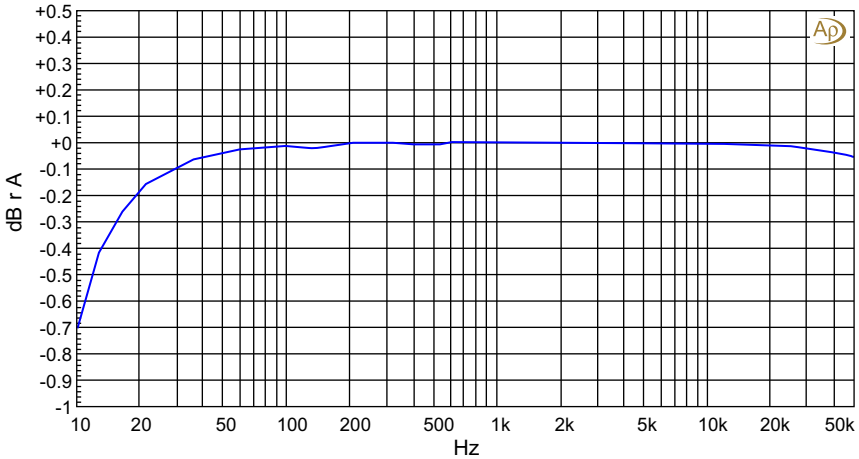


Figure 2.16: The measured roll-off of the high-pass test circuit for examining electrolytic capacitor distortion

www.raguvaran.puzl.com

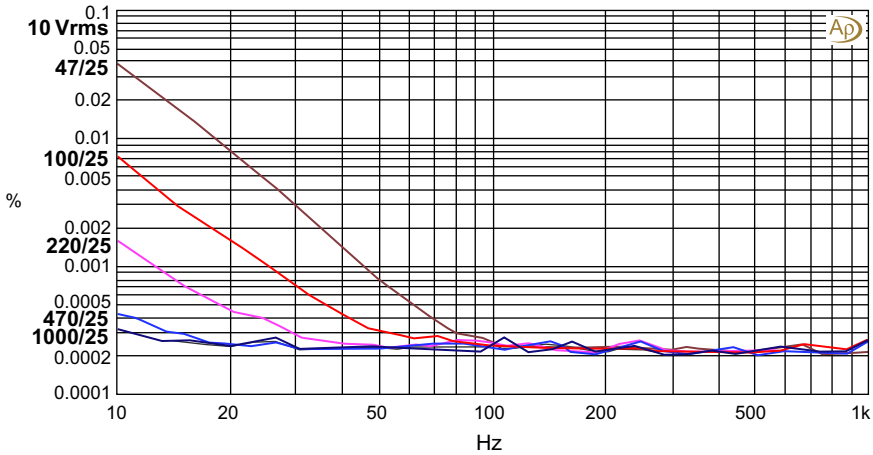


Figure 2.17: Reducing electrolytic capacitor distortion by increasing the capacitor value. Input 10 Vrms

voltage across the capacitor, the amount produced drops more than proportionally. The result is seen in Figure 2.17, for increasing capacitor values with a 10 Vrms input.

Replacing C1 with a 100 µF 25 V capacitor drops the distortion at 20 Hz from 0.0080% to 0.0017%, an improvement of 4.7 times; the voltage across the capacitor at 20 Hz has been reduced from 1.66 Vrms to 790 mVrms. A 220 µF 25 V capacitor reduces the voltage across itself to 360 mV, and gives another very welcome reduction to 0.0005% at 20 Hz, but it is necessary to go to 1000 µF 25 V to obtain the bottom trace, which is indistinguishable from the noise floor of the AP-2702 test system. The voltage across the capacitor at 20 Hz is now only 80 mV. From this data, it appears that the AC voltage across an electrolytic capacitor should be limited to below 80 mVrms if you want to avoid distortion. I would emphasize that these are ordinary 85°C rated electrolytic capacitors, and in no sense special or premium types.

This technique can be seen to be highly effective, but it naturally calls for larger and somewhat more expensive capacitors, and larger footprints on a PCB. This can be to some extent countered by using capacitors of lower voltage, which helps to bring back down the *CV* product and hence the capacitor volume. I tested 1000 µF 16 V and 1000 µF 6V3 capacitors, and both types gave exactly the same results as the 1000 µF 25 V part in Figure 2.17, with useful reductions in *CV* product and can size. This does of course assume that the capacitor is, as is usual, being used to block small voltages from op-amp offsets to prevent switch clicks and pot noises rather than for stopping a substantial DC voltage.

The use of large coupling capacitors in this way does require a little care, because we are introducing a long time-constant into the circuit. Most op-amp circuitry is pretty much free of big DC voltages, but if there are any, the settling time after switch-on may become undesirably long.

www.raguvaran.puzl.com

More information on capacitor distortion in specific applications can be found in chapters 3 and 5.

Inductors

For several reasons, inductors are unpopular with circuit designers. They are relatively expensive, often because they need to be custom made. Unless they are air-cored (which limits their inductance to low values) the core material is a likely source of non-linearity. Some types produce substantial external magnetic fields, which can cause crosstalk if they are placed close together, and similarly they can be subject to the induction of interference from other external fields. In general they deviate from being an ideal circuit element much more than resistors or capacitors.

It is rarely, if ever, essential to use inductors in signal-processing circuitry. Historically they were used in tone controls, before the Baxandall configuration swept all before it, and their last applications were probably in mid EQ controls for mixing consoles and in LCR filters for graphic equalizers. These too were gone by the end of the 1970s, being replaced by active filters and gyrators, to the considerable relief of all concerned (except inductor manufacturers).

The only place where inductors are essential is when the need for galvanic isolation, or enhanced EMC immunity, makes input and output transformers desirable, and even then they need careful handling (see Chapters 14 and 15 on line-in and line-out circuitry).

References

[1] D. Self, Ultra-low-noise amplifiers and granularity distortion, JAES (November 1987), pp. 907–915.
[2] D. Self, Audio Power Amplifier Handbook, fifth ed, Focal Press, 2009, p. 43 (amp output cap).
[3] C. Bateman, Capacitor sound? Parts 1–6, Electronics World (July 2002 to March 2003).
[4] D. Self, Audio Power Amplifier Design Handbook, fifth ed., Focal Press, 2009, p. 205 (C0G cap).
[5] C. Bateman, Capacitor sound? Part 3, Electronics World (October 2002), p. 16, 18.
[6] C. Bateman, Capacitor sound? Part 4, Electronics World (November 2002) p. 47.

www.raguvaran.puzl.com

This page intentionally left blank

www.raguvaran.puzl.com

CHAPTER 3

Discrete Transistor Circuitry

This chapter deals with small-signal design using discrete transistors, specifically BJTs. Many things found in standard textbooks are skated over quickly. It concentrates on audio issues, and gives information that I do not think appears anywhere else, including the distortion behavior of various configurations.

Why Use Discrete Transistor Circuitry?

Circuitry made with discrete transistors is not obsolete. It is appropriate when:

1. a load must be driven to higher voltages than the op-amp can sustain between the supply rails. Op-amps are mostly restricted to supply voltages of ±18 or ±20 V. Hybrid-construction amplifiers, typically packaged in TO-3 cans, will operate from rails as high as ±100 V, but they are very expensive, and not optimized for audio use in parameters like crossover distortion. Discrete op-amps provide a viable alternative;

2. a load requires more drive current, because of its low impedance, than an op-amp can provide without overheating or current limiting, e.g. any audio power amplifier;

3. the best possible noise performance is required. Discrete bipolar transistors can outperform op-amps, particularly with low source resistances, say 500 Ω or less. The commonest examples are moving-coil head amps and microphone preamplifiers. These almost invariably use a discrete input device or devices, with the open-loop gain (for linearity) and load-driving capability provided by an op-amp which may itself have fairly humble noise specs;

4. the best possible distortion performance is demanded. Most op-amps have Class-B or -AB output stages, and many of them (though certainly not all) show clear crossover artefacts on the distortion residual. A discrete op-amp can dissipate more power than an IC, and so can have a Class-A output stage, sidestepping the crossover problem completely;

5. it would be necessary to provide a low-voltage supply to run just one or two op-amps. The cost of extra transformer windings, rectifiers, reservoirs, and regulators will buy a lot of discrete transistors. For example, if you need a buffer stage to drive a power amplifier from a low impedance, it may be more economical, and save space and weight, to use a discrete

Small Signal Audio Design; ISBN: 9780240521770
Copyright © 2010 Elsevier Ltd; All rights of reproduction, in any form, reserved.

www.raguvaran.puzl.com

emitter-follower running from the same rails as the power amplifier. In these days of auto-insertion; fitting the extra parts on the PCB will cost very little;

6. purely for marketing purposes, as you think you can mine a vein of customers that don't trust op-amps.

When studying the higher reaches of discrete design, the most fruitful source of information is paradoxically papers on analog IC design. This applies with particular force to design with BJTs. The circuitry used in ICs can rarely be directly adapted for use with discrete semiconductors, because some features such as multiple collector transistors and differing emitter areas simply do not exist in the discrete transistor world; it is the basic principles of circuit operation that can be useful. A good example is a paper by Erdi, dealing with a unity-gain buffer with a slew rate of 300 V/μs [1]. Another highly informative discourse is by Barry Hilton [2], which also deals with a unity-gain buffer.

Bipolars and FETs

This chapter only deals with bipolar transistors. Their high transconductance and predictable operation make them far more versatile than FETs. The highly variable V_g values of an FET can be dealt with by expedients such as current-source biasing, but the low gain, which means low feedback and poor linearity, remains a problem. FETs have their uses when super-high input impedances are required, and an example of a JFET working with an op-amp that provides loop gain can be found in Chapter 13 on microphone amplifiers.

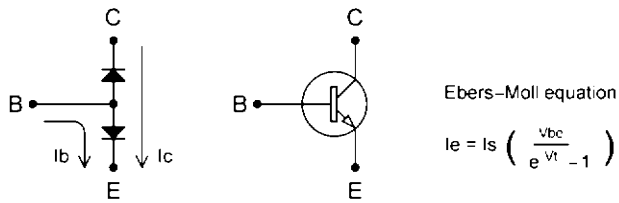
Bipolar Junction Transistors

There is one thing to get straight first: *the bipolar junction transistor is a voltage-operated device*.

What counts is the base-emitter voltage, or V_{be}. Certainly a BJT needs base current to flow for it to operate, but this is really an annoying imperfection rather than the basis of operation. I appreciate this may take some digesting; far too many discussions of transistor action say something like 'a small current flowing into the base controls a much larger current flowing into the collector'. In fact the only truly current-operated amplifying device that comes to mind is the Hall-effect multiplier, and you don't come across those every day. I've certainly never seen one used in audio – could be a market niche there.

Transistor operation is thus: if the base is open-circuit, then no collector current flows, as the collector–base junction is effectively a reverse-biased diode, as seen in Figure 3.1. There is

www.raguvaran.puzl.com



Ie = Ic + Ib

Is is the reverse saturation current of the base-emitter diode (of the order of 10^{-15} to 10^{-12} amp)

Vbe is the base-emitter voltage

Vt is the thermal voltage kT/q (approx 26 mV at 25°C room temperature).

k is Boltsmann's constant 1.381×10^{-23} joule/K

T is absolute temperature in K

q is the electron charge 1.6022×10^{-19} coulomb

Figure 3.1: Current flow through a bipolar transistor and the fundamental transistor equation

a little leakage through from the collector to the emitter, but with modern silicon BJTs you can usually ignore it.

When the base is forward biased by taking it about 600 mV above the emitter, charge carriers are launched into the base region. Since the base region is narrow, the vast majority shoot through into the collector, to form the collector current I_c. Only a small proportion of these carriers are snared in the base and become the base current I_b, which is clearly a result and not the cause of the base-emitter voltage. I_b is normally just a nuisance.

The Transistor Equation

Every bipolar transistor obeys the Ebers–Moll transistor equation shown in Figure 3.1 with startling accuracy over nine or ten decades of I_c, which is a pretty broad hint that we are looking at the fundamental mechanism. In contrast, beta varies with I_c, temperature, and just about everything else you can think of. The collector current is to a first approximation independent of collector voltage – in other words, it is a current-source output. The qualifications to this are as follows:

1. This only holds for V_{ce} above, say, 2 V.
2. It is not a perfect current source; even with a high V_{ce}, I_c increases slowly with V_{ce}. This is called the Early Effect, after Jim Early [3], and has nothing to do with timing or punctuality. It is a major consideration in the design of stages with high voltage gain. The same effect when the transistor is operated in reverse mode – a perversion that will not concern us here – has sometimes been called the Late Effect. Ho-ho.

www.raguvaran.puzl.com

Beta

Beta (or h_{fe}) is the ratio of the base current I_b to the collector current I_c. It is not a fundamental property of a BJT. Never design circuits that depend on beta, unless of course you're making a transistor tester.

Here are some of the factors that affect beta. This should convince you that it is a shifty and thoroughly untrustworthy parameter:

- Beta varies with I_c. First it rises as I_c increases, reaching a broad peak, and then it falls off as I_c continues to increase.
- Beta increases with temperature. This seems to be relatively little known. Most things, like leakage currents, get worse as temperature increases, so this makes a nice change.
- Beta is lower for high-current transistor types.
- Beta is lower for high V_{ceo} transistor types. This is a major consideration when you are designing the small-signal stages of power amplifiers with high supply rails.
- Beta varies widely between nominally identical examples of the same transistor type.

A very good refutation of the beta-centric view of BJTs is given by Barrie Gilbert [4].

Unity-Gain Buffer Stages

A buffer stage is used to isolate two portions of circuitry from each other. It has a high input impedance and low output impedance; typically it prevents things downstream from loading things upstream. The use of the word 'buffer' normally implies 'unity-gain buffer' because otherwise we would be talking about an amplifier or gain stage. The gain with the simpler discrete implementations is in fact slightly less than 1. The simplest discrete buffer circuit-block is the one-transistor emitter-follower; it is less than ideal both in its mediocre linearity and its asymmetrical load-driving capabilities. If we permit ourselves another transistor, the complementary feedback pair (CFP) configuration gives better linearity. Both versions can have their load-driving performance much improved by replacing the emitter resistor with a constant-current source or a push–pull Class-A output arrangement.

If the CFP stage is not sufficiently linear, the next stage in sophistication is to combine an input differential pair with an output emitter-follower, using three transistors. This arrangement, often called the Schlotzaur configuration [5], can be elaborated until it gives a truly excellent distortion performance. Its load-driving capability can be enhanced in the same way as in the simpler configurations.

www.raguvaran.puzl.com

The Simple Emitter-Follower

The simplest discrete circuit-block is the one-transistor emitter-follower. This count of one does not include extra transistors used as current sources, etc., to improve load-driving ability. It does not have a gain of exactly 1, but it is usually pretty close.

Figure 3.2 shows a simple emitter-follower with a 2k7 emitter resistor R2, giving a quiescent current of 8.8 mA with ±24 V supply rails. Biasing is by a high-value resistor R1 connected to 0 V. Note the polarity of the output capacitor; the output will sit at about −0.6 V due to the V_{be} drop, plus a little lower due to the voltage drop caused by the base current I_b flowing through R1.

The input impedance is approximately that of the emitter resistor in parallel with an external loading on the stage multiplied by the transistor beta:

$$R_{in} = \beta(R_e \| R_{load}) \qquad \text{(Equation 3.1)}$$

Don't expect the output impedance to be as low as an op-amp with plenty of negative feedback (NFB). The output impedance is approximately that of the source resistance divided by beta:

$$R_{out} = R_s/\beta \qquad \text{(Equation 3.2)}$$

The gain of an emitter-follower is always slightly less than unity, because of the finite transconductance of the transistor. Essentially the intrinsic emitter resistance r_e (not to be confused with the physical component Re) forms a potential divider with the output load. It is simple to work out the small-signal gain at a given operating point. The value of r_e is given by

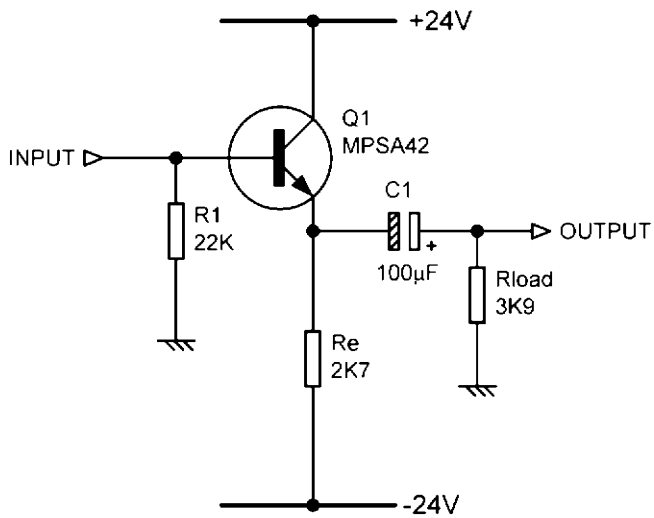


Figure 3.2: The simple emitter-follower circuit running from ±24 V supply rails

www.raguvaran.puzl.com

$25/I_c$ (for I_c in mA). Since the value of r_e is inversely proportional to I_c, it varies with large signals, and it is one cause of the rather imperfect linearity of the simple emitter-follower. Heavier external loading increases the modulation of I_c, increases the gain variation, and so increases distortion.

Figure 3.2 shows the emitter-follower with an AC-coupled load. Its load-driving capability is not very good. While the transistor can, within limits, source as much current as required into the load, the current-sinking ability is limited by the emitter resistor Re, which forms a potential divider with the load resistance R_{load}. With the values shown here, negative clipping occurs at about −10 V, severely limiting the maximum output amplitude. The circuit shows a high-voltage low-beta transistor type running from ±24 V rails, to give worst-case performance results and to exploit the ability of discrete circuitry to run from high-voltage rails.

The simple emitter-follower has several factors that affect its distortion performance:

- Distortion is reduced as the emitter resistor is reduced, for a given load impedance.
- Distortion is reduced as the DC bias level is raised above the mid-point.
- Distortion increases as the load impedance is reduced.
- Distortion increases monotonically with output level.
- Distortion does *not* vary with the beta of the transistor. This statement assumes low-impedance drive, which may not be the case for an emitter-follower used as a buffer. If the source impedance is significant then beta is likely to have a complicated effect on linearity – and not always for the worse.

Emitter-follower distortion is mainly second harmonic, except when closely approaching clipping. This is entirely predictable, as the circuit is asymmetrical. Only symmetrical configurations, such as the differential pair, restrict themselves to generating odd harmonics only, and then only when they are carefully balanced [6]. Symmetry is often praised as desirable in an audio circuit, but this is subject to Gershwin's Law: 'It ain't necessarily so'. Linearity is what we want in a circuit, and symmetry is not necessarily the best way to get it.

Because the distortion is mostly second harmonic, its level is proportional to amplitude, as seen in Figure 3.3. At 2 Vrms with no load it is about 0.006%, rising to 0.013% at 4 Vrms. External loading always makes the distortion worse, and more rapidly as the amplitude approaches the clipping point; the THD is more than doubled from 0.021% to 0.050% at 6 Vrms, just by adding a light 6k8 load. For these tests Re was 2k7, as in Figure 3.2.

For both the emitter-follower (EF) and CFP circuits, distortion is flat across the audio band so no THD/frequency plots are given.

Insight into what's happening can be gained by using SPICE to plot the incremental gain over the output swing, as in Figure 3.4. As loading increases the curvature of the gain characteristic

www.raguvaran.puzl.com

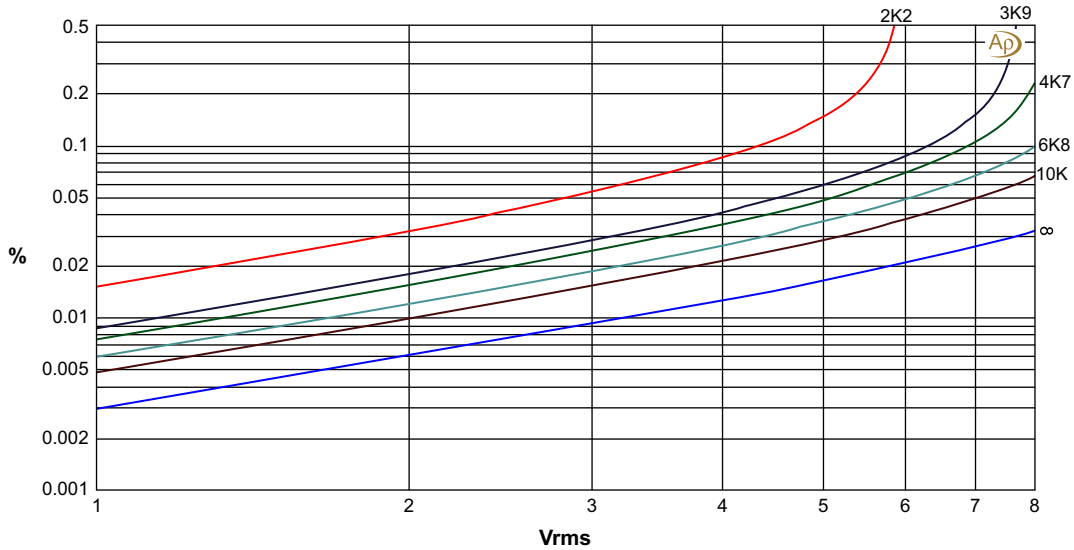


Figure 3.3: How various external loads degrade the linearity of the simple emitter-follower. Re = 2k7

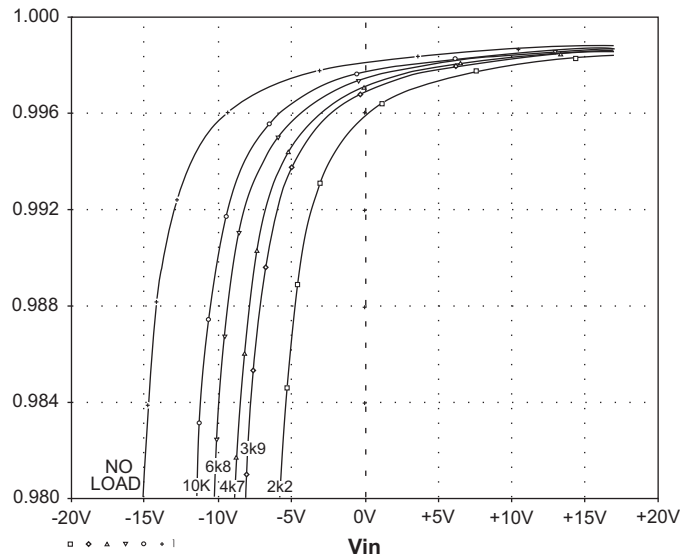


Figure 3.4: Incremental gain of the circuit in Figure 3.2, with a different loads. A distortionless circuit would have constant gain and so give horizontal lines. SPICE simulation

becomes greater for a given voltage swing. It is obvious that the circuit is much more linear on the positive side of 0 V, explaining why emitter-followers give less distortion when biased above the mid-point. This trick can be very useful if the full output swing is not required.

www.raguvaran.puzl.com

Most amplifier stages are biased so the quiescent output voltage is at the mid-point of the operating region, to allow the maximum symmetrical voltage swing. However, the asymmetry of the simple emitter-follower's output current capability means that if there is significant loading, a greater symmetrical output swing is often possible if the stage is biased positive of 0 V. If the output is loaded with 2k2, negative clipping occurs at −8 V, which allows a maximum output amplitude of only 5.6 Vrms. The unloaded output capability is about 12 Vrms. If the bias point is raised from 0 to +5 V, the output capability becomes roughly symmetrical and the maximum amplitude is increased to 9.2 Vrms. Don't forget to turn the output capacitor around.

The measured noise output of this stage with a 40 Ω source resistance is a commendably low −122.7 dBu (22 Hz to 22 kHz) but with a base-stopper resistor (see below) of 1 kΩ this degrades to −116.9 dBu. A higher stopper of 2k7 gives −110.4 dBu.

The Constant-Current Emitter-Follower

The simple emitter-follower can be greatly improved by replacing the sink resistor Re with a constant-current source, as shown in Figure 3.5. The voltage across a current source does not (to a first approximation) affect the current through it, so if the sink current is large enough a load can be driven to the full voltage swing in both directions.

The current source Q2 is biased by D1, D2. One diode cancels the V_{be} drop of Q2, while the other sets up 0.6 V across the 100 Ω resistor R2, establishing the quiescent current at 6 mA.

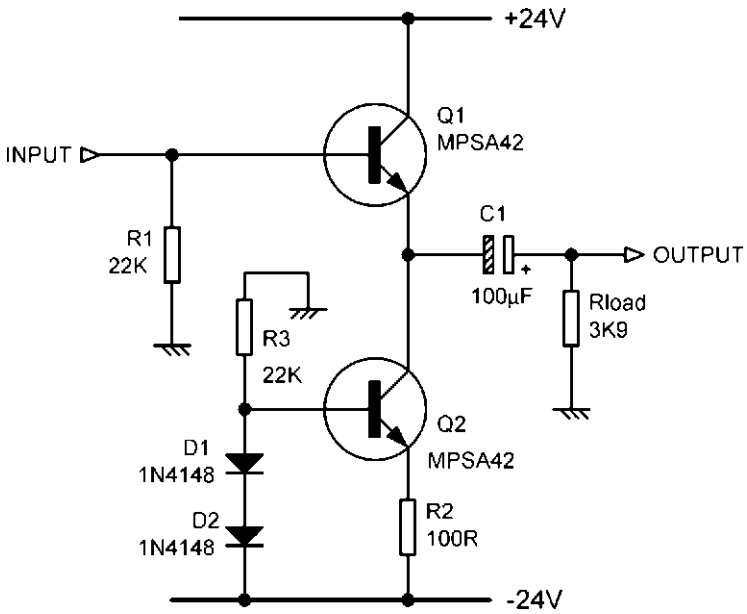


Figure 3.5: Emitter-follower with a constant-current source replacing the emitter resistor

www.raguvaran.puzl.com

The 22 kΩ resistor R3 in turn biases the diodes. This simple bias system works quite adequately if the supply rails are regulated, but might require filtering if they are not. The 22 kΩ value is non-critical; so long as the diode current exceeds the I_b of Q2 by a reasonable factor (say 10 times) there will be no problem.

Figure 3.6 shows that distortion is much reduced. With no external load, 0.013% at 4 Vrms has become less than 0.0003%, the measurement system noise floor. This is because the amount by which the collector current is modulated is very much reduced. The linearity of this emitter-follower is still degraded by increasing loading, but to a much lesser extent; with a significant external load of 4k7, 0.036% at 4 Vrms becomes 0.006%. The steps in the bottom (no load) trace are artefacts of the AP SYS-2702 measuring system.

The noise performance of this stage is exactly as for the simple emitter-follower above.

The Push–Pull Emitter-Follower

This is an extremely useful and trouble-free form of push–pull output; I have used it many times in preamplifiers, mixers, etc. I derived the notion from the valve-technology White cathode-follower, described by Nelson-Jones in a long-ago *Wireless World* [7]. The original reference is a British patent taken out by Eric White in 1940 [8].

Figure 3.7 shows a push–pull emitter-follower. When the output is sourcing current, there is a voltage drop through the upper sensing resistor R5, so its lower end goes downwards in

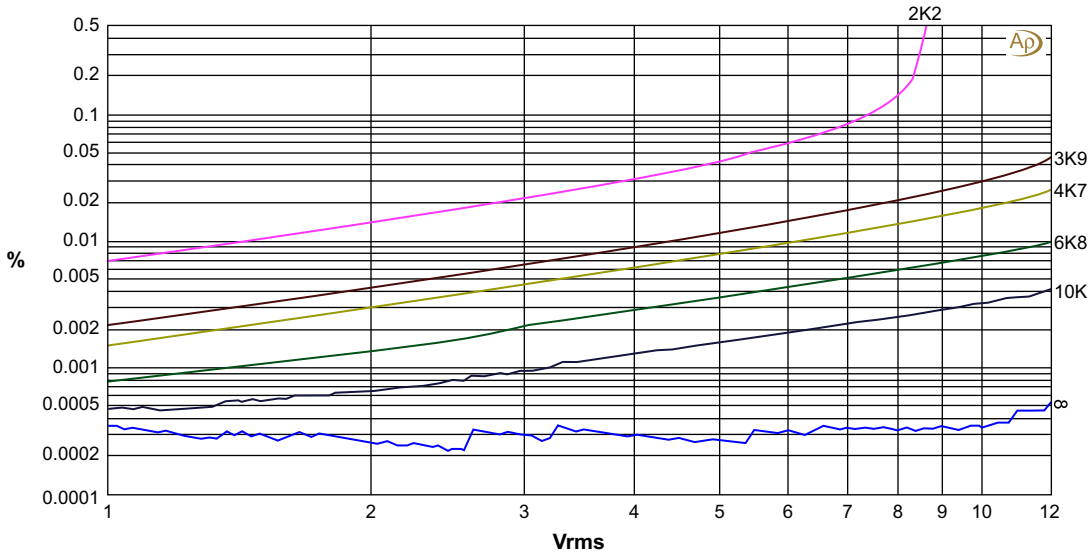


Figure 3.6: How various loads degrade the linearity of the current-source emitter-follower. The quiescent current was 6 mA

www.raguvaran.puzl.com

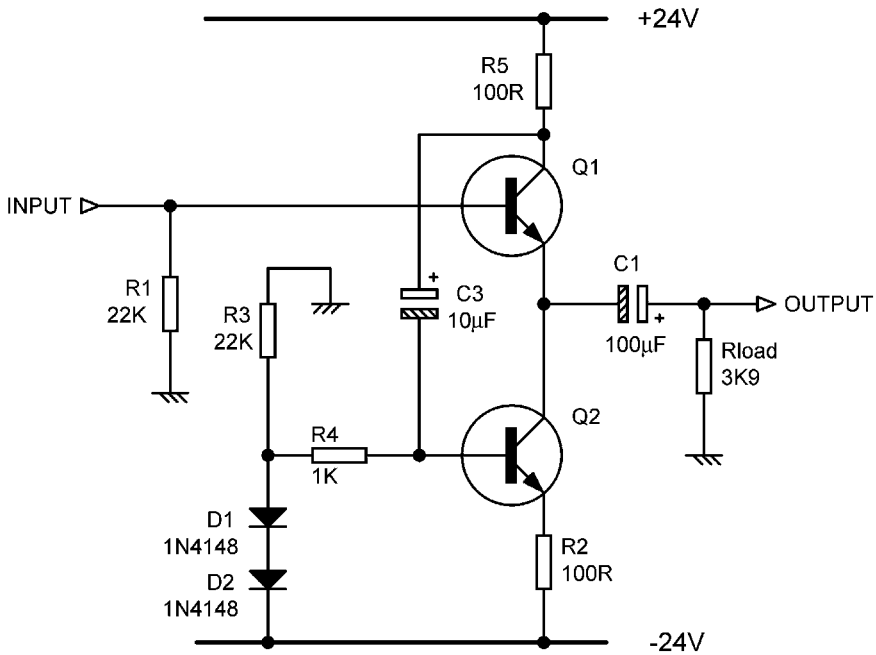


Figure 3.7: Circuit of push–pull emitter-follower. Quiescent current still 6 mA as before, but the load-driving capability is twice as great

voltage. This is coupled to the current source Q2 through C3, and tends to turn it off. Likewise, when the current through Q1 falls, Q2 is turned on more. This is essentially a negative-feedback loop with an open-loop gain of unity, and so by simple arithmetic the current variations in Q1, Q2 are halved, and this stage can sink twice the current of the constant-current version described above, while running at the same quiescent current. The effect of loading on linearity is once again considerably reduced, and only one resistor and one capacitor have been added.

This configuration needs fairly clean supply rails to work, as any upper-rail ripple or disturbance is passed directly through C3 to the current source, modulating the quiescent current and disrupting the operation of the circuit.

Push–pull action further improves the linearity of load driving; the THD with a 4k7 external load is halved, from 0.006% at 4 Vrms to 0.003%, at the same quiescent current of 6 mA, as seen in Figure 3.8. This is pretty good linearity for such simple circuitry.

Emitter-Follower Stability

The emitter-follower is about as simple as an amplifier gets, and it seems highly unlikely that it could suffer from obscure stability problems. However, it can, and often does. Emitter-followers

www.raguvaran.puzl.com

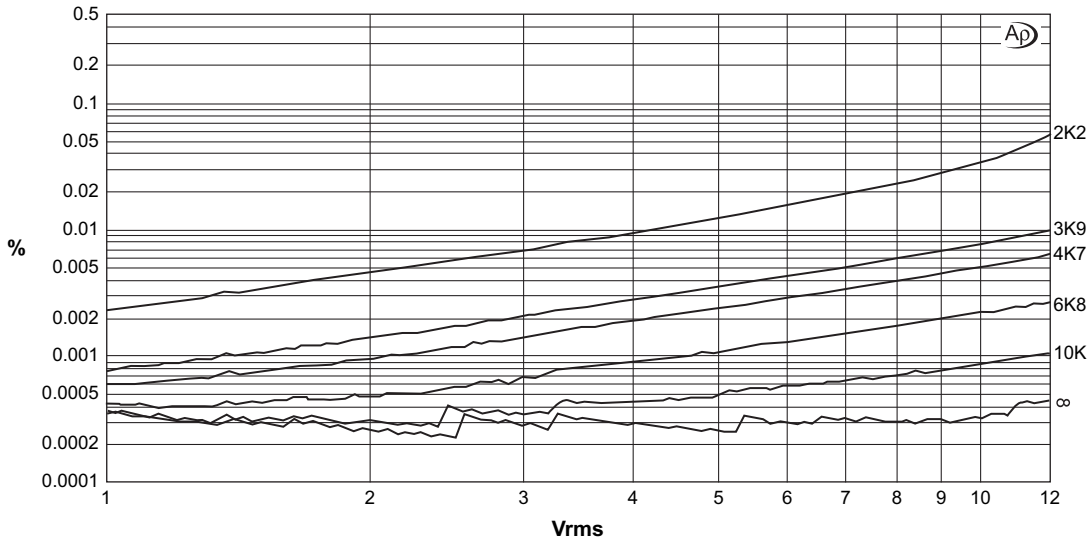


Figure 3.8: How loading degrades linearity of the push–pull emitter-follower. Loads from 10 kΩ to 2k2

are liable to RF oscillation when fed from inductive source impedances. This oscillation is in the VHF region, usually in the area 100–400 MHz, and will be quite invisible on the average oscilloscope; however, a sure sign of this problem is unusually high distortion that varies strongly when the transistor is touched with a probing finger. One way to stop this is to put a 'base-stopper' resistor directly in series with the base. This should come after the bias resistor to prevent loss of gain. Depending on the circuit conditions, the resistor may be as low as 100 Ω or as high as 2k7. The latter generates −119.6 dBu of Johnson noise, which in itself is inconvenient in low-noise circuitry, but the effects of the transistor noise current flowing through it are likely to be even worse. Base-stopper resistors are not shown in the following diagrams to aid clarity, but you should always be aware of the possible need for them. This also applies to the CFP configuration, which is equally, if not more, susceptible to the problem.

The instability is due to the fact that the typical emitter-follower is fed from a source with some inductance, and has some capacitive loading, even if it is only due to stray capacitance, as in Figure 3.9(a), where the transistor internal base-emitter capacitance C_{be} is included. If this is redrawn as in Figure 3.9(b), it is the classic circuit of a Colpitts oscillator.

For more information on this phenomenon, see Feucht [9] and de Lange [10].

CFP Emitter-Followers

The simple emitter-follower is lacking both in linearity and load-driving ability. The first shortcoming can be addressed by adding a second transistor to increase the negative-feedback

www.raguvaran.puzl.com

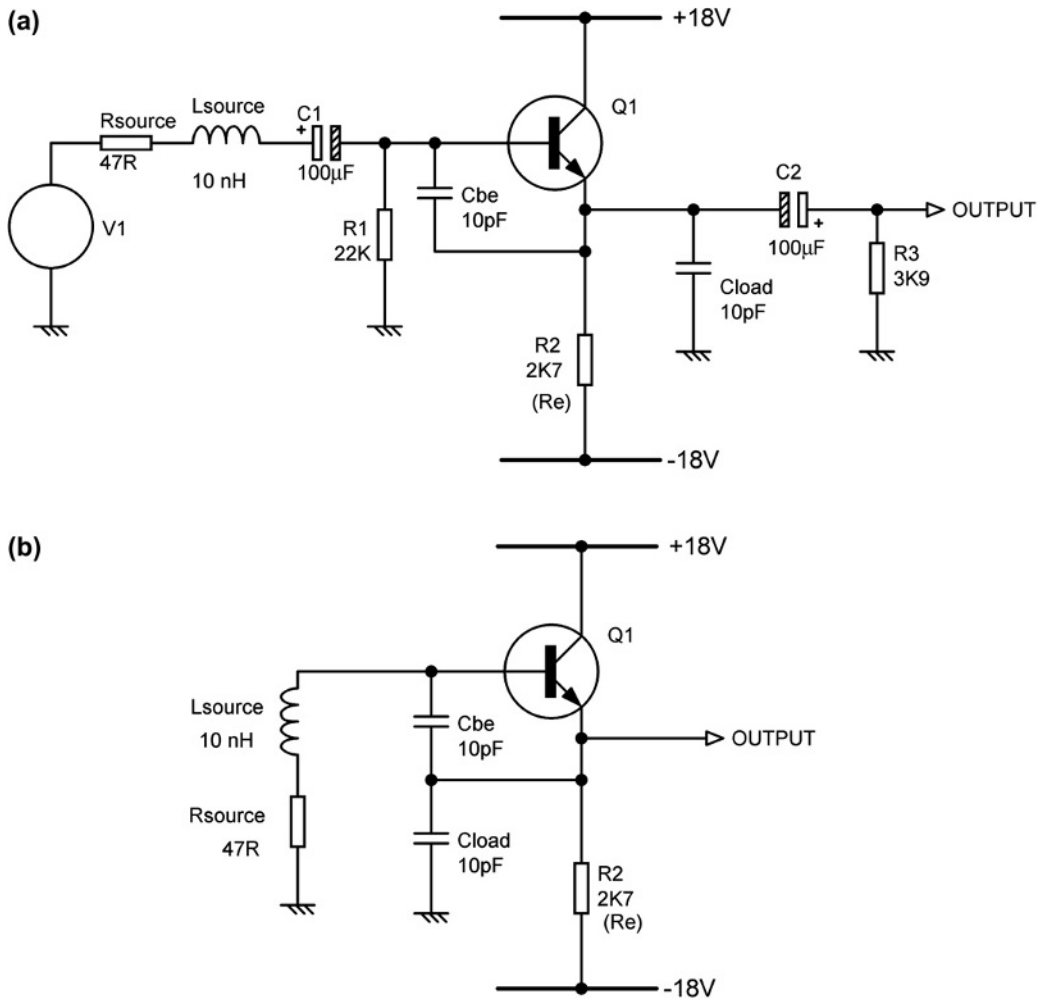


Figure 3.9: Emitter-follower oscillation: the effective circuit at (a) bears a startling resemblance to the Colpitts oscillator at (b)

factor by increasing the open-loop gain. This also allows the stage to be configured to give voltage gain, as the output and feedback point are no longer inherently the same. This arrangement is usually called the complementary feedback pair (CFP), though it is sometimes known as the Szilaki configuration. This circuit can be modified for constant-current or push–pull operation exactly as for the simple emitter-follower.

Figure 3.10 shows an example. The emitter resistor Re is the same value as in the simple emitter-follower to allow meaningful comparisons. The value of R4 is crucial to good linearity, as it sets the I_c of the first transistor, and determines its collector loading. The value of 3k3 shown here is a good compromise.

www.raguvaran.puzl.com

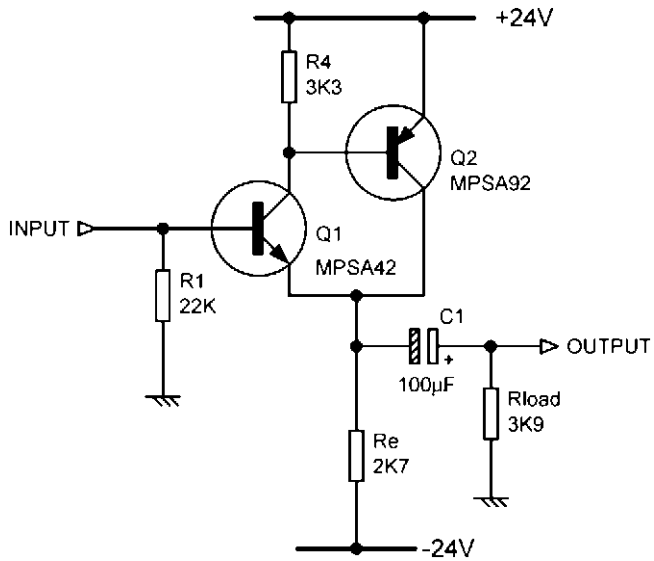


Figure 3.10: The CFP emitter-follower. The single transistor is replaced by a pair with 100% voltage feedback to the emitter of the first transistor

This circuit is also susceptible to emitter-follower oscillation, particularly if it sees some load capacitance, and will probably need a base stopper. If 1 kΩ does not do the job, try adding a series output resistor of 100 Ω close to the stage to isolate it from load capacitance.

Figure 3.11 shows the improved linearity; Figure 3.12 is the corresponding SPICE simulation. The measured noise output with a 100 Ω base stopper is −116.1 dBu.

If we replace Re with a 6 mA current source, as in Figure 3.13, we once more get improved linearity and load-driving capability, as shown in Figure 3.14. The 6 Vrms, 6k8 THD is now only just above the noise at 0.0005% (yes, three zeros after the point: three-transistor circuitry can be rather effective).

Converting the constant-current CFP to push–pull operation as in Figure 3.15 gives another improvement in linearity and load driving. Figure 3.16 shows that now only the results for 2k2 and 3k9 loading are above the measurement floor.

Improved Unity-Gain Buffers

There is often a need for a unity-gain buffer with very low distortion. If neither the simple emitter-follower, made with one transistor, nor the CFP configuration with two transistors is adequately linear, we might ponder the advantages of adding a third transistor to improve performance, without going to the complexity of the discrete op-amps described later in this chapter. (In our transistor count we are ignoring current sources used to create the active output loads that so improve linearity into significant external loading.)

www.raguvaran.puzl.com

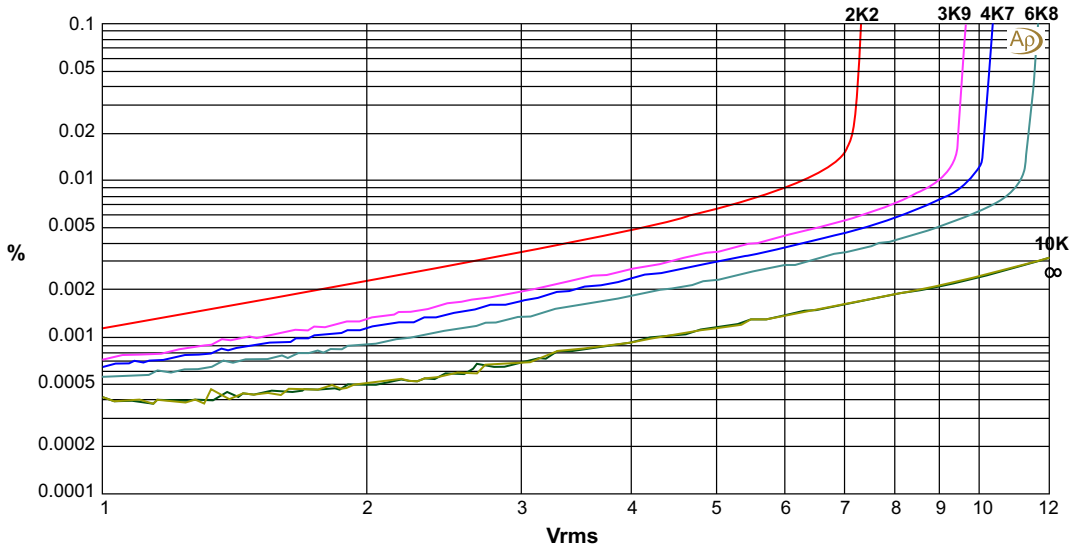


Figure 3.11: How loading affects the distortion of a CFP emitter-follower. THD at 6 Vrms, 6k8 load is only 0.003% compared with 0.05% for the simple EF. Re is 2k7

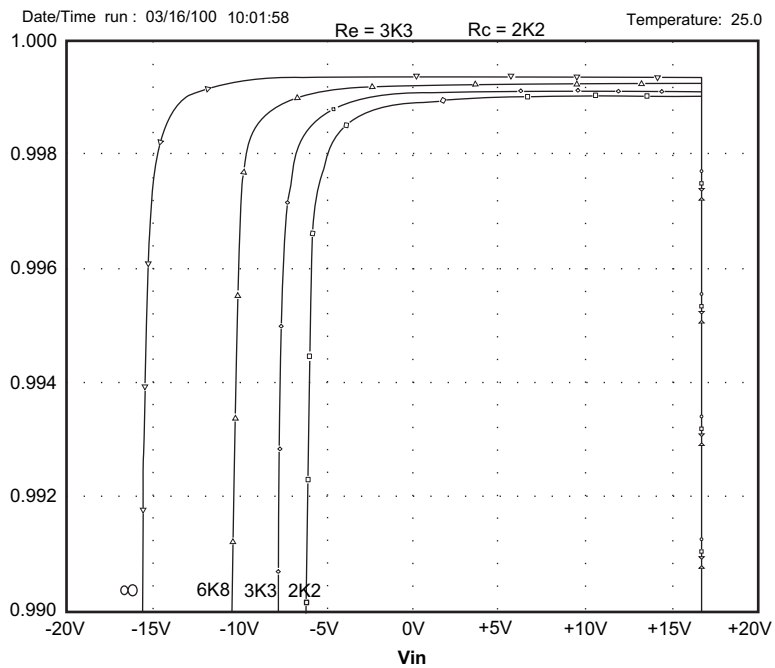


Figure 3.12: SPICE simulation of the circuit in Figure 3.10, for different load resistances. The curves are much flatter than those in Figure 3.4, even though the vertical scale has been expanded

www.raguvaran.puzl.com

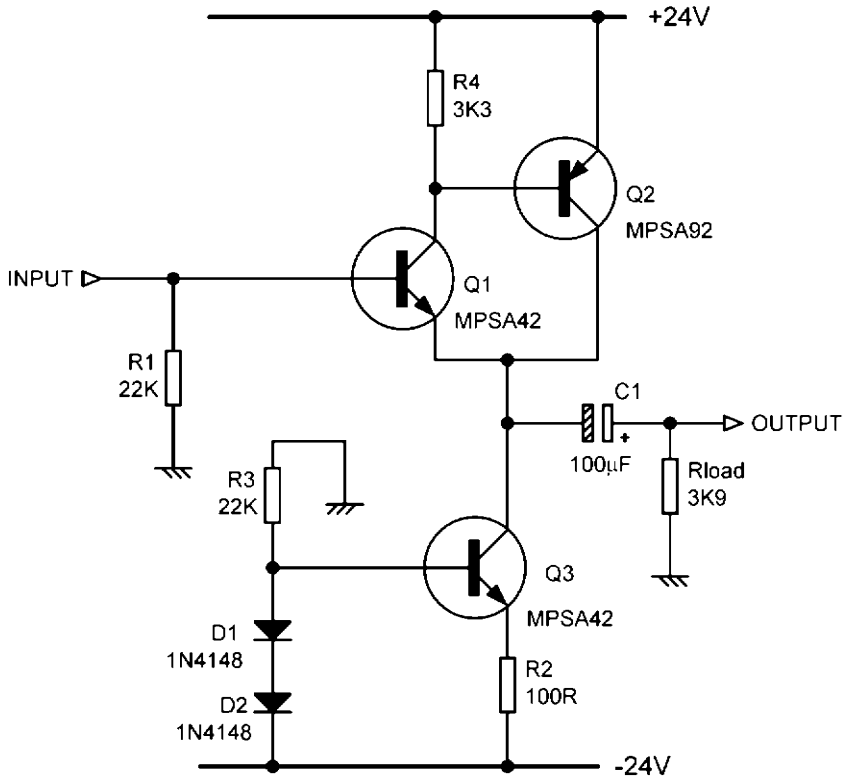


Figure 3.13: Constant-current CFP follower. Once more the resistive emitter load is replaced by a constant-current source to improve current sinking

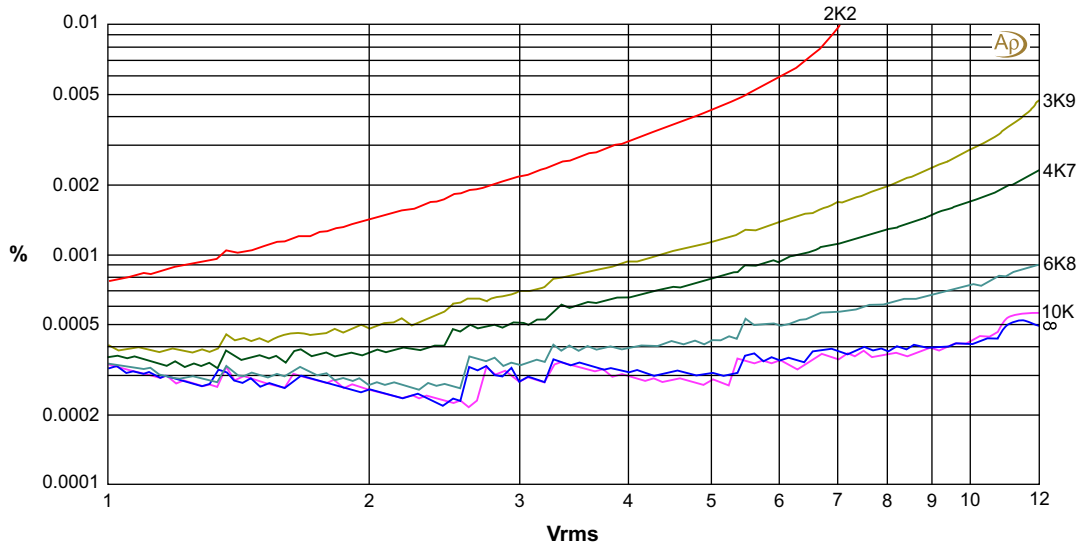


Figure 3.14: Distortion and loading effects on the CFP emitter-follower with a 6 mA current source. The steps on the lower traces are artefacts caused by the measurement system gain-ranging as it attempts to measure the THD of pure noise. Note change of scale

www.raguvaran.puzl.com

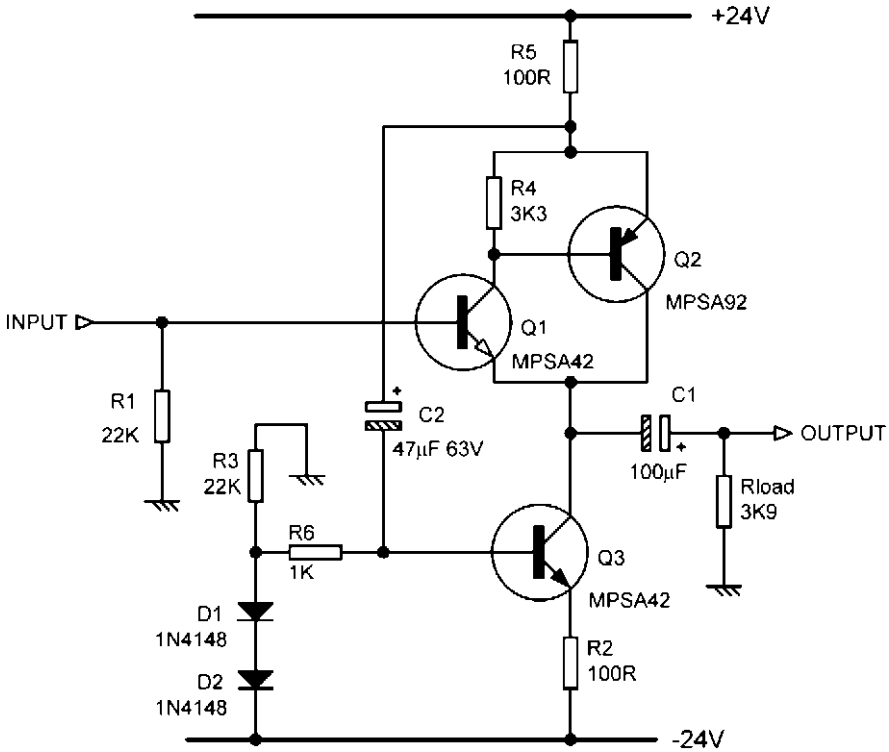


Figure 3.15: Circuit of a push–pull CFP follower. This version once more gives twice the load-driving capability for no increase in standing current

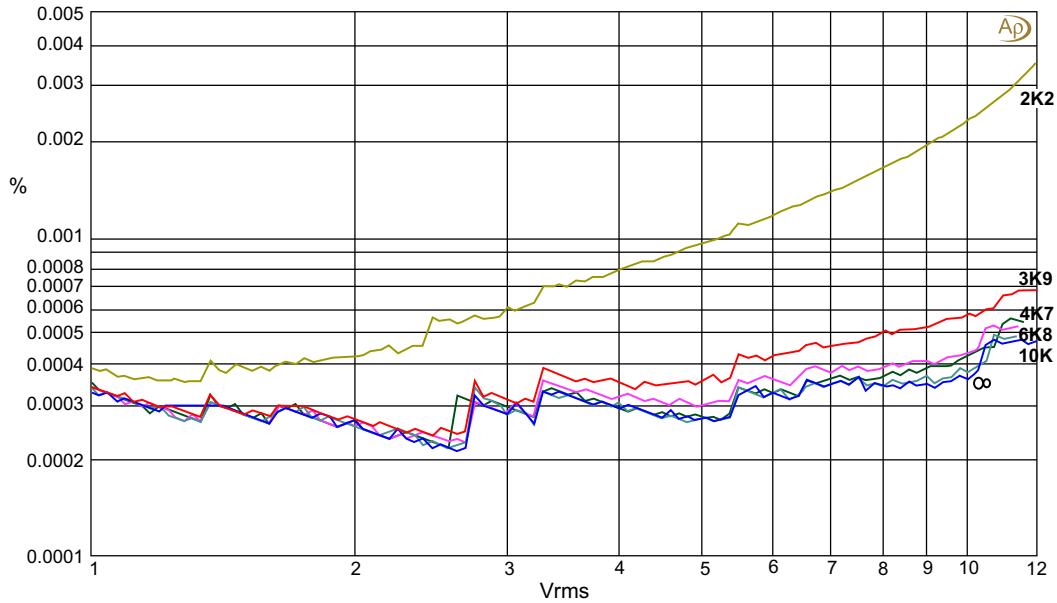


Figure 3.16: Distortion and loading effects on the push–pull CFP emitter-follower, still with 6 mA

www.rasuwaran.puzl.com

One promising next step is a three-transistor configuration that is often called the Schlotzaur configuration (see Feucht [5] and Staric and Margan [11]). The single input transistor in the CFP emitter-follower is now replaced with a long-tail pair with 100% feedback to Q2, as in Figure 3.17(a). Because long-tail pair have the property of canceling out their even-order distortion [6], you might expect a considerable improvement. You would be wrong; running on ±15 V rails we get 0.014% at 6 Vrms unloaded, almost flat across the audio band. The linearity is much inferior to the CFP emitter-follower. The open-loop gain, determined by measuring the error voltage between Q1, Q2 bases, is 249 times.

It's not a promising start, but we will persist! Replace the collector load R2 with a current source of half the value of the tail current source as in Figure 3.17(b) and the linearity is transformed, yielding 0.00075% at 6 Vrms unloaded on ±15 V rails. Increasing the rails to ±18 V gives 0.00066% at 6 Vrms; and a further increase to ±24 V, as high-voltage operation is part of what discrete design is all about, gives 0.00042% at 6 Vrms (unloaded). It's the increase in the positive rail that gives the improvement. Reducing the measurement bandwidth from 80 to 22 kHz for a 1 kHz signal eliminates some noise and gives a truer figure of 0.00032% at 6 Vrms. The open-loop gain is increased to 3400 times.

Adding loading to the buffer actually has very little effect on the linearity, but its output capability is clearly limited by the use of R3 to sink current, just as for a simple emitter-follower. Replacing R3 with a 6 mA constant-current source, as for previous circuits, much improves drive capability and also improves linearity somewhat (see Figure 3.17(c)). On ±30 V rails we get a reading of 0.00019% at 5 Vrms with a 2k2 load, and that is mostly the noise in a 22 kHz bandwidth.

Finally we remove current source I2 and replace it with a simple current mirror in the input pair collectors, as in Figure 3.17(d), in the pious hope that the open-loop gain will be doubled and distortion halved. On ±30 V rails there is a drop in THD from 0.00019% to 0.00017% at 5 Vrms with a 2k2 load (22 kHz bandwidth), but that is almost all noise and I am pushing the limits of even the magnificent Audio Precision SYS-2702. This goes to show that there are other ways of designing low-distortion circuitry apart from hefting a bucket of 5532s.

Figure 3.18 shows the distortion plot at 5 Vrms. Using an 80 kHz bandwidth so the HF end is meaningful means the readings are higher, and virtually all noise below 10 kHz. It also gives another illustration of the distortion generated by undersized coupling capacitors. C1 started as 22 μF, but you can see that 220 μF is required to eliminate distortion at 10 Hz with a 2k2 load.

SPICE analysis shows that the collector currents of the input pair Q1, Q2 are somewhat unbalanced by the familiar base-current errors of a simple current mirror. Replacing the simple current mirror with the well-known Wilson improved mirror might get us further improvement, if we could measure it. Work in progress …

The circuit we have now could still be regarded as a much-enhanced emitter-follower, but it is probably more realistic to consider it as a two-stage discrete op-amp.

www.raguvaran.puzl.com

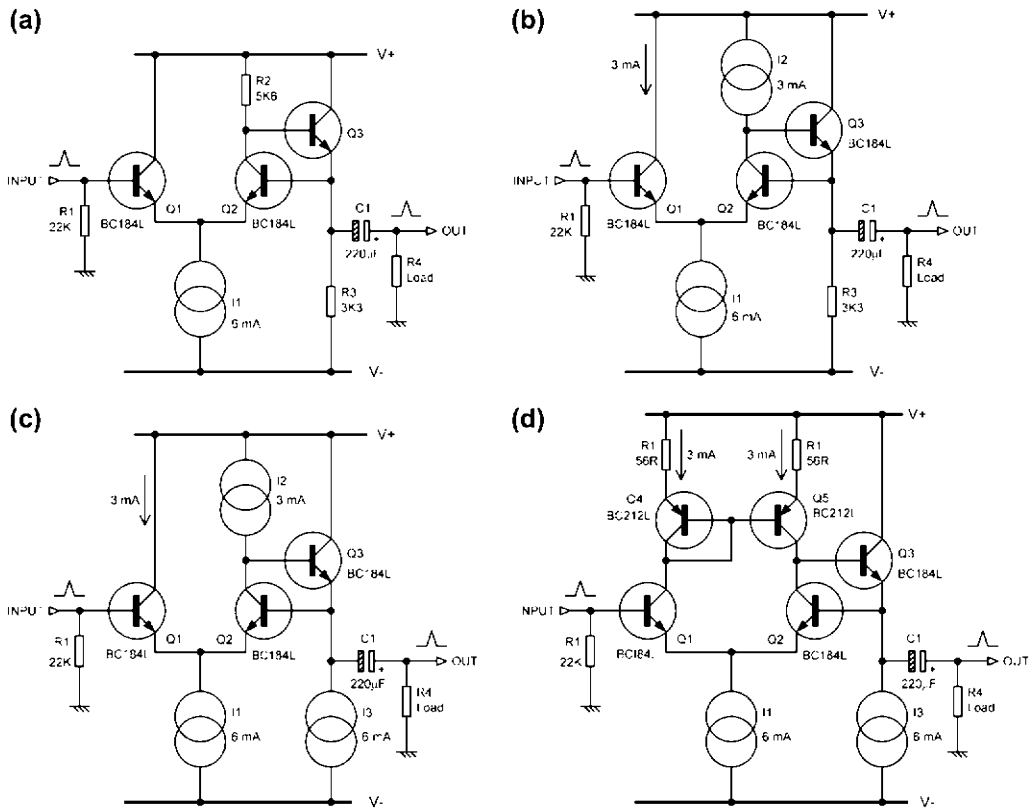


Figure 3.17: Developing a unity-gain buffer design, which replaces the single input transistor with a long-tailed pair. (a) Simple Schlotzaur circuit. (b) Collector load R2 replaced with current source. (c) Output current source added. (d) Inserting current mirror in input pair collectors

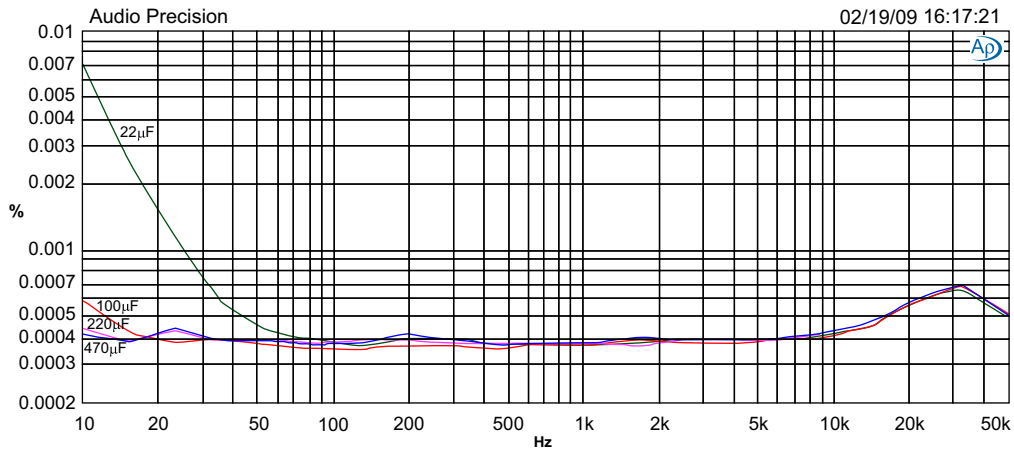


Figure 3.18: THD plot for the final version of the buffer, at 5 Vrms with a 2k2 load. The rising LF curves illustrate the distortion generated by undersized output capacitors. Bandwidth 80 kHz

www.raguvaran.puzl.com

Gain Stages

This section covers any discrete transistor stage that can give voltage gain. It may be as simple as a single transistor or as complex as an op-amp implemented with discrete components. A single transistor can give voltage gain in either series or shunt mode.

One-Transistor Shunt-Feedback Gain Stages

Single-transistor shunt-feedback gain stages are inherently inverting, and of very poor linearity by modern standards. The circuit in Figure 3.19 is inevitably a collection of compromises. The collector resistor R4 should be high in value to maximize the open-loop gain, but this reduces the collector current of Q1, and thus its transconductance, and hence reduces open-loop gain once more. The collector resistor must also be reasonably low in value as the collector must drive external loads directly. Resistor R2, in conjunction with R3, sets the operating conditions. This stage has only a modest amount of shunt feedback via R3, and the input can hardly be called a virtual earth. However, such circuits were once very common in low-end discrete preamplifiers, back in the days when the cost of an active device was a serious matter. Such stages are still occasionally found doing humble jobs like driving VU meters, but the cost advantage over an op-amp section is small, if it exists at all.

The gain of the stage in Figure 3.19 is at a first look 220k/68k = 3.23 times, but the actual gain is 2.3× (with no load) due to the small amount of open-loop gain available. The mediocre distortion performance that must be expected even with this low gain is shown in Figure 3.20.

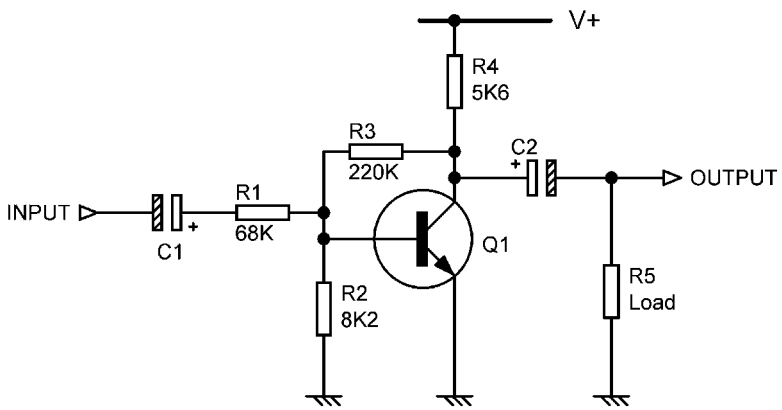


Figure 3.19: Circuit of single-transistor gain stage, shunt-feedback version

www.raguvaran.puzl.com

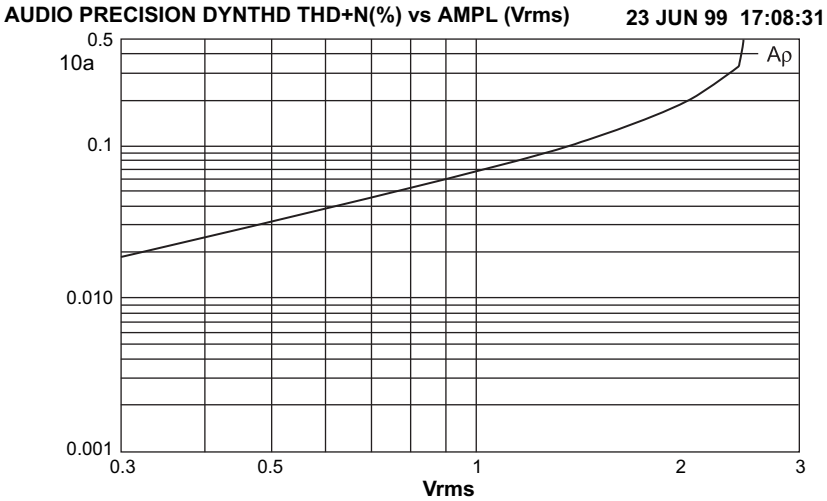


Figure 3.20: Single-transistor shunt-feedback gain stage, distortion versus level. Gain is 2.3 times

One-Transistor Series-Feedback Gain Stages

Single-transistor series-feedback gain stages are made by creating a common-emitter amplifier with a feedback resistor that gives series voltage feedback to the emitter. The gain is the ratio of the collector and emitter resistors, if loading is negligible. Note that unlike op-amp-based series-feedback stages, this one is inherently inverting. To make a non-inverting stage with a gain more than unity requires at least two transistors, because of the inversion in a single common-emitter stage.

This very simple stage, shown in Figure 3.21, naturally has disadvantages. The output impedance is high, being essentially the value of the collector resistor R3. The output is neither good at sinking nor sourcing current.

The distortion performance as seen in Figure 3.22 is indifferent, giving 0.3% THD at 1 Vrms out. Compare this with the shunt-feedback one-transistor gain stage above, which gives 0.07% under similar conditions.

Two-Transistor Shunt-Feedback Gain Stages

Before the advent of the op-amp, inverting stages were required for tone controls and virtual-earth summing amplifiers. The one-transistor amplifier stage already described is deficient in distortion and load-driving capability. See Figure 3.24 for distortion performance. A much better amplifier can be made with two transistors, as in Figure 3.23. The voltage gain is generated by Q1, which has a much higher collector resistor R4 and so much higher gain. This is possible because Q2 buffers it from external loading, and allows a higher NFB factor.

www.raguvaran.puzl.com

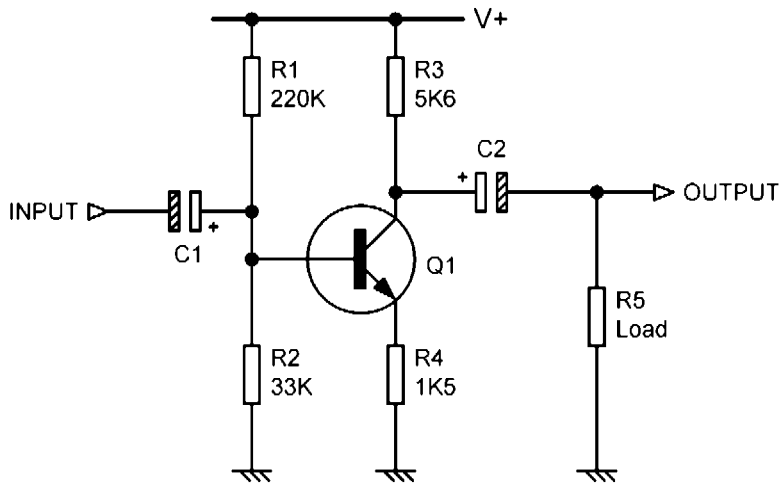


Figure 3.21: Circuit of single-transistor series-feedback gain stage

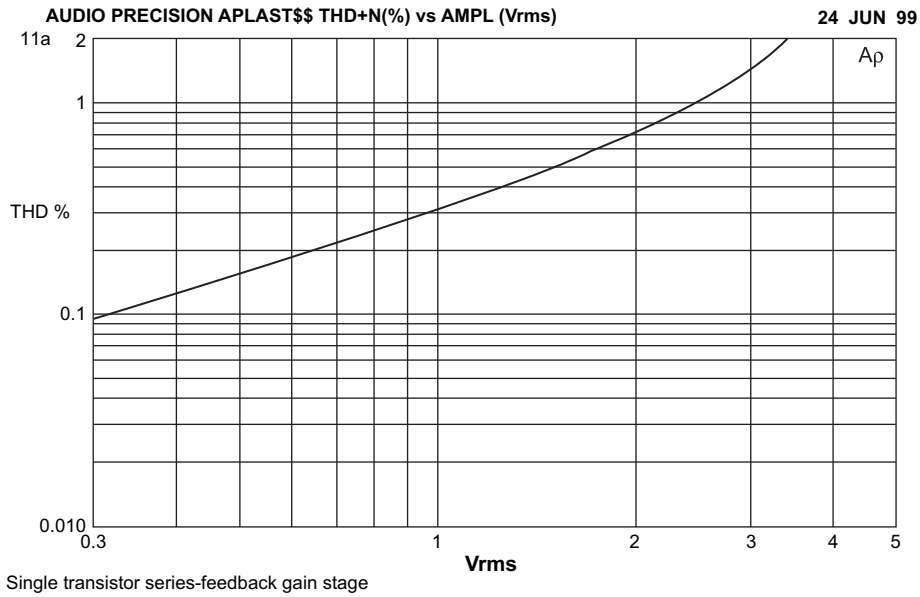


Figure 3.22: Single-transistor series-feedback gain stage, distortion versus input level. Gain is 3

With the addition of bootstrapping, as shown in Figure 3.25, the two-transistor stage has its performance transformed. Figure 3.26 shows how THD is reduced by a factor of 10; 0.15% at 1 Vrms in, 3 Vrms out becomes 0.015%, which is much more respectable. The improvement is due to the increased voltage gain of the first stage, giving a higher NFB factor. THD is still approximately proportional to level, as the distortion products are mainly second harmonic.

www.raguvaran.puzl.com

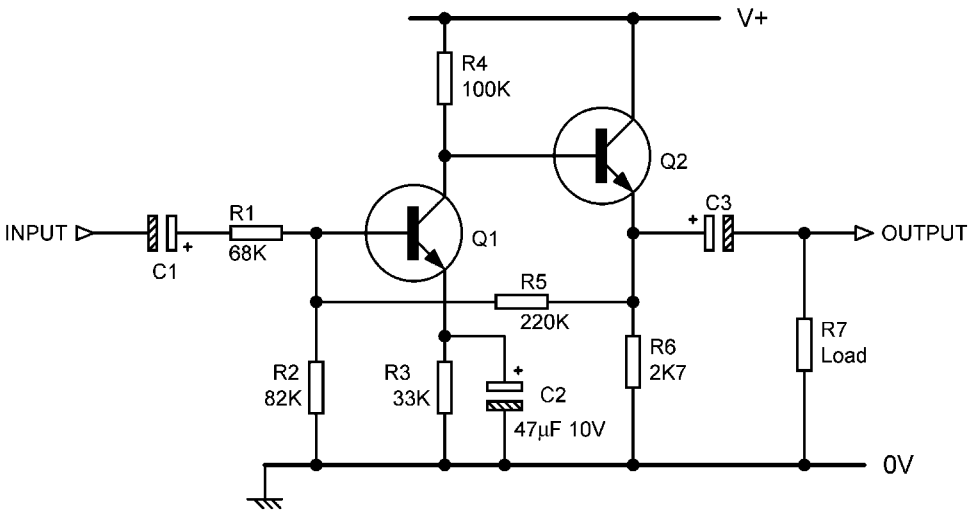


Figure 3.23: Two-transistor gain stage with shunt-feedback. Gain is 3

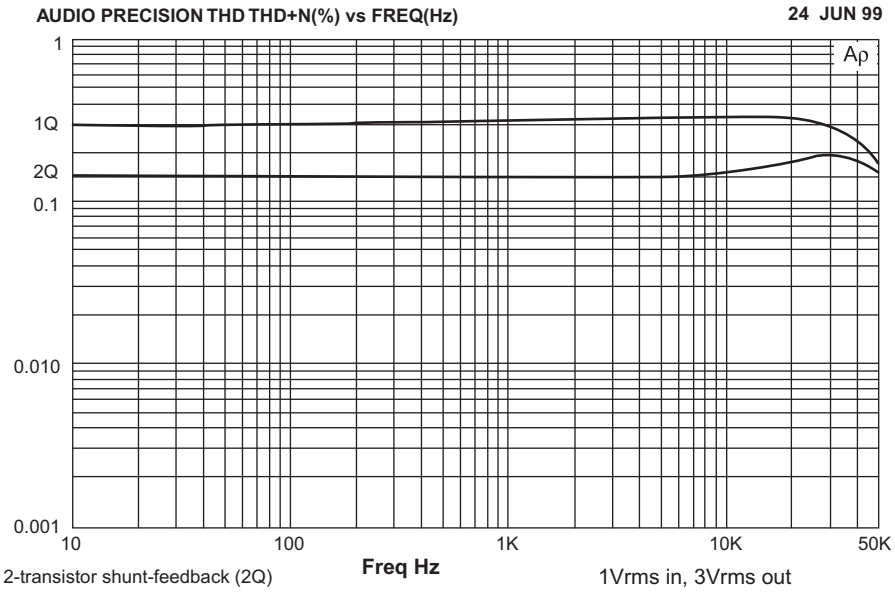


Figure 3.24: Distortion of two-transistor shunt stage versus frequency (2Q). Distortion of the one-transistor version (1Q) is also shown. The two-transistor version is only twice as good, which seems a poor return for the extra active device

Clipping occurs abruptly at 2.9 Vrms in, 8.7 Vrms out; abrupt clipping onset is characteristic of stages with a high NFB factor.

Stages like this were commonly used as virtual-earth summing amplifiers in mixing consoles before acceptable op-amps were available at reasonable cost.

www.raguvaran.puzl.com

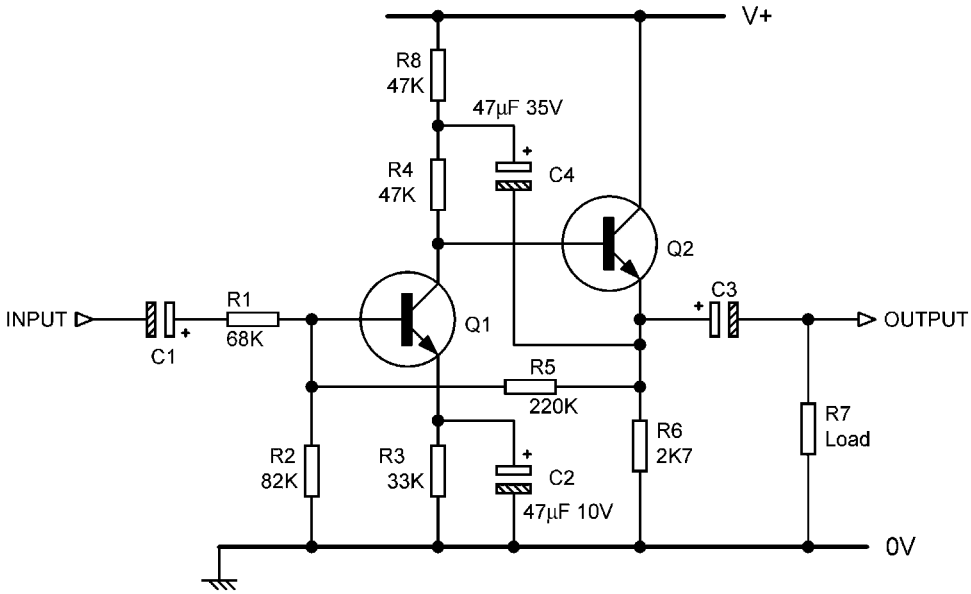


Figure 3.25: Two-transistor shunt-feedback stage, with bootstrapping added to the first stage to improve linearity

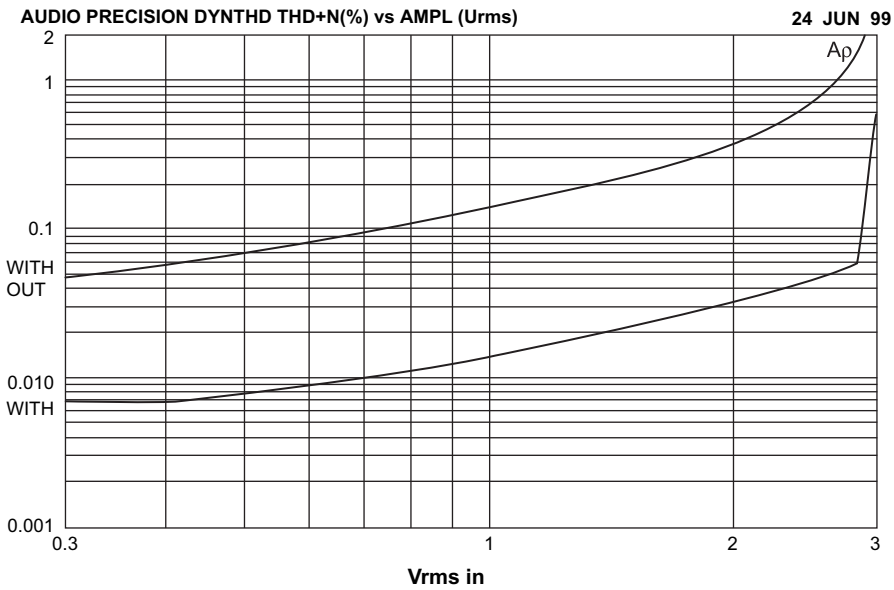


Figure 3.26: Distortion of the two-transistor shunt stage versus level, with and without bootstrapping. *X*-axis is input level; output is three times this

www.raguvaran.puzl.com

Figure 3.27 shows that the distortion is further reduced to 0.002% at 3 Vrms out if the impedance of the input and feedback networks is reduced by 10 times. SPICE simulation confirms that this is because the signal currents flowing in R1 and R5 are now larger compared with the non-linear currents drawn by the base of Q1. There is also a noise advantage.

Increasing the supply reduces the distortion for a given output level in this kind of stage. A version with the low-impedance feedback values gave 0.0044% at 1 kHz for a 3 Vrms output with a +24V rail; increasing the supply voltage to +30V reduced this to 0.0032%, while +40V gave 0.0023%. The biasing system adjusts itself to these changes. A creditable performance for such a simple circuit.

Distortion can also be improved by tweaking the value of the emitter-follower load R6, to partially cancel the curvatures of the first and second stages. With a +24V rail and R6 at 2k7 as shown, THD is 0.0044% at 3 Vrms out, but reducing R6 to 2k2 drops the distortion to 0.0038%, while 2k0 gives 0.0035%. The latter figure is 79% lower, the only cost being a slight increase in power consumption.

Another important point is that the bootstrap capacitor needs to be larger than you might think. Reducing it to 2.2 μF here gives rising distortion at LF, with a ten-fold increase at 10 Hz. This is because the gain increase given by bootstrapping is very sensitive to the AC voltage at the top of the capacitor.

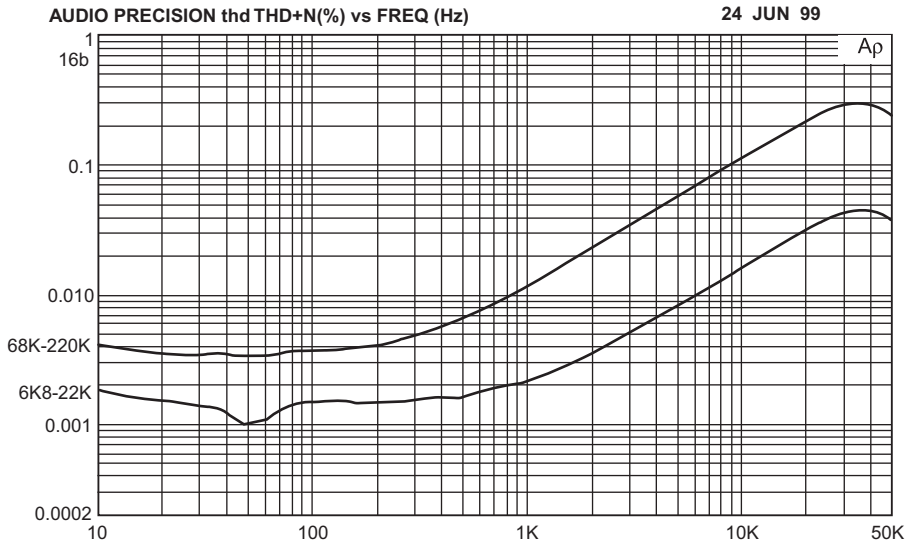


Figure 3.27: The distortion is reduced by a further factor of at least 5 if the impedance of the input and feedback networks is reduced by 10 times. Output 3 Vrms

www.raguvaran.puzl.com

Figure 3.28 shows how output drive capability can be increased, as before, by replacing R6 with a 6 mA current source. The input and feedback resistors R1, R5 have again been scaled down by a factor of 10. Push–pull operation can also be simply implemented as before. The EIN of this version is -116 dBu.

The constant-current and push–pull options can also be added to more complex discrete stages. Some good examples can be found in a preamplifier design of mine [12].

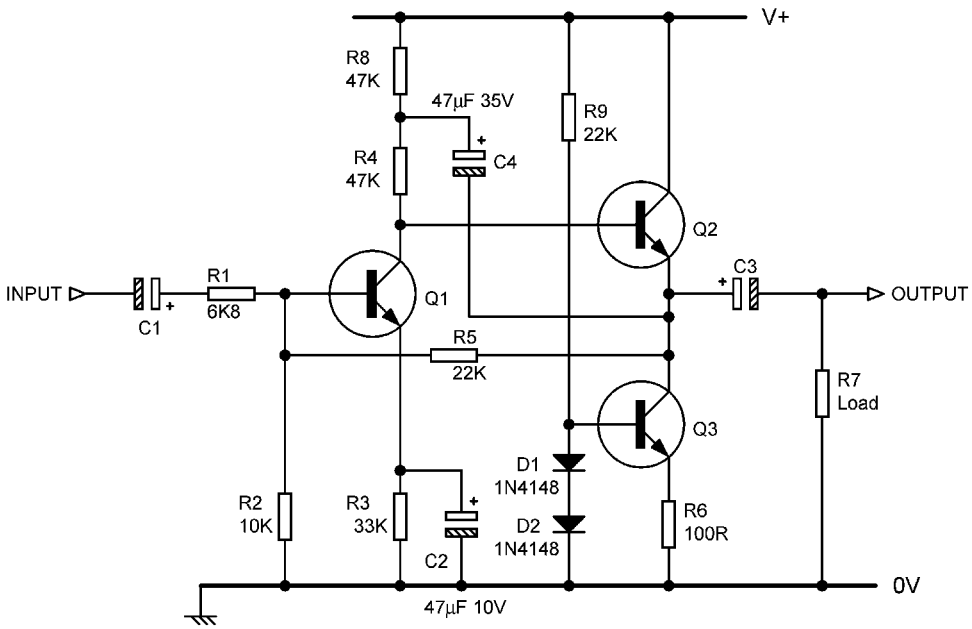


Figure 3.28: Two-transistor bootstrapped shunt-feedback stage, with low-impedance feedback network, and with current-source output to enhance load-driving capability

Two-Transistor Series-Feedback Gain Stages

This circuit clearly has a close family relationship with the CFP emitter-follower, which is simply one of these stages configured for unity gain. The crucial difference here is that the output is separated from the input emitter, so the closed-loop gain is set by the R3–R2 divider ratio. See Figure 3.29.

Only limited NFB is available, so closed-loop gains of two or three times are usually the limit. It is less easy to adapt this circuit to improve load-driving capability, because the feedback resistive divider must be retained. Figure 3.30 shows that distortion performance depends strongly on the value of R4, the collector load for Q1. The optimal value for linearity and noise performance is around 4k7.

www.raguvaran.puzl.com

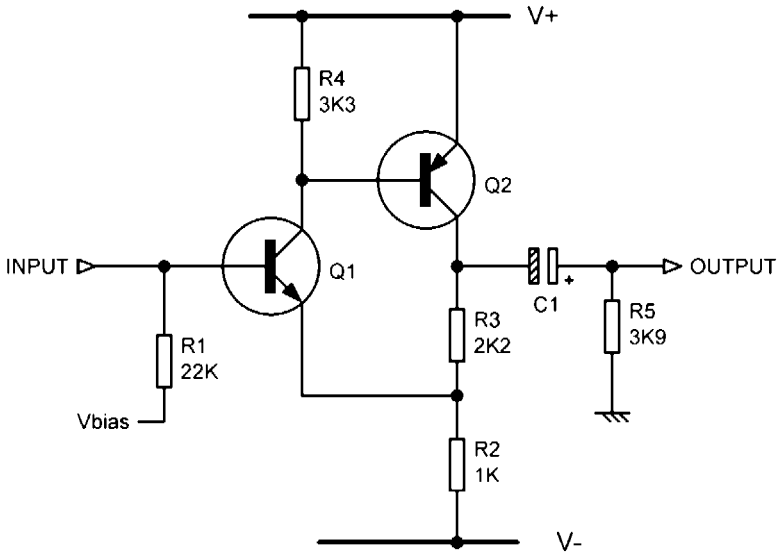


Figure 3.29: Two-transistor gain stage, series feedback. Gain once more is approximately 3

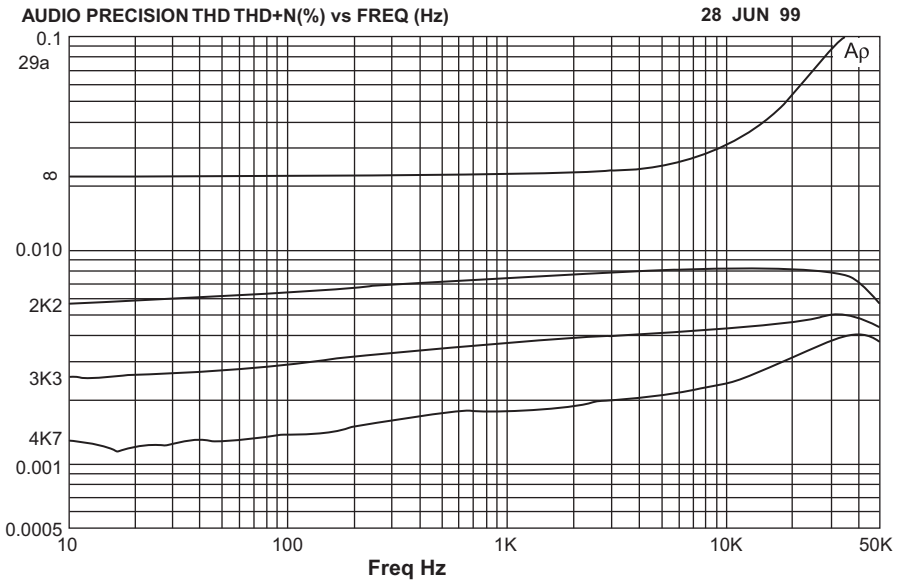


Figure 3.30: Two-transistor series-feedback stage. THD varies strongly with value of R4

Discrete Op-Amp Design

When the previously described circuits do not show enough linearity or precision for the job in hand, or a true differential input is required, it may be time to turn to op-amp design using discrete transistors. This section does not pretend to be an exhaustive guide, but

www.raguvaran.puzl.com

concentrates on audio issues such as distortion. Much that can be found in standard op-amp textbooks is taken for granted. It does, however, give information that I do not think appears anywhere else.

The op-amp in Figure 3.31 is conventional in design; its resemblance to a power amplifier circuit is obvious. It consists of three stages: the differential input stage Q1–Q3, the voltage-amplifier stage (VAS) Q4, Q5, and a unity-gain output stage Q7, Q8.

The long-tailed pair first stage Q2, Q3 subtracts the input and feedback voltages; it is a transconductance amplifier, i.e. it turns a voltage input into a current output. The voltage/current gain curve of this stage peaks at the center, where the collector currents of Q2, Q3 are equal, and is its most linear part. As a result the value of R2 is crucial; if the I_c values are unbalanced there is a wholly avoidable rise in second-harmonic THD at high frequencies. Since global NFB establishes 0.6 V across R2, simply making it equal to R3 does not do the job; it needs to be approximately equal to half the value of the resistor R1 that sets the tail current. This is demonstrated in Figure 3.32.

This rapid rise in first-stage distortion with frequency is because:

- at high frequencies more current flows in and out of the dominant-pole capacitor C3. At the input side of the capacitor this can only be provided by the input stage, which has to work harder – in other words, the error voltage is greater so the output current has a larger swing over more of the transfer characteristic;

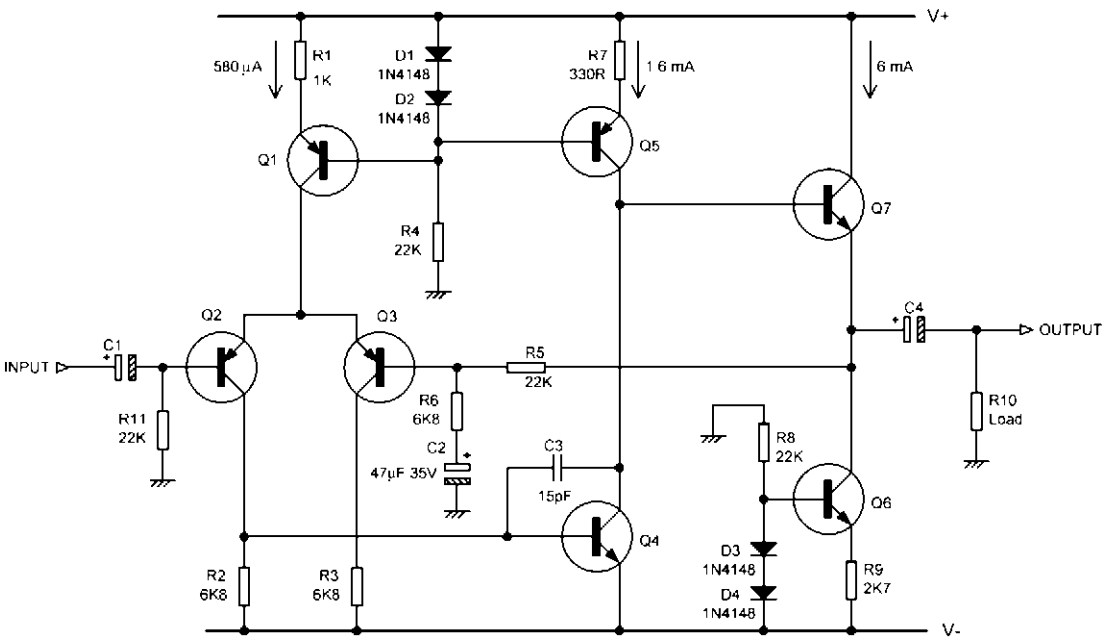


Figure 3.31: A typical discrete op-amp circuit, with current-source output

www.raguvaran.puzl.com

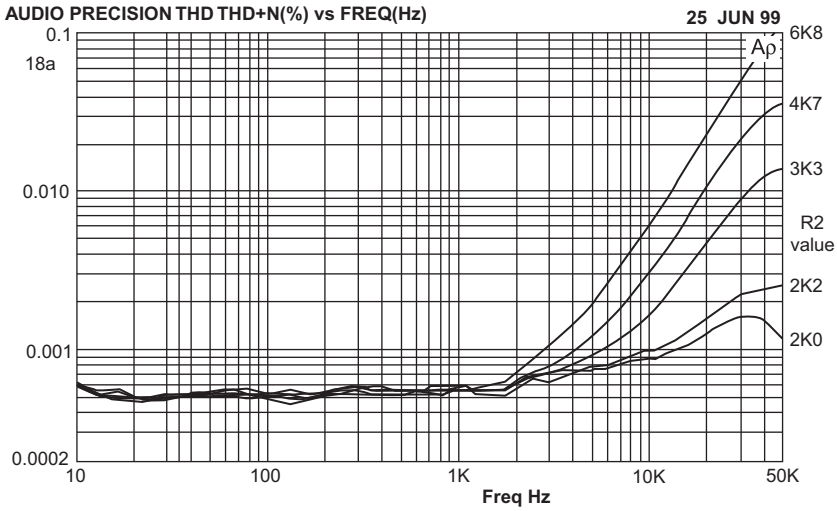


Figure 3.32: Reduction in high-frequency THD as input stage approaches balance with R2 = 2 kΩ

- once the input stage is balanced, it creates only third-harmonic distortion, which rises as the square of LTP input voltage;
- there is less global NFB at high frequencies due to C3.

The first-stage output current flows into the second voltage-amplifier stage Q4, Q5. This is a transadmittance amplifier; current in becomes voltage out. The input is at low impedance (a sort of virtual earth) because of local NFB through dominant-pole capacitor C3, and so is well adapted to accepting the output current from the first stage.

The third stage is a unity-gain buffer, isolating the VAS collector from external loads. The higher the quiescent current in this stage (here 6 mA as before), the lower the load impedance that can be driven symmetrically.

An improved circuit is shown in Figure 3.33. R2 has been given its correct value; R3 has now been altered to match but this change is purely cosmetic. The distortion performance is shown in Figure 3.34. Much more information on LTP input stage design can be found in Chapter 4 of Ref. [6].

The operating conditions of the op-amp are set by the three currents shown in the figures. The tail current of the first stage is probably the most critical, as this sets:

- the transconductance of the input pair. This in turn affects the open-loop gain of the amplifier and therefore stability when the loop is closed;
- the maximum output slew rate;

www.raguvaran.puzl.com

- the noise performance. This is normally determined by a combination of the source resistance (of the external source that is driving the amplifier) and the collector current. The lower the source resistance, the higher the I_c required for minimal noise;

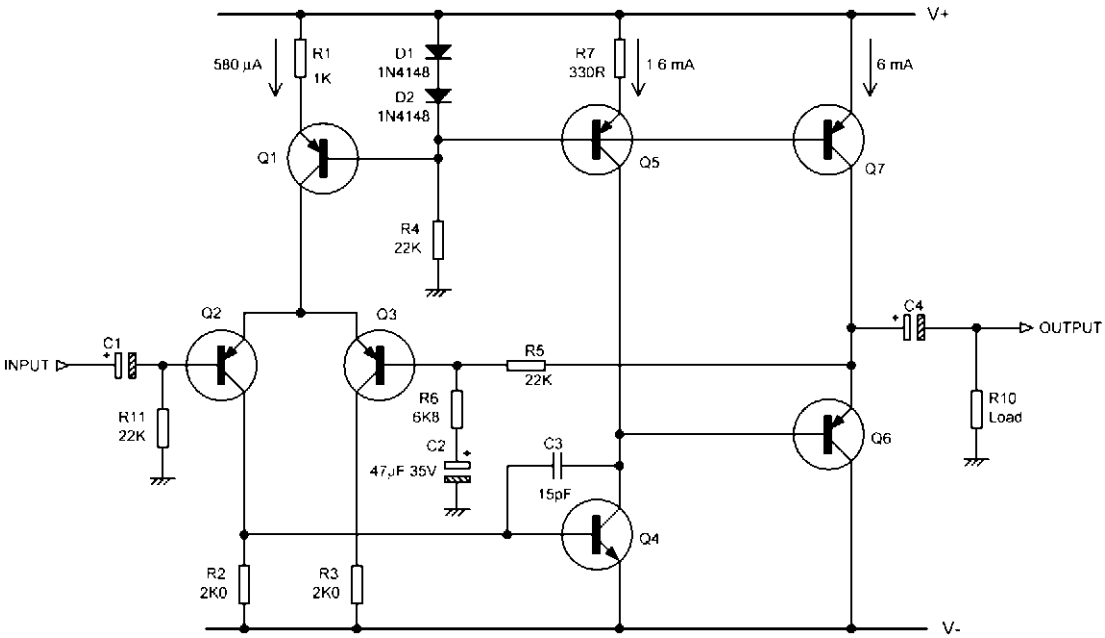


Figure 3.33: An improved discrete op-amp, with the input stage balanced and the output stage inverted to share the biasing diodes, and so reduce component count

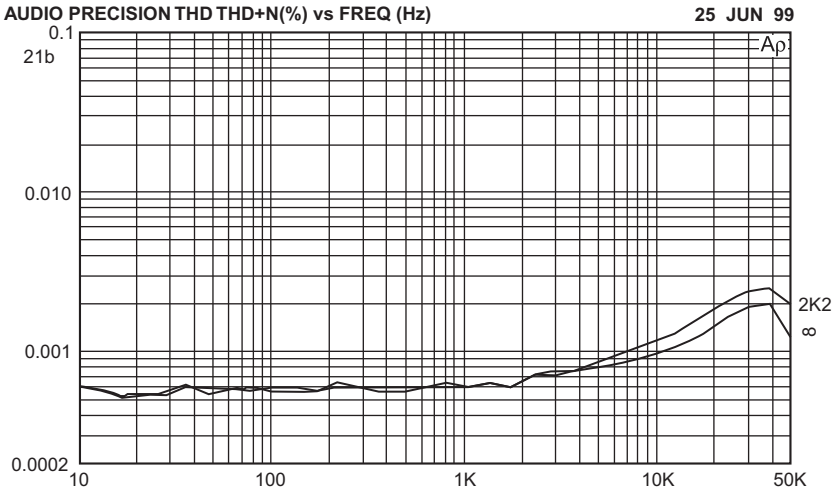


Figure 3.34: THD of the improved op-amp circuit, showing that the effect of a 2k2 load on linearity is small. Output level 8 Vrms

www.raguvaran.puzl.com

- the input bias currents, and hence DC accuracy in some circumstances. Input transistor beta also has a major influence on this.

A good example of discrete op-amp usage is a preamplifier I designed a while back [13].

High-Input-Impedance Bipolar Stages

Transistors are commonly thought of as low-impedance devices, but in fact they can be used to create amplifiers with extremely high input impedances. FETs are of course the obvious choice for high-input-impedance amplifiers, but their lack of transconductance is a drawback. This section was inspired by an article by T.D. Towers in 1968 that is still well worth reading [14]. The circuits that follow deliberately use high-voltage, low-beta transistors, as higher than normal rails are one of the reasons for using discrete circuitry – ±24V rails used. The impedances given were measured and also checked with SPICE.

Figure 3.35(a) shows a simple emitter-follower biased by R_{bias}, which at 100 kΩ is about as high as you would want it to be, because of the voltage drop due to the base current. That limits the input impedance to 100 kΩ of course. The First Principle of high-impedance design is to bootstrap the bias resistor as in Figure 3.35(a), which raises the input impedance to 500 kΩ. You may be wondering: why not more? One reason is that Q1 is only a simple emitter-follower and its voltage gain is distinctly less than 1, limiting the efficacy of the bootstrapping.

The Second Principle of high-impedance design with simple discrete stages is that any loading on the output reduces the input impedance because, as noted earlier:

$$R_{in} = \beta(R_e \| R_{load}) \qquad \text{(Equation 3.3)}$$

and the bootstrap capacitor C2 is driving the load of R1 even if there is no external loading. Adding an external 10 kΩ R_{load} to represent the following stage reduces the input impedance to 350 kΩ.

Having noted that both R_e and R_{load} pull down the input impedance, we will take steps to increase their effective values. Figure 3.35(b) shows R_e made very high by replacing the emitter resistor with a current source Q2. R_{load} is made high by adding simple emitter-follower Q3 to drive the bootstrap and the output. This gives a 6.1 MΩ input impedance even when driving an external 10 kΩ load output.

A further increase in input impedance can be obtained by increasing the β term in Equation 3.3, by using a Darlington configuration, where one emitter-follower feeds another, as in Figure 3.35(c). Q1 must have a reasonable collector current to operate at a good β – more than the base current of Q2. An emitter resistor to the negative rail would increase the loading and defeat the object, so R4 is used, with its lower end bootstrapped from the emitter of Q2. This

www.raguvaran.puzl.com

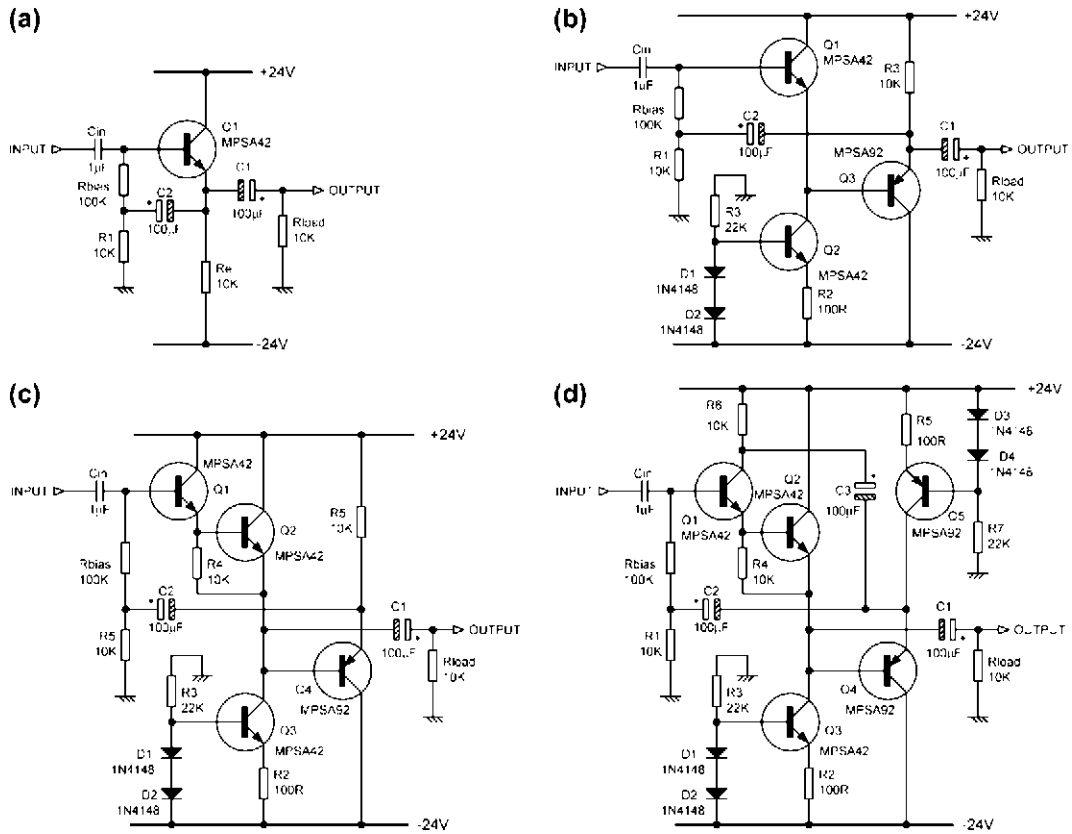


Figure 3.35: High-impedance input stages. (a) Simple emitter-follower with bootstrapped biasing. (b) Emitter-follower with current source and bootstrap driver stage. (c) Darlington with current source and bootstrap driver stage. (d) Darlington with current source and Q1 collector bootstrapped

technique is also used in power amplifier circuitry [15]. This gives an input impedance of 21 MΩ with an external 10 kΩ load. The output is now taken from the Q2 emitter once more.

To significantly further raise the input impedance, we need to take on board the Third Principle of high-impedance design: bootstrap the input transistor collector, as in Figure 3.35(d). The collector resistance r_c of Q1 and the base-collector capacitance c_{bc} are both effectively in parallel with the input; the former can also be regarded as Early effect causing I_b to vary. Their effects are reduced by bootstrapping Q1 collector using R6 and C3. (Note that collector bootstrapping cannot be used with a single-transistor stage [16].) The emitter-follower Q4 also has its emitter resistor replaced by current source Q5 to make its gain nearer 1 for more effective bootstrapping and help with driving R1 and R6. The output point has also been shifted back to the second emitter-follower. The result is an input impedance of 60 MΩ, or 50 MΩ with an external 10 kΩ load.

www.raguvaran.puzl.com

We have achieved this using only five transistors, so the BJT is clearly not an inherently low-impedance device. There are many more technical possibilities if you need a really astronomical input impedance; the record in 1968 appears to have been no less than 20,000 MΩ [17].

The circuits in Figure 3.35 are not optimized for linearity. They all give substantially more distortion when fed from a high source impedance such as 1 MΩ, as the base currents drawn are not linear.

References

[1] G. Erdi, 300 V/μs monolithic voltage follower, IEE Journal of Solid-State Circuits SC-14 6 (December 1979), p. 1059.
[2] B. Hilton, in: J. Williams (Ed.), Analog Circuit Design: Art, Science & Personalities, Butterworth-Heinemann, 1991, p. 193 (Chapter 21).
[3] J. Early, Effects of space-charge layer widening in junction transistors, Proc. IRE 40 (1952), pp. 1401–1406.
[4] B. Gilbert, in: Tomazou, Lidgey, Haigh (Eds), Analogue Design: The Current-Mode Approach, IEE Circuits and Systems Series 2, 1990, p. 12 et seq.
[5] D.L. Feucht, The Handbook of Analog Circuit Design, Academic Press, 1990, p. 484 et seq.
[6] D. Self, Audio Power Amplifier Design Handbook, fifth ed, Focal Press, 2009, p. 78.
[7] L. Nelson-Jones, Wideband oscilloscope probe, Wireless World (August 1968), p. 276.
[8] E. White, Improvements in or relating to thermionic valve amplifier circuit arrangements, British patent no. 564,250 (1940).
[9] D.L. Feucht, The Handbook of Analog Circuit Design, Academic Press, 1990, p. 310 et seq.
[10] de Lange, Flat wideband buffers, Electronics World (October 2003), p. 19.
[11] P. Staric, E. Margan, Wideband Amplifiers, Springer, 2006, Chapter 5, p. 5.117 (Schlotzaur).
[12] D. Self, A high-performance preamplifier design, Wireless World (February 1979), p. 41.
[13] D. Self, An advanced preamplifier design, Wireless World (November 1976), p. 41.
[14] T.D. Towers, High input-impedance amplifier circuits, Wireless World (July 1968), p. 197.
[15] D. Self, Audio Power Amplifier Design Handbook, fifth ed, Focal Press, 2009, p. 119.
[16] P.A. Johnson, Comment on “High input-impedance amplifier circuits”, Letters, Wireless World (September 1968), p. 303.
[17] G.W. Horn, Feedback reduces bio-probe’s input capacitance, Electronics (March 1968), pp. 97–98.

www.raguvaran.puzl.com

CHAPTER 4

Op-Amps and Their Properties

Audio design has for many years relied on a very small number of op-amp types; the TL072 and the 5532 dominated the audio small-signal scene for many years. The TL072, with its JFET inputs, was used wherever its negligible input bias currents and low cost were important. For a long time the 5534/5532 was much more expensive than the TL072, so the latter was used wherever feasible in an audio system, despite its inferior noise, distortion, and load-driving capabilities. The 5534 was reserved for critical parts of the circuitry. Although it took many years, the price of the 5534 is now down to the point where you need a very good reason to choose any other type of op-amp for audio work.

The TL072 and the 5532 are dual op-amps; the single equivalents are TL071 and 5534. Dual op-amps are used almost universally, as the package containing two is usually cheaper than the package containing one, simply because it is more popular.

There are, however, other op-amps, some of which can be useful, and a selected range is covered here.

A Very Brief History of Op-Amps

The op-amp is today thought of as quintessentially a differential amplifier, responding to the difference of the input voltages while (hopefully) ignoring any common-mode component. The history of differential amplifiers goes back to that great man Alan Blumlein, and his 1936 patent [1] for a pair of valves with their cathodes connected to ground through a common resistor. However, the first valve-based operational amplifiers, i.e. those intended to be capable of performing a mathematical operation, were in fact not differential at all, having only one input. That had to be an inverting input, of course, so you could apply negative feedback.

The first op-amp to get real exposure in the UK was the Fairchild uA709, designed by the renowned Bob Widlar and introduced in 1965. It was a rather awkward item that required quite complicated external compensation and was devoid of output short-circuit protection. One slip of the probe and an expensive IC was gone. It was prone to latch-up with high common-mode voltages and did not like capacitative loads. I for one found all this most discouraging, and gave up on the 709 pretty quickly. If you're going to quit, do it early, I say.

Small Signal Audio Design; ISBN: 9780240521770
Copyright © 2010 Elsevier Ltd; All rights of reproduction, in any form, reserved.

www.raguvaran.puzl.com

The arrival of the LM741 was a considerable relief. To my mind, it was the first really practical op-amp, and it was suddenly possible to build quite complex circuitry with a good chance of it being stable, doing what it should do, and not blowing up at the first shadow of an excuse. I have given some details of it in this chapter for purely historical reasons. There is also an interesting example of how to apply the LM741 appropriately in Chapter 17.

The first IC op-amps opened up a huge new area of electronic applications, but after the initial enthusiasm for anything new, the audio market greeted these devices with less than enthusiasm. There were good reasons for this. The LM741 worked reliably; the snag with using it for audio was the leisurely slew rate of 0.5 V/μs, which made full output at 20 kHz impossible. For a period of at least 5 years, roughly from 1972 to 1977, the only way to obtain good performance in a preamp was to stick with discrete transistor Class-A circuitry, and this became recognized as a mark of high quality. The advent of the TL072 and the 5532 changed this situation completely, but there is still marketing cachet to be gained from a discrete design.

An excellent and detailed history of operational amplifiers can be found in Ref. [2].

Op-Amp Properties: Noise

There is no point in regurgitating manufacturers' data sheets, especially since they are readily available on the internet. Here I have simply ranked the op-amps most commonly used for audio in order of voltage noise (Table 4.1).

TABLE 4.1 Op-amps ranked by voltage noise density (typical)

Op-amp	e_n (nV/$\sqrt{\text{Hz}}$)	i_n (pA/$\sqrt{\text{Hz}}$)	Input device type	Bias cancel?
TL072	18	0.01	FET	No
OPA604	11	0.004	FET	No
NJM4556	8	Not specified	BJT	No
OPA2134	8	0.003	FET	No
OP275	6	1.5	BJT + FET	No
OPA627	5.2	0.0025	FET	No
5532A	5	0.7	BJT	No
LM833	4.5	0.7	BJT	No
MC33078	4.5	0.5	BJT	No
5534A	3.5	0.4	BJT	No
OP270	3.2	0.6	BJT	No
OP27	3	0.4	BJT	Yes
LM4562	2.7	1.6	BJT	No
AD797	0.9	2	BJT	No
LT1028	0.85	1	BJT	Yes

www.raguvaran.puzl.com

The great divide is between JFET input op-amps and BJT input op-amps. The JFET op-amps have more voltage noise but less current noise than bipolar input op-amps, the TL072 being particularly noisy. If you want the lowest voltage noise, it has to be a bipolar input. The difference, however, between a modern JFET-input op-amp such as the OPA2134 and the old faithful 5532 is only 4 dB, but the JFET part is a good deal more costly. The bipolar AD797 seems to be out on its own here, but it is a specialized and expensive part. The LT1028 is not suitable for audio use for reasons described later. The LM741, which is included in this chapter for purely historical reasons, is omitted from Table 4.1 because there are no noise specs on its data sheets.

Op-amps with bias-cancellation circuitry are normally unsuitable for audio use due to the extra noise this creates. The amount depends on circuit impedances, and is not taken into account in Table 4.1. The general noise behavior of op-amps in circuits is dealt with in Chapter 1.

Op-Amp Properties: Slew Rate

Slew rates vary more than most parameters; a range of 100:1 is shown in Table 4.2. The slowest is the 741, which is the only type not fast enough to give full output over the audio band. There are faster ways to handle a signal, such as current-feedback architectures, but they

TABLE 4.2 Op-amps ranked by slew rate (typical)

Op-amp	V/µs
LM741	0.5
OP270	2.4
OP27	2.8
NJM4556	3
MC33078	7
LM833	7
5532A	9
LT1028	11
TL072	13
5534A	13
OPA2134	20
LM4562	20
AD797	20
OP275	22
OPA604	25
OPA627	55

www.raguvaran.puzl.com

usually fall down on linearity. In any case, a maximum slew rate greatly in excess of what is required appears to confer no benefits whatever.

The 5532 slew rate is typically ±9 V/µs. This version is internally compensated for unity-gain stability, not least because there are no spare pins for compensation when you put two op-amps in an eight-pin dual package. The single-amp version, the 5534, can afford a couple of compensation pins, and so is made to be stable only for gains of 3× or more. The basic slew rate is therefore higher at ±13 V/µs.

Compared with power-amplifier specs, which often quote 100 V/µs or more, these speeds may appear rather sluggish. In fact they are not; even ±9 V/µs is more than fast enough. Assume you are running your op-amp from ±18 V rails, and that it can give a ±17 V swing on its output. For most op-amps this is distinctly optimistic, but never mind. To produce a full-amplitude 20 kHz sine wave you only need 2.1 V/µs, so even in the worst case there is a safety margin of at least four times. Such signals do not of course occur in actual use, as opposed to testing. More information on slew limiting is given in the section on op-amp distortion.

Op-Amp Properties: Common-Mode Range

This is simply the range over which the inputs can be expected to work as proper differential inputs. It usually covers most of the range between the rail voltages, with one notable exception. The data sheet for the TL072 shows a common-mode (CM) range that looks a bit curtailed at −12 V. This bland figure hides the deadly trap this IC contains for the unwary. Most op-amps, when they hit their CM limits, simply show some sort of clipping. The TL072, however, when it hits its negative limit, promptly inverts its phase, so your circuit either latches up, or shows nightmare clipping behavior with the output bouncing between the two supply rails. The positive CM limit is, in contrast, trouble-free. This behavior can be especially troublesome when TL072s are used in high-pass Sallen-and-Key filters.

Op-Amp Properties: Input Offset Voltage

A perfect op-amp would have its output at 0 V when the two inputs were exactly at the same voltage. Real op-amps are not perfect and a small voltage difference – usually a few millivolts – is required to zero the output. These voltages are large enough to cause switches to click and pots to rustle, and DC blocking is often required to keep them in their place. This issue is examined in depth in Chapter 11.

The typical offset voltage for the 5532A is ±0.5 mV typical, ±4 mV maximum at 25°C; the 5534A has the same typical spec but a lower maximum at ±2 mV. The input offset voltage of the new LM4562 is only ±0.1 mV typical, ±4 mV maximum at 25°C.

www.raguvaran.puzl.com

Op-Amp Properties: Bias Current

Bipolar-input op-amps not only have larger noise currents than their JFET equivalents; they also have much larger bias currents. These are the base currents taken by the input transistors. This current is much larger than the input offset current, which is the difference between the bias current for the two inputs. For example, the 5532A has a typical bias current of 200 nA, compared with a much smaller input offset current of 10 nA. The LM4562 has a lower bias current of 10 nA typical, 72 nA maximum. In the case of the 5532/4 the bias current flows into the input pins as the input transistors are NPN.

Bias currents are a considerable nuisance; when they flow through variable resistors they make them noisy when moved. They will also cause significant DC offsets when they flow through high-value resistors.

It is often recommended that the effect of bias currents can be canceled out by making the resistance seen by each op-amp input equal. Figure 4.1(a) shows a shunt-feedback stage with a 22 kΩ feedback resistor. When 200 nA flows through this it will generate a DC offset of 4.4 mV, which is rather more than we would expect from the input offset voltage error.

If an extra resistance R_{compen}, of the same value as the feedback resistor, is inserted into the non-inverting input circuit then the offset will be canceled. This strategy works well and is done almost automatically by many designers. However, there is a snag. The resistance R_{compen} generates extra Johnson noise, and to prevent this it is necessary to shunt the resistance with a capacitor, as in Figure 4.1(b). This extra component costs money and takes up PCB space, so it is questionable if this technique is actually very useful for audio work. It is usually more economical to allow offsets to accumulate in a chain of op-amps, and then remove the DC voltage with a single output blocking capacitor. This assumes that there are no stages with a large DC gain, and that the offsets are not large enough to significantly reduce the available

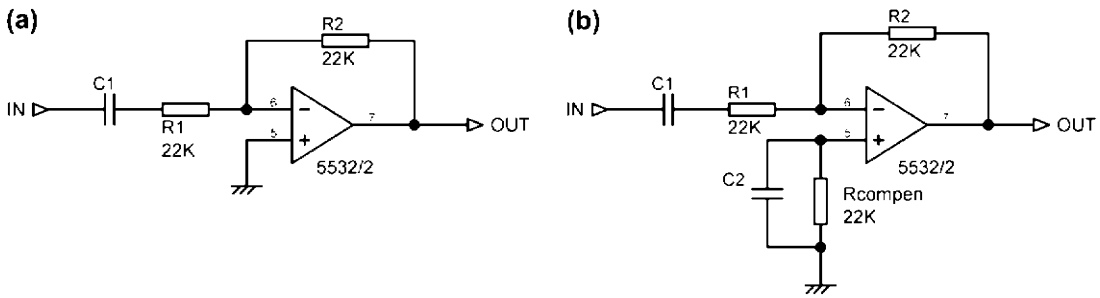


Figure 4.1: Compensating for bias-current errors in a shunt-feedback stage. The compensating resistor must be bypassed by a capacitor C2 to prevent it adding Johnson noise to the stage

www.raguvaran.puzl.com

voltage swing. Care must also be taken if controls are involved, because even a small DC voltage across a potentiometer will cause it to become crackly, especially as it wears.

FET input op-amps have very low bias current at room temperature; however, it doubles for every 10°C rise. This is pretty unlikely to cause trouble in most audio applications, but a combination of high internal temperatures and high-value pots could lead to some unexpected crackling noises.

Op-Amp Properties: Cost

While it may not appear on the data sheet, the price of an op-amp is obviously a major factor in deciding whether or not to use it. Table 4.3 was derived from the averaged prices for 1+ and 25+ quantities across a number of UK distributors. At the time of writing (September 2009) the cheapest popular op-amps are the TL072 and the 5532, and these happened to come out at exactly the same price, so their price is taken as unity and used as the basis for the price ratios given.

Table 4.3 was compiled using prices for DIL packaging and the cheapest variant of each type. Price is per package and not per op-amp section. It is obviously only a rough guide. Purchasing in large quantities or in different countries may change the rankings somewhat (even going from 1+ to 25+ causes some changes) but the basic look of things will not alter too much. One thing is obvious – the 5532 is one of the great op-amp bargains of all time.

TABLE 4.3 Op-amps ranked by price (2009) relative to 5532 and TL072

Quantity	Format	Price ratio 1+		Price ratio 25+
LM833	Dual	1.45	5532	**1.00**
5532	Dual	1.64	TL072	**1.00**
MC33078	Dual	1.97	LM833	1.12
TL072	Dual	2.45	MC33078	1.27
OPA604	Single	5.09	TL052	2.55
OPA2134PA	Dual	5.55	OP275GP	3.42
TL052	Dual	5.76	OPA2134PA	4.45
OP275GP	Dual	7.18	OPA604	5.03
OP27	Single	8.67	OP27	6.76
LM4562	Dual	12.45	LM4562	9.06
AD797	Single	25.73	AD797	13.09
OP270	Dual	29.85	LT1028	17.88
LT1028	Single	30.00	OPA270	24.42
OPA627	Single	51.91	OPA627	48.42

www.raguvaran.puzl.com

Op-Amp Properties: Distortion

Relatively few discussions of op-amp behavior deal with non-linear distortion, perhaps because it is a complex business. Op-amp 'accuracy' is closely related, but the term is often applied only to DC operation. Accuracy here is often specified in terms of bits, so '20-bit accuracy' means errors not exceeding one part in 2^{20}, which is −120 dB or 0.0001%. Audio signal distortion is of course a dynamic phenomenon, very sensitive to frequency, and DC specs are of no use at all in estimating it.

Distortion is always expressed as a ratio, and can be quoted as a percentage, as number of decibels, or in parts per million (ppm). With the rise of digital processing, treating distortion as the quantization error arising from the use of a given number of bits has become more popular. Figure 4.2 hopefully provides a way of keeping perspective when dealing with these different metrics.

There are several different causes of distortion in op-amps. We will now examine them.

Op-Amp Internal Distortion

This is what might be called the basic distortion produced by the op-amp you have selected. Even if you scrupulously avoid clipping, slew-limiting, and common-mode issues, op-amps are not distortion free, though some types such as the 5532 and the LM4562 have very low levels. If distortion appears when the op-amp is run with shunt feedback, to prevent common-mode voltages on the inputs, and with very light output loading, then it is probably wholly internal and there is nothing to be done about it except pick a better op-amp.

If the distortion is higher than expected, the cause may be internal instability provoked by putting a capacitative load directly on the output, or neglecting the supply decoupling. The classic example of the latter effect is the 5532, which shows high distortion if there is not a capacitor across the supply rails close to the package; 100 nF is usually adequate. No actual HF oscillation is visible on the output with a general-purpose oscilloscope, so the problem may be instability in one of the intermediate gain stages.

Slew-Rate-Limiting Distortion

While this is essentially an overload condition, it is wholly the designer's responsibility. If users whack up the gain until the signal is within a hair of clipping, they should still be able to assume that slew limiting will never occur, even with aggressive material full of high frequencies.

Arranging this is not too much of a problem. If the rails are set at the usual maximum voltage, i.e. ±18 V, then the maximum possible signal amplitude is 12.7 Vrms, ignoring the saturation

www.raguvaran.puzl.com

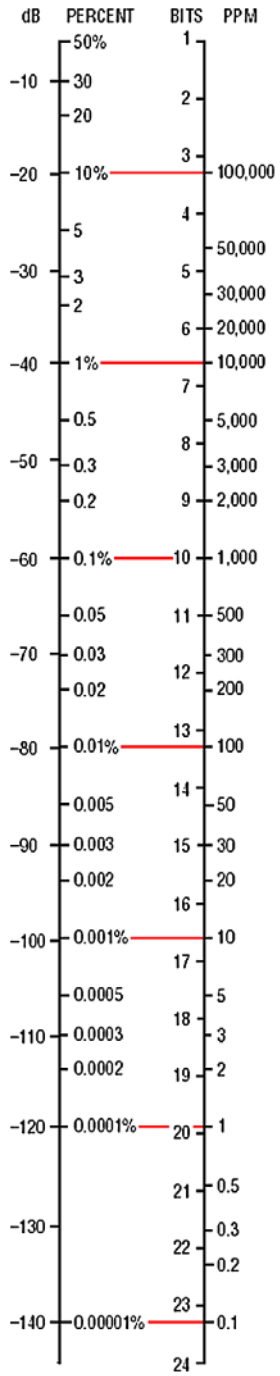


Figure 4.2: The relation between different ways of quoting THD: decibels, percentages, bit accuracy, and parts per million

www.raguvaran.puzl.com

voltages of the output stage. To reproduce this level cleanly at 20 kHz requires a minimum slew rate of only 2.3 V/μs. Most op-amps can do much better than this, though with the OP27 (2.8 V/μs) you are sailing rather close to the wind. The old LM741 looks as though it would be quite unusable, as its very limited 0.5 V/μs slew rate allows a full output swing only up to 4.4 kHz.

Horrific as it may now appear, audio paths full of LM741s were quite common in the early 1970s. Entire mixers were built with no other active devices, and what complaints there were tended to be about noise rather than distortion. The reason for this is that full-level signals at 20 kHz simply do not occur in reality; the energy at the HF end of the audio spectrum is well known to be much lower than that at the bass end.

This assumes that slew limiting has an abrupt onset as level increases, rather like clipping. This is in general the case. As the input frequency rises and an op-amp gets closer to slew limiting, the input stage is working harder to supply the demands of the compensation capacitance. There is an absolute limit to the amount of current this stage can supply, and when you hit it the distortion shoots up, much as it does when you hit the supply rails and induce voltage clipping. Before you reach this point, the linearity may be degraded, but usually only slightly until you get close to the limit. It is not normally necessary to keep big margins of safety when dealing with slew limiting. If you are employing the usual suspects of the audio op-amp world – the 5532 and TL072, with maximal slew rates of 9 and 13 V/μs respectively – you are most unlikely to suffer any slew-rate non-linearity.

Distortion Due to Loading

Output stage distortion is always worse with heavy output loading because the increased currents flowing exacerbate the gain changes in the Class-B output stage. These output stages are not individually trimmed for optimal quiescent conditions (as are audio power amplifiers) and so the crossover distortion produced by op-amps tends to be both higher and more variable between different specimens of the same chip. Distortion increases with loading in different ways for different op-amps. It may rise only at the high-frequency end (e.g. the OP2277) or there may be a general rise at all frequencies. Often both effects occur, as in the TL072.

The lowest load that a given op-amp can be allowed to drive is an important design decision. It will typically be a compromise between the distortion performance required and opposing factors such as number of op-amps in the circuit, cost of load-capable op-amps, and so on. It even affects noise performance, for the lower the load resistance an amplifier can drive, the lower the resistance values in the negative feedback can be, and hence the lower the Johnson noise they generate. There are limits to what can be done in noise reduction by this method, because Johnson noise is proportional to the square root of circuit resistance, and so improves only slowly as op-amp loading is increased.

www.raguvaran.puzl.com

Thermal Distortion

Thermal distortion is that caused by cyclic variation of the properties of the amplifier components due to the periodic release of heat in the output stage. The result is a rapid rise in distortion at low frequencies, which gets worse as the loading becomes heavier. Those who have read my work on audio power amplifiers will be aware that I am highly sceptical – in fact totally sceptical – about the existence of thermal distortion in amplifiers built from discrete components [3]. The power devices are too massive to experience per-cycle parameter variations, and there is no direct thermal path from the output stage to the input devices. There is no rise, rapid or otherwise, in distortion at low frequencies in a properly designed discrete power amplifier. The situation is quite different in op-amps, where the output transistors have much less thermal inertia and are also on the same substrate as the input devices. Nonetheless, op-amps do not normally suffer from thermal distortion; there is generally no rise in low-frequency distortion, even with heavy output loading. Integrated-circuit power amplifiers are another matter, and the much greater amounts of heat liberated on the substrate do appear to cause serious thermal distortion, rising at 12 dB/octave below 50 Hz. I have never seen anything resembling this in any normal op-amp.

Common-Mode Distortion

This is the general term for extra distortion that appears when there is a large signal voltage on both the op-amp inputs. The voltage difference between these two inputs will be very small, assuming the op-amp is in its linear region, but the common-mode (CM) voltage can be a large proportion of the available swing between the rails.

It appears to be by far the least understood mechanism, and gets little or no attention in op-amp textbooks, but it is actually one of the most important influences on op-amp distortion. It is simple to separate this effect from the basic forward-path distortion by comparing THD performance in series and shunt-feedback modes; this should be done at the same noise gain. The distortion is usually a good deal lower for the shunt-feedback case where there is no common-mode voltage. Bipolar and JFET input op-amps show different behavior and they are treated separately below.

Bipolar input op-amps

Figure 4.4 shows the distortion from a 5532 working in shunt mode with low-value resistors of 1 kΩ and 2k2 setting a gain of 2.2 times, at an output level of 5 Vrms. This is the circuit of Figure 4.3(a) with R_s set to zero; there is no CM voltage. The distortion is well below 0.0005% up to 20 kHz; this underlines what a superlative bargain the 5532 is.

www.raguvaran.puzl.com

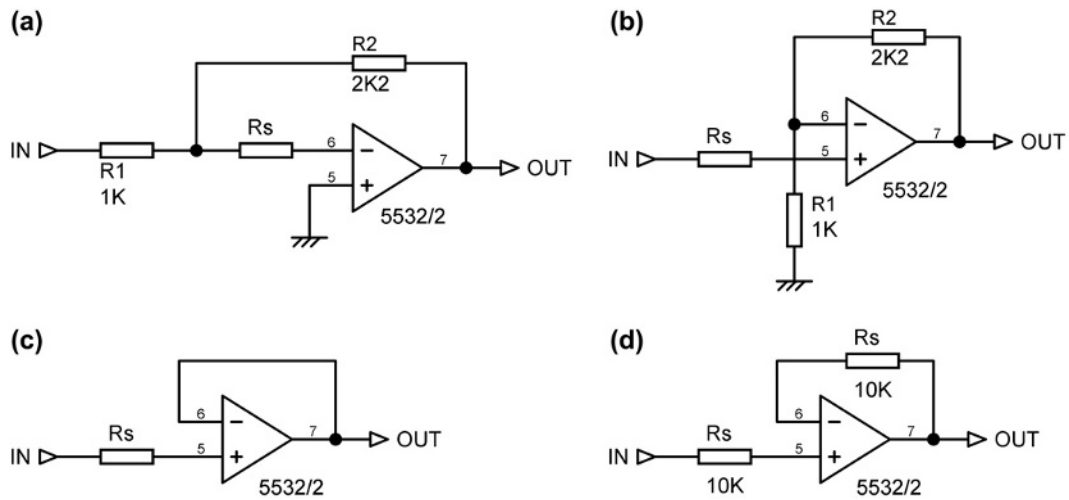


Figure 4.3: Op-amp test circuits with added source resistance R_s. (a) Shunt. (b) Series. (c) Voltage-follower. (d) Voltage-follower with cancellation resistor in feedback path

Figure 4.5 shows the same situation but with the output increased to 10 Vrms (the clipping level on ±18 V rails is about 12 Vrms) and there is now significant distortion above 10 kHz, though it only exceeds 0.001% at 18 kHz.

This remains the case when R_s in Figure 4.3(a) is increased to 10 kΩ and 47 kΩ – the noise floor is higher but there is no real change in the audio-band distortion behavior. The significance of this will be seen in a moment.

We will now connect the 5532 in the series-feedback configuration, as in Figure 4.3(b); note that the stage gain is greater at 3.2 times but the op-amp is working at the same noise gain. The

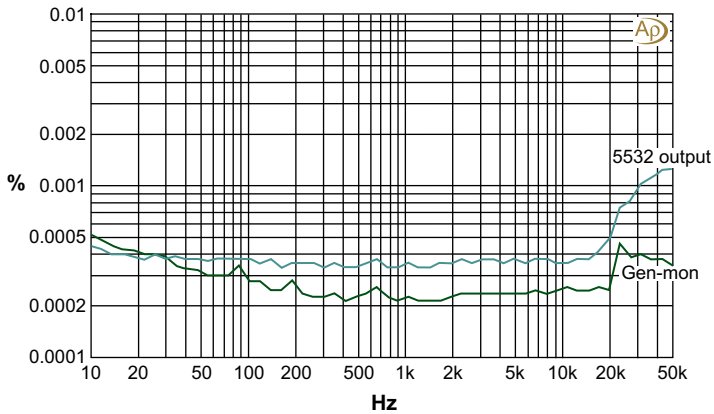


Figure 4.4: 5532 distortion in a shunt-feedback circuit at 5 Vrms out. This shows the AP SYS-2702 output (lower trace) and the op-amp output (upper trace). Supply ±18 V

www.raguvaran.puzl.com

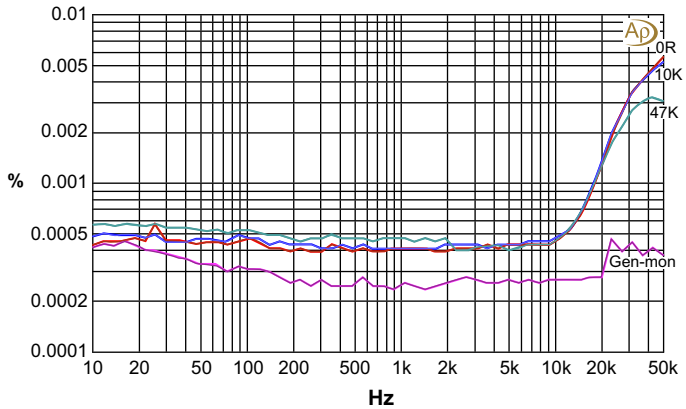


Figure 4.5: 5532 distortion in the shunt-feedback circuit of Figure 4.3(b). Adding extra resistances of 10 kΩ and 47 kΩ in series with the inverting input does not degrade the distortion at all, but does bring up the noise floor a bit. Test level 10 Vrms out, supply ±18 V

CM voltage is 3.1 Vrms. With a 10 Vrms output we can see in Figure 4.6 that even with no added source resistance the distortion starts to rise from 2 kHz, though it does not exceed 0.001% until 12 kHz. But when we add some source resistance R_s, the picture is radically worse, with serious mid-band distortion rising at 6 dB/octave, and roughly proportional to the amount of resistance added. We will note it is 0.0085% at 10 kHz with $R_s = 47$ kΩ.

The worst case for CM distortion is the voltage-follower configuration, as in Figure 4.3(c), where the CM voltage is equal to the output voltage. Figure 4.7 shows that even with a CM voltage of 10 Vrms, the distortion is no greater than for the shunt mode. However, when source resistance is inserted in series with the input, the distortion mixture of second, third, and other low-order harmonics increases markedly. It increases with output level, approximately

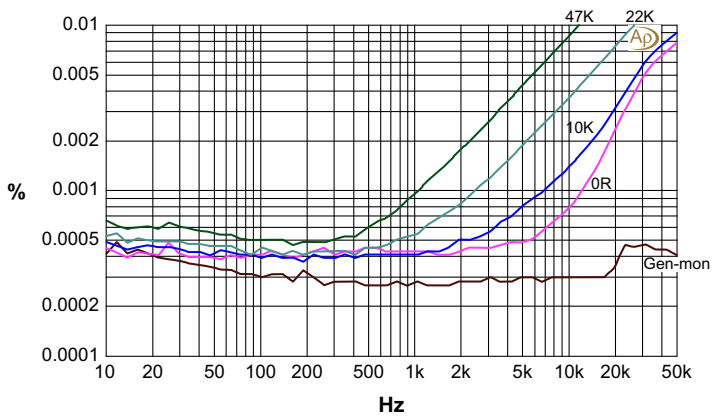


Figure 4.6: 5532 distortion in a series-feedback stage with 2k2 and 1k feedback resistors, and varying source resistances. Output 10 Vrms

www.raguvaran.puzl.com

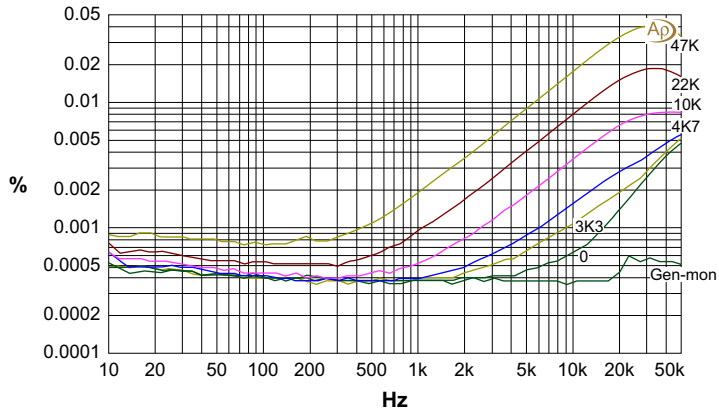


Figure 4.7: 5532 distortion in a voltage-follower circuit with a selection of source resistances. Test level 10 Vrms, supply ±18 V. The lowest trace is the analyzer output measured directly, as a reference

quadrupling as level doubles. The THD is now 0.018% at 10 kHz with $R_s = 47$ kΩ, more than twice that of the series-feedback amplifier above, due to the increased CM voltage.

It is would be highly inconvenient to have to stick to the shunt-feedback mode, because of the phase inversion and relatively low input impedance that comes with it, so we need to find out how much source resistance we can live with. Figure 4.8 zooms in on the situation with resistance of 10 kΩ and below; when the source resistance is below 2k2, the distortion is barely distinguishable from the zero source resistance trace. This is why the low-pass Sallen-and-Key filters in Chapter 5 have been given series resistors that do not in total exceed this figure.

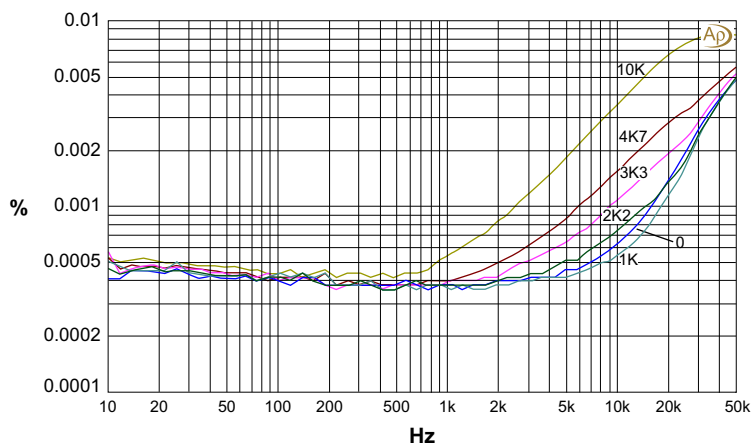


Figure 4.8: A closer look at 5532 distortion in a voltage-follower with relatively low source resistances; note that a 1 kΩ source resistance actually gives less distortion than none. Test level 10 Vrms, supply ±18 V

www.raguvaran.puzl.com

Close examination reveals the intriguing fact that a 1 kΩ source actually gives *less* distortion than no source resistance at all, reducing THD from 0.00065% to 0.00055% at 10 kHz. Minor resistance variations around 1 kΩ make no difference. This must be due to the cancellation of distortion from two different mechanisms. It is hard to say whether it is repeatable enough to be exploited in practice.

So, what's going on here? Is it simply due to non-linear currents being drawn by the op-amp inputs? Audio power amplifiers have discrete input stages that are very simple compared with those of most op-amps, and draw relatively large input currents. These currents show appreciable non-linearity even when the output voltage of the amplifier is virtually distortion free, and, if they flow through significant source resistances, will introduce added distortion [4].

If this was the case with the 5532 then the extra distortion would manifest itself whenever the op-amp was fed from a significant source resistance, no matter what the circuit configuration. But we have just seen that it only occurs in series-feedback situations; increasing the source resistance in a shunt-feedback does not perceptibly increase distortion. The effect may be present but if so it is very small, no doubt because op-amp signal input currents are also very small, and it is lost in the noise.

The only difference is that the series circuit has a CM voltage of about 3 Vrms, while the shunt circuit does not, and the conclusion is that with a bipolar input op-amp you must have *both* a CM voltage and a significant source resistance to see extra distortion. The input stage of a 5532 is a straightforward long-tailed pair (see Figure 4.21 below) with a simple tail-current source, and no fancy cascoding, and I suspect that the Early effect operates on it when there is a large CM voltage, modulating the quite high input bias currents, and this is what causes the distortion. The signal input currents are much smaller, due to the high open-loop gain of the op-amp, and as we have seen appear to have a negligible effect.

FET op-amps

FET-input op-amps behave differently from bipolar-input op-amps. Take a look at Figure 4.9, taken from a TL072 working in shunt and in series configuration with a 5 Vrms output. The circuits are as in Figure 4.3(a) and (b), except that the resistor values have to be scaled up to 10 and 22 kΩ because the TL072 is nothing like so good at driving loads as the 5532. This unfortunately means that the inverting input is seeing a source resistance of 10k||22k = 6.9k, which introduces a lot of CM distortion in the series case – five times as much at 20 kHz as for the shunt case. Adding a similar resistance in the input path cancels out this distortion, and the trace then is the same as the 'Shunt' trace in Figure 4.9. Disconcertingly, the value that achieved this was not 6.9k, but 9k1. That means adding −113 dBu of Johnson noise, so it's not always appropriate.

It's worth mentioning that the flat part of the 'Shunt' trace below 10 kHz is not noise, as it would be for the 5532; it is distortion.

www.raguvaran.puzl.com

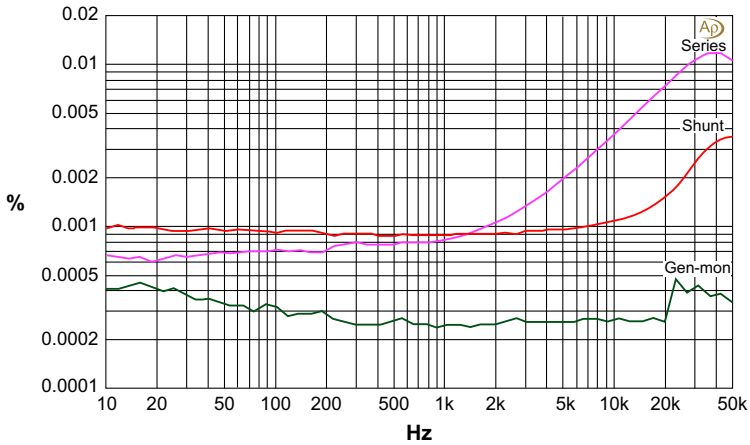


Figure 4.9: A TL072 shunt-feedback stage using 10 and 22 kΩ resistors shows low distortion. The series version is much worse due to the impedance of the NFB network, but it can be made the same as the shunt case by adding cancellation source resistance in the input path. No external loading, test level 5 Vrms, supply ±18 V

A voltage-follower has no inconvenient medium-impedance feedback network, but it does have a much larger CM voltage. Figure 4.10 shows a voltage-follower working at 5 Vrms. With no source resistance the distortion is quite low, due to the 100% NFB, but as soon as a 10 kΩ source resistance is added we are looking at 0.015% at 10 kHz.

Once again, this can be cured by inserting an equal resistance in the feedback path of the voltage-follower, as in Figure 4.3(d) above. This gives the 'Cancel' trace in Figure 4.10. Adding resistances for distortion cancellation in this way has the obvious disadvantage that they introduce extra Johnson noise into the circuit. Another point is that stages of this kind are

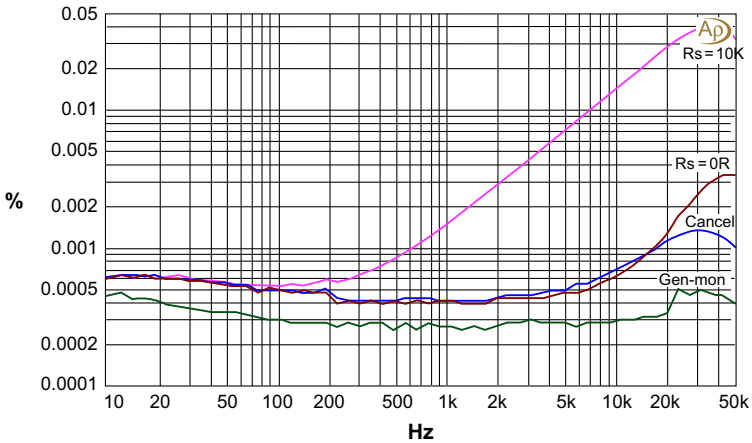


Figure 4.10: A TL072 voltage-follower working at 5 Vrms with a low source resistance produces little distortion (R_s = 0R), but adding a 10 kΩ source resistance makes things much worse (R_s = 10k). Putting a 10 kΩ resistance in the feedback path as well gives complete cancellation of this extra distortion (Cancel). Supply ±18 V

www.raguvaran.puzl.com

often driven from pot wipers, so the source impedance is variable, ranging between zero and one-quarter of the pot track resistance. Setting a balancing impedance in the other op-amp input to a mid-value, i.e. one-eighth of the track resistance, should reduce the average amount of input distortion, but it is inevitably a compromise.

With JFET inputs the problem is not the operating currents of the input devices themselves, which are negligible, but the currents drawn by the non-linear junction capacitances inherent in field-effect devices. These capacitances are effectively connected to one of the supply rails. For P-channel JFETs, as used in the input stages of most JFET op-amps, the important capacitances are between the input JFETs and the substrate, which is normally connected to the V− rail (see Jung [5]).

According to the Burr-Brown data sheet for the OPA2134, 'The P-channel JFETs in the input stage exhibit a varying input capacitance with applied CM voltage.' It goes on to recommend that the input impedances should be matched if they are above 2 kΩ.

Common-mode distortion can be minimized by running the op-amp off the highest supply rails permitted, though the differences are not large. In one test on a TL072, going from ±15 to ±18 V rails reduced the distortion from 0.0045% to 0.0035% at 10 kHz.

Rail bootstrapping to reduce CM distortion

So what do you do if you need a really high-impedance low-distortion voltage-follower and you have a significant source resistance, but you don't want the added noise that would come from adding a cancellation resistor? We noted above that the non-linear input capacitances that cause the trouble with JFET op-amp voltage-followers are effectively connected to the V− supply rail or substrate. This suggests a way to remove the problem: if the supply rails are bootstrapped so they go up and down with the inputs, the signal voltage across the non-linear input capacitances is zero, no current can flow through them, and no extra distortion is generated.

Figure 4.11(a) shows the idea. The resistor–Zener chain R4, D1, D2, R5 creates ±5 V rails that are moved up and down by op-amp A4, and buffered by A2, A3. A1 expects reasonably low supply-rail impedances at HF, and attempting to run it directly from the outputs of A2, A3 does not work – the signal disappears in a fog of HF oscillation. The two resistors R2, R3 prevent this by isolating C1 from A2, A3 outputs, while capacitor C1 across A1 supply pins keeps the HF rail impedance low. Since the ±5 V rails of A1 have to remain inside the fixed ±15 V supply rails, the possible swing of the supplies is limited and the maximum output is reduced compared with a basic voltage-follower. The circuit of Figure 4.11(a) clips at 6.7 Vrms (1 kHz). This could be increased somewhat by using ±17 or ±18 V fixed rails.

Figure 4.12 shows the result of basic bootstrapping while handling a 5 Vrms signal, which as we saw earlier, is enough to cause serious CM distortion. The increase in linearity is encouraging; the distortion is promptly halved.

www.raguvaran.puzl.com

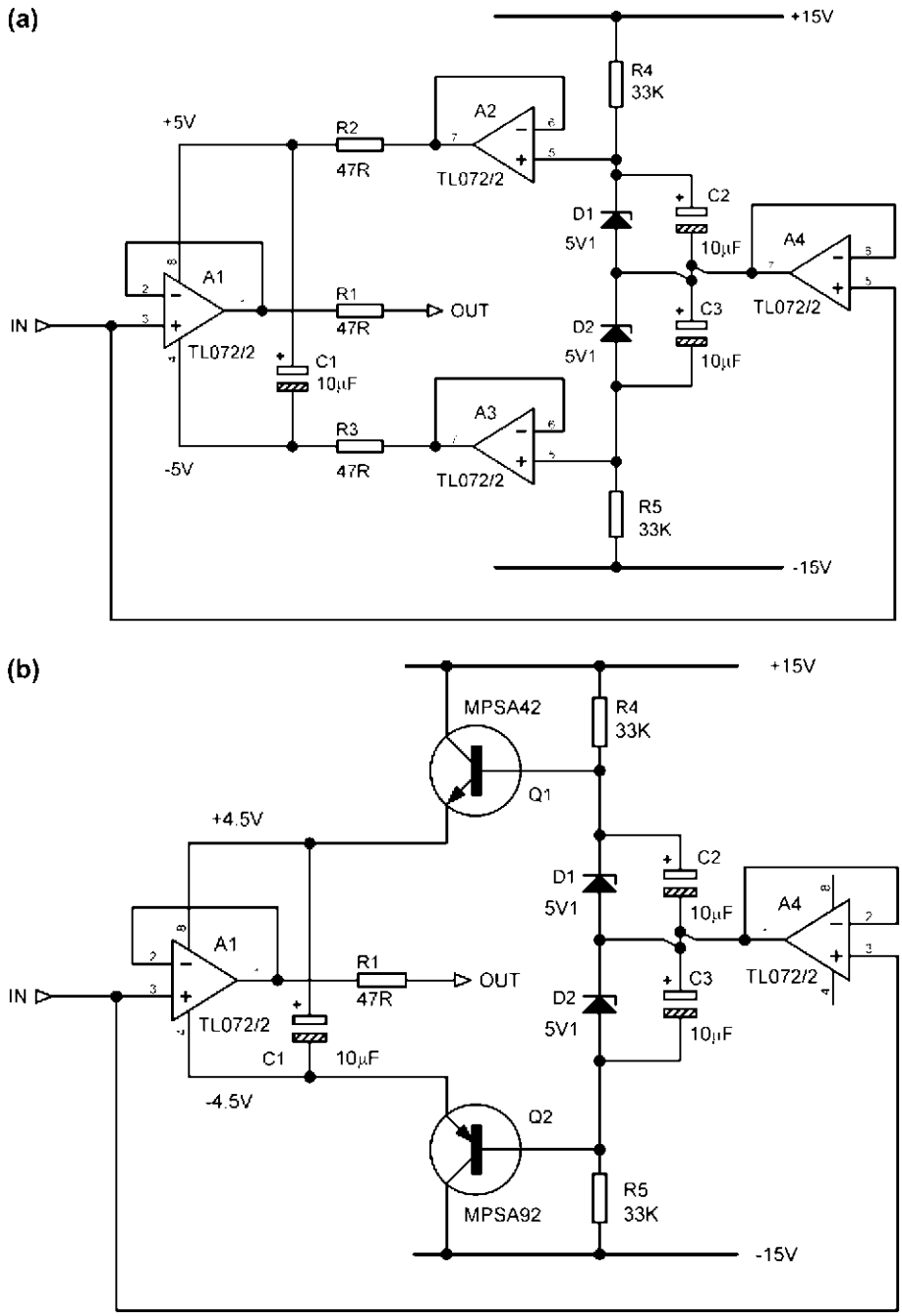


Figure 4.11: Bootstrapping the supply rails of voltage-follower A1 by moving them up and down with the input signal: (a) using op-amps; (b) using transistors

www.raguvaran.puzl.com

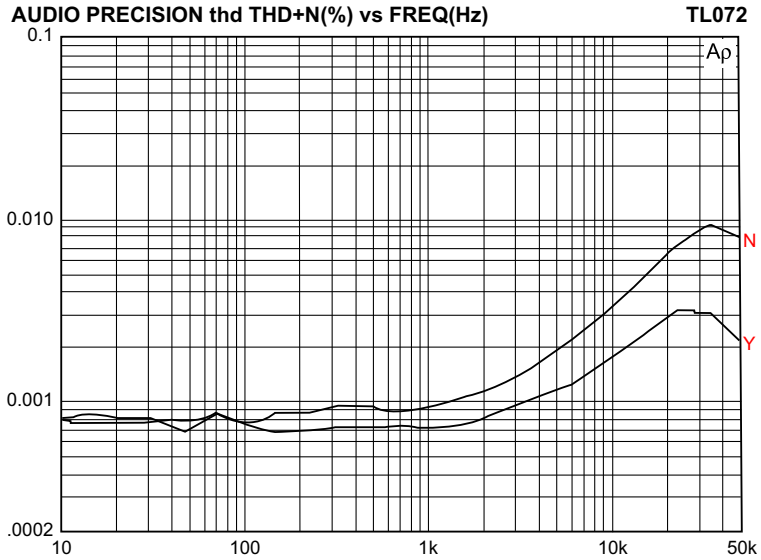


Figure 4.12: TL072 voltage-follower distortion with (Y) and without (N) rail bootstrapping. Test level 5 Vrms, supply ±15 V

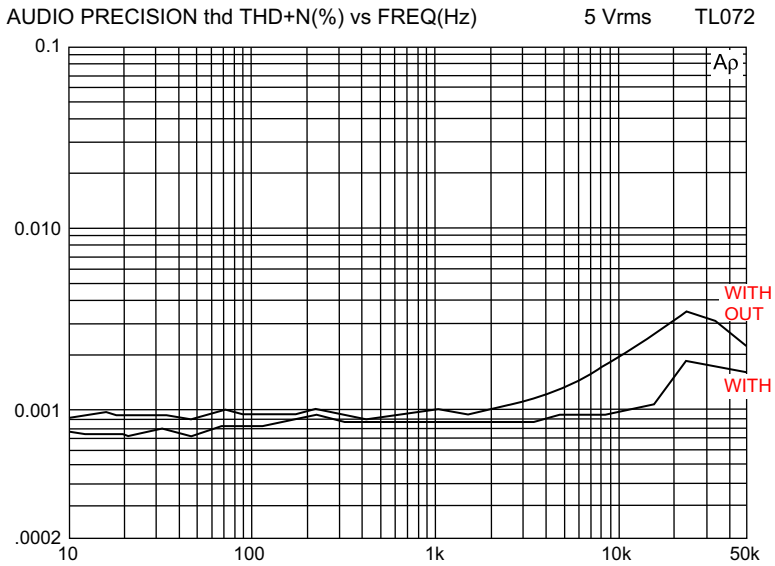


Figure 4.13: Rail bootstrapping is much enhanced by adding capacitors C2, C3. TL072, test level 5 Vrms, supply ±15 V. The 'WITH' trace is essentially the distortion of the test-gear alone

Figure 4.13, however, shows that we can do better by adding C2, C3. These are in parallel with the effective slope resistance of the Zeners, and improve the accuracy of the rail bootstrapping. The lower trace marked 'WITH' is once again indistinguishable from that of the test-gear alone.

www.raguvaran.puzl.com

Figures 4.12 and 4.13 were taken with near-zero source resistance, and show that internal CM distortion has been dealt with. But what happens when a 10 kΩ source resistance is reintroduced?

Figure 4.14 gives the answer: adding a 10 kΩ source resistance now makes almost no difference. Note that no cancellation resistor has been put in the feedback path.

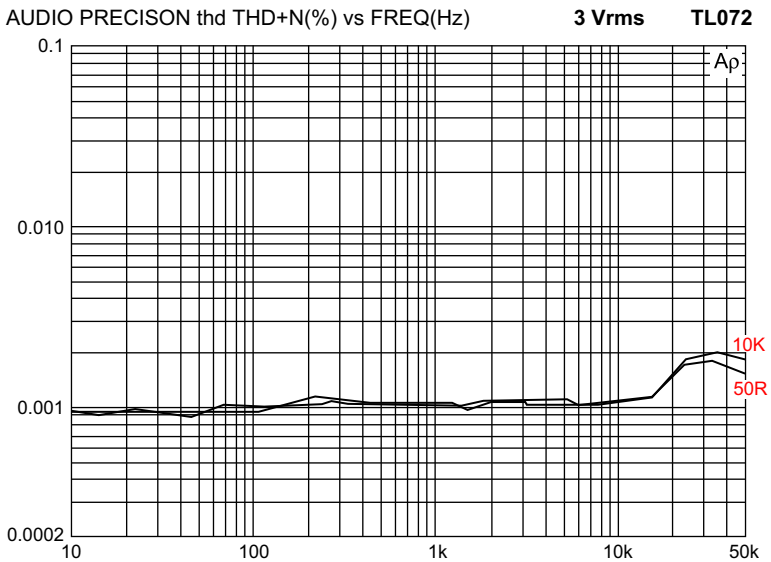


Figure 4.14: TL072 Voltage-follower distortion with 10 kΩ and 50 Ω source resistances, and no cancellation. Test level 3 Vrms, supply ±15 V

Simpler rail bootstrapping

On contemplating Figure 4.11(a), it may occur to you that using three op-amps to make a friendly environment for one is a bit over-complex. You are quite right. It is always good to simplicate and add lightness when you can, and A2 and A3 can in fact be replaced by simple emitter-followers with no detectable loss in performance, as in Figure 4.11(b). The two 47 Ω resistors have been removed, but C1 is retained. This seems to be reliably stable. The total supply voltage to A1 has been reduced by two V_{be} drops, or 1.2 V; it could be restored by increasing the Zener voltages if required. The simpler version also uses less power as we no longer need to supply the quiescent currents of A2 and A3.

The attentive reader will recall that the troublesome non-linear capacitances are effectively connected to the substrate, which is usually the V− supply rail. Would it not be possible to bootstrap just that rail, and leave V+ connected to a fixed +15 V rail? It would – it works, but the results with an OPA2134, while more linear than with a conventional voltage-follower, are worse than bootstrapping both rails. Before we were keeping the magnitude of the A1 supply

www.raguvaran.puzl.com

voltage substantially constant, although it was sailing up and down. If only one rail is bootstrapped the actual supply voltage is being modulated, so it is hardly surprising that linearity suffers.

The rail bootstrapping concept was also tested with the TL052 and the OPA2134 at 5 Vrms, and similar dramatic reductions in CM distortion were found.

In the previous section we saw that CM distortion is also generated by bipolar input op-amps, though by a different mechanism, so rail bootstrapping ought to work for these types of op-amp as well. Figure 4.15 shows that it does. Adding a 10 kΩ source resistance now causes virtually no extra distortion.

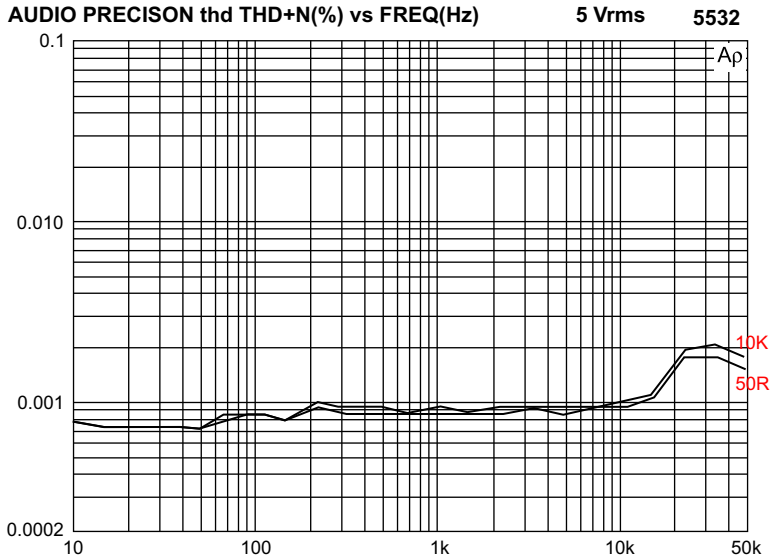


Figure 4.15: Rail bootstrapping works for 5532 voltage-followers as well; 10 kΩ and 50 Ω source resistances, and no cancellation. Test level 5 Vrms, supply ±15 V

Bootstrapping series-feedback JFET op-amp stages

The voltage-follower is the worst case for CM distortion, as the full output voltage exists on both inputs. In contrast, it is the best case for output loading, as there is no resistive feedback network at all to drive – just a high-impedance input pin. Similarly, the shunt-feedback amplifier is the best case for CM distortion as there is no significant signal on the inputs.

Series-feedback amplifier stages fall between these two cases. For a +10 dB amplifier stage, the signal on the inputs is one-third that of the output, and so the input distortion is less, but still very definitely present, as we saw in Figures 4.6 and 4.9. Amplifier stages like this can have a mixture of distortion mechanisms. The impedance of the NFB network, as seen from the inverting input of the amplifiers, is 22 kΩ in parallel with 10 kΩ, i.e. 6.87 kΩ. We have seen

www.raguvaran.puzl.com

above that this is enough to cause serious non-linearity unless the other input sees the same impedance, and it might be thought that reducing the impedance level of the NFB network would be a good way to deal with this, not least because it would minimize the Johnson noise produced by the network. Figure 4.16 shows that this does not work for the TL072; if the feedback network impedance is reduced by a factor of 10 the distortion gets worse rather than better, due to the heavier loading on the output. Input distortion has been replaced by a larger amount of output distortion; this is not a good exchange. Lowering the NFB network impedance is, however, likely to be successful with JFET op-amps having better load-driving capability than the TL072.

Rail bootstrapping is once more a possible answer. We drive the op-amp supply rails up and down with the same signal as the *input* – not the output. The only modification required is to take the increased output swing into account by increasing the A1 supply voltage to ±10 V (see Figure 4.17). The Zeners have been replaced with simple resistive dividers. This works just as well, and is a good thing as Zeners are more expensive than resistors. Figure 4.18 shows the excellent results.

Selecting the Right Op-Amp

In audio work, the 5532 is pre-eminent. It is found in almost every mixing console, and in a large number of preamplifiers. Distortion is almost unmeasurably low, even when driving

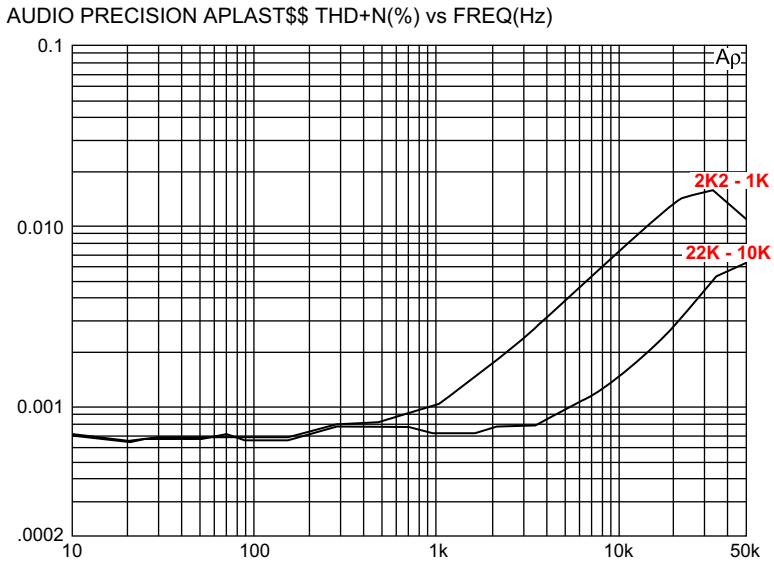


Figure 4.16: With the TL072, reducing the impedance of the negative-feedback network may reduce input distortion, but output distortion more than makes up for it because of the extra loading. Upper trace 2k2 – 1 kΩ, lower trace 22 kΩ – 10 kΩ in feedback network

www.raguvaran.puzl.com

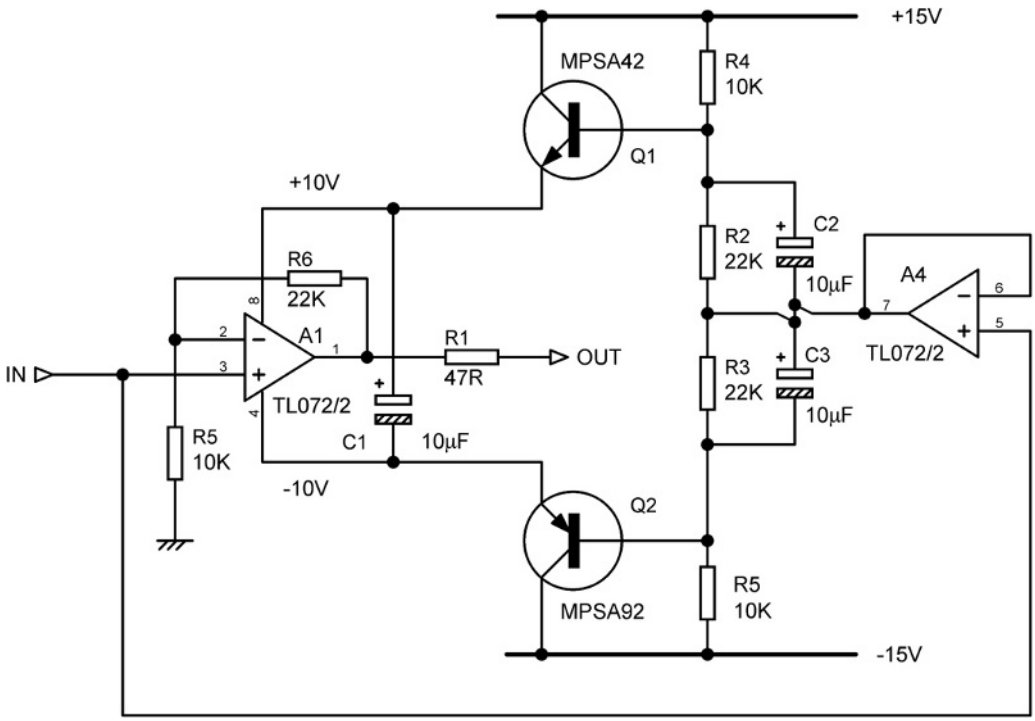


Figure 4.17: Bootstrapping the rails of a series-feedback amplifier from the input of A1. The Zeners have been replaced by resistors

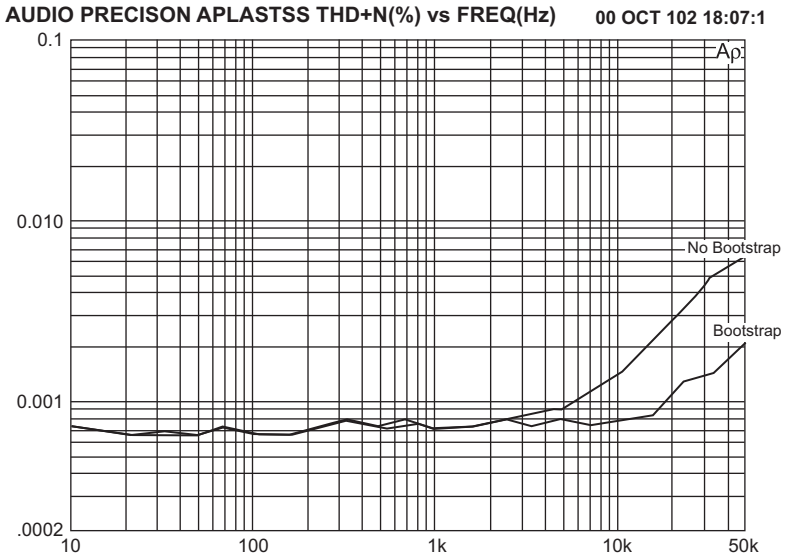


Figure 4.18: The benefit of bootstrapping the rails of a series-feedback amplifier with a gain of 3.2×. The lower trace is essentially that of the THD from the test equipment

www.raguvaran.puzl.com

600 Ω loads. Noise is very low, and the balance of voltage and current noise in the input stage is well matched to moving-magnet phono cartridges; in many applications discrete devices give no significant advantage. Large-quantity production has brought the price down to a point where a powerful reason is required to pick any other device.

The 5532 is not, however, perfect. It suffers common-mode distortion. It has high bias and offset currents at the inputs, as an inevitable result of using a bipolar input stage (for low noise) without any sort of bias-cancellation circuitry. The 5532 is not in the forefront for DC accuracy, though it's not actually that bad. The offset voltage spec is 0.5 mV typical, 4 mV max, compared with 3 mV typical, 6 mV max for the popular TL072. I have actually used 5532s to replace TL072s when offset voltage was a problem, but the increased bias current was acceptable.

With horrible inevitability, the very popularity and excellent technical performance of the 5532 has led to it being criticized by subjectivists who have contrived to convince themselves that they can tell op-amps apart by listening to music played through them. This always makes me laugh, because there is probably no music on the planet that has not passed through a hundred or more 5532s on its way to the consumer.

In some applications, such as low-cost mixing consoles, bipolar-style bias currents are a real nuisance because keeping them out of EQ pots to prevent scratching noises requires an inappropriate number of blocking capacitors. There are plenty of JFET-input op-amps around with negligible bias currents, but there is no obviously superior device that is the equivalent of the 5532. The TL072 has been used in this application for many years but its HF linearity is not first-class and distortion across the band deteriorates badly as output loading increases. However, the op-amps in many EQ sections work in the shunt-feedback configuration with no CM voltage on the inputs, and this reduces the distortion considerably. When low bias currents are needed with superior performance then the OPA2134 is often a good choice, though it is at least four times as expensive as the TL072.

The question of op-amp selection is examined in much more detail in the rest of this chapter, where the most popular types are surveyed.

Op-Amps Surveyed: BJT Input Types

The rest of this chapter looks at some op-amp types and examines their performance, with the 5532 the usual basis for comparison. The parts shown here are not necessarily intended as audio op-amps, though some, such as the OP275 and the OPA2134, were specifically designed as such. They have, however, all seen use, in varying numbers, in audio applications. Bipolar input op-amps are dealt with first.

www.raguvaran.puzl.com

The LM741 Op-Amp

The LM741 is only included here for its historical interest; in its day it was a most significant development and, to my mind, the first really practical op-amp. It was introduced by Fairchild in 1968 and is considered a second-generation op-amp, the 709 being first generation.

The LM741 had (and indeed has) effective short-circuit protection and internal compensation for stability at unity gain, and was much easier to make work in a real circuit than its predecessors. It was clear that it was noisy compared with discrete circuitry, and you sometimes had to keep the output level down if slew limiting was to be avoided, but with care it was usable in audio. Probably the last place the LM741 lingered was in the integrators of state-variable EQ filters, where neither indifferent noise performance nor poor slewing capability is a serious problem (see Chapter 10 for more details on this application). The LM741 is a single op-amp. The dual version is the LM747.

Figure 4.19 shows a region between 100 Hz and 4 kHz where distortion rises at 6dB/octave. This is the result of the usual dominant-pole Miller compensation scheme. When slew limiting begins, the slope increases and THD rises rapidly with frequency.

The NE5532/5534 Op-Amp

The 5532 is a low-noise, low-distortion bipolar dual op-amp, with internal compensation for unity-gain stability. The 5534 is a single version internally compensated for gains down to 3, and an external compensation capacitor can be added for unity-gain stability; 22 pF is the usual value. The common-mode range of the inputs is a healthy ±13 V, with no phase

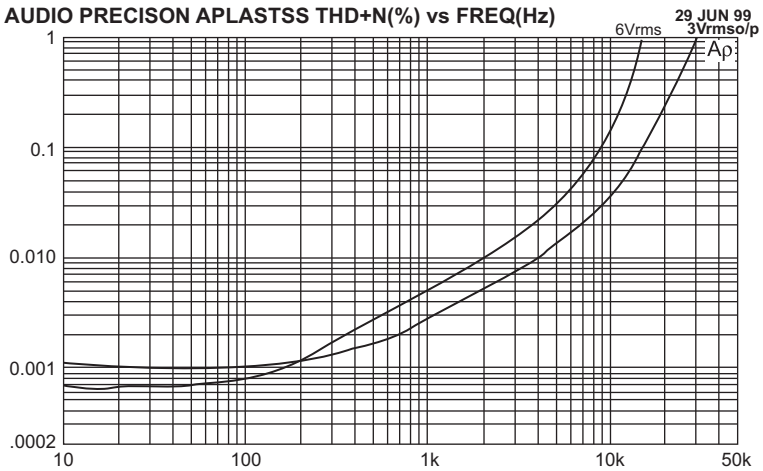


Figure 4.19: The THD performance of an LM741 working at a gain of 3×, on ±15 V rails, giving 3 and 6 Vrms outputs, with no load. At 6 Vrms, slew distortion exceeds 1% before 20 kHz is reached; there is visible slew limiting in the waveform. THD is, however, very low at 100 Hz, due to the high NFB factor at low frequencies

www.raguvaran.puzl.com

inversion problems if this is exceeded. The 5532 has a distinctly higher power consumption than the TL072, drawing approximately 4 mA per op-amp section when quiescent. The DIL version runs perceptibly warm when quiescent on ±17 V rails.

Figure 4.20 shows that the 5532 deals well with loads up to its maximum 500 Ω. Its distortion performance is studied in detail in the section above on common-mode distortion.

The 5534/5532 has bipolar transistor input devices. This means it gives low noise with low source resistances, but draws a relatively high bias current through the input pins. The input devices are NPN, so the bias currents flow into the chip from the positive rail. If an input is fed through a significant resistance then the input pin will be more negative than ground due to the voltage drop caused by the bias current. The inputs are connected together with back-to-back diodes for reverse-voltage protection, and should not be forcibly pulled to different voltages. The 5532 is intended for linear operation, and using it as a comparator is not recommended.

As can be seen from Figure 4.20, the 5532 is almost distortion free, even when driving the maximum 500 Ω load. The internal circuitry of the 5532 has never been publicly explained, but appears to consist of nested Miller loops that permit high levels of internal negative feedback. The 5532 is the dual of the 5534, and is more commonly used than the single as it is cheaper per op-amp and does not require an external compensation capacitor when used at unity gain.

The 5532 and 5534 type op-amps require adequate supply decoupling if they are to remain stable, otherwise they appear to be subject to some sort of internal oscillation that degrades linearity without being visible on a normal oscilloscope. The essential requirement is that the

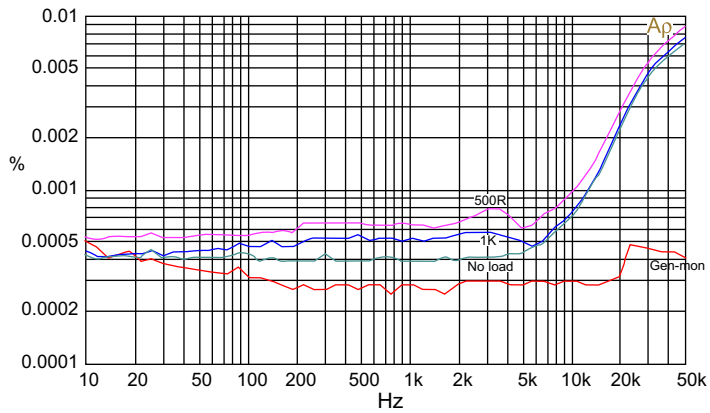


Figure 4.20: Distortion is very low from the 5532, though loading makes a detectable difference. Here it is working in series feedback mode at the high level of 10 Vrms with 500 Ω, 1 kΩ loads and no load. The 'Gen-mon' trace is the output of the distortion analyzer measured directly. Gain of 3.2×, supply ±18 V

www.raguvaran.puzl.com

positive and negative rails should be decoupled with a 100 nF capacitor between them, at a distance of not more than a few millimeters from the op-amp; normally one such capacitor is fitted per package as close to it as possible. It is *not* necessary, and often not desirable, to have two capacitors going to ground; every capacitor between a supply rail and ground carries the risk of injecting rail noise into the ground.

Deconstructing the 5532

To the best of my knowledge, virtually nothing has been published about the internal operation of the 5532. This is surprising given its unique usefulness as a high-quality audio op-amp. I believe the secret of the 5532's superb linearity is the use of nested negative feedback inside the circuit, in the form of traditional Miller compensation.

Figure 4.21 shows the only diagram of the internal circuitry that has been released; the component and node numbers are mine. This has been in the public domain for at least 20 years, so I hope no one is going to object to my impertinent comments on it. The circuit initially looks like a confusing sea of transistors, and there is even a solitary JFET lurking in there, but it breaks down fairly easily. There are three voltage-gain stages, plus a unity-gain output stage to increase drive capability. This has current-sensing overload protection. There is also a fairly complex bias generator that establishes the operating currents in the various stages.

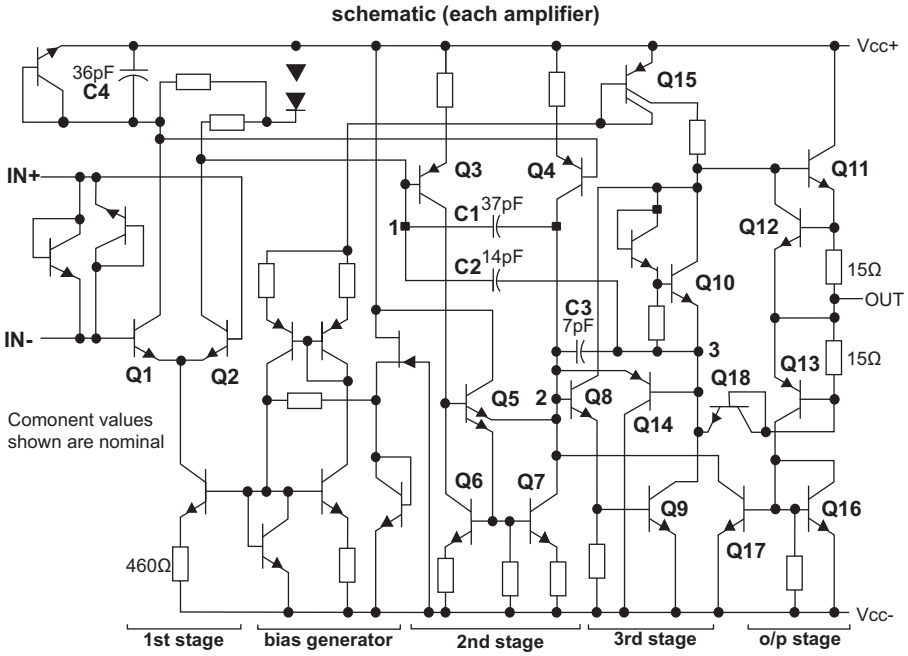


Figure 4.21: The internal circuitry of the 5532

www.raguvaran.puzl.com

In all conventional op-amps there are two differential input signals that have to be subtracted to create a single output signal, and the node at which this occurs is called the 'phase-summing point'.

Q1, Q2 make up the input differential amplifier. They are protected against reverse biasing by the diode-connected transistors across the input pins. Note there are no emitter degeneration resistors, which would linearize the input pair at the expense of degrading noise. Presumably high open-loop gain (note there are three gain stages, whereas a power amplifier normally only has two) means that the input pair is handling very small signal levels, so its distortion is not a problem.

Q3, Q4 make up the second differential amplifier; emitter degeneration is now present. Phase summing occurs at the output of this stage at node 2. C1 is the Miller capacitor around this stage, from node 2 to node 1. Q5, Q6, Q7 are a Wilson current mirror, which provides a driven current source as the collector load of Q4. The function of C4 is obscure but it appears to balance C1 in some way.

The third voltage-amplifier stage is basically Q9 with split-collector transistor Q15 as its current-source load. Q8 increases the basic transconductance of the stage, and C3 is the Miller capacitor around it, feeding from node 3 to node 2 – note that this Miller loop does not include the output stage. Things are a bit more complicated here as it appears that Q9 is also the sink half of the Class-B output stage. Q14 looks very mysterious as it seems to be sending the output of the third stage back to the input; possibly it's some sort of clamp to ensure clean clipping, but to be honest I haven't a clue. Q10 plus associated diode generates the bias for the Class-B output stage, just as in a power amplifier.

The most interesting signal path is the semi-local Miller loop through C2, from node 3 to node 1, which encloses both the second and third voltage amplifiers; each of these has its own local Miller feedback, so there are two nested layers of internal feedback. This is probably the secret of the 5532's low distortion.

Q11 is the source side of the output stage and, as mentioned above, Q9 appears to be the sink. Q12, Q13 implement overcurrent protection. When the voltage drop across the 15 Ω resistor becomes too great, Q12 turns on and shunts base drive away from Q11. In the negative half-cycle, Q13 is turned on, which in turn activates Q17 to shunt drive away from Q8.

The biasing circuit shows an interesting point. Bipolar bias circuits tend not to be self-starting; no current flows anywhere until some flows somewhere, so to speak. Relying on leakage currents for starting is unwise, so here the depletion-mode JFET provides a circuit element that is fully on until you bias it off, and can be relied upon to conduct as the power rails come up from zero.

The LM4562 Op-Amp

The LM4562 is a new op-amp, which first became freely available at the beginning of 2007. It is a National Semiconductor product. It is a dual op-amp – there is no single or quad version. It costs about 10 times as much as a 5532.

www.raguvaran.puzl.com

The input noise voltage is typically 2.7 nV/$\sqrt{\text{Hz}}$, which is substantially lower than the 4 nV/$\sqrt{\text{Hz}}$ of the 5532. For suitable applications with low source impedances this translates into a useful noise advantage of 3.4 dB. The bias current is 10 nA typical, which is very low and would normally imply that bias cancellation, with its attendant noise problems, was being used. However, in my testing I have seen no sign of excess noise, and the data sheet is silent on the subject. No details of the internal circuitry have been released so far, and quite probably never will be.

It is not fussy about decoupling and, as with the 5532, 100 nF across the supply rails close to the package should ensure HF stability. The slew rate is typically ±20 V/µs, more than twice as quick as the 5532.

The first THD plot in Figure 4.22 shows the LM4562 working at a closed-loop gain of 2.2× in shunt-feedback mode, at a high level of 10 Vrms. The top of the THD scale is 0.001%, something you will see with no other op-amp in this survey. The no-load trace is barely distinguishable from the AP SYS-2702 output, and even with a heavy 500 Ω load driven at 10 Vrms there is only a very small amount of extra THD, reaching 0.0007% at 20 kHz.

Figure 4.23 shows the LM4562 working at a gain of 3.2× in series-feedback mode, both modes having a noise gain of 3.2×. The extra distortion from 500 Ω loading is barely detectable.

For Figures 4.23 and 4.24 the feedback resistances were 2k2 and 1 kΩ, so the minimum source resistance presented to the inverting input is 687 Ω. In Figure 4.25 extra source resistances were then put in series with the input path (as was done with the 5532 in the section above on common-mode distortion) and this revealed a remarkable property of the LM4562 – it is much

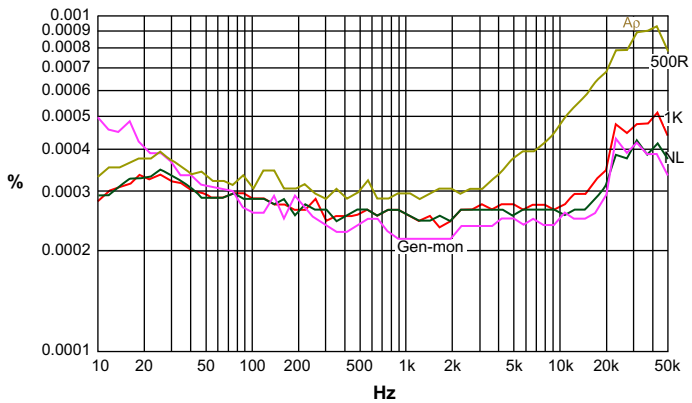


Figure 4.22: The LM4562 in shunt-feedback mode, with 1 kΩ, 2k2 feedback resistors giving a gain of 2.2×. Shown for no load (NL) and 1 kΩ, 500 Ω loads. Note the vertical scale ends at 0.001% this time. Output level is 10 Vrms, ±18 V supply rails

www.raguvaran.puzl.com

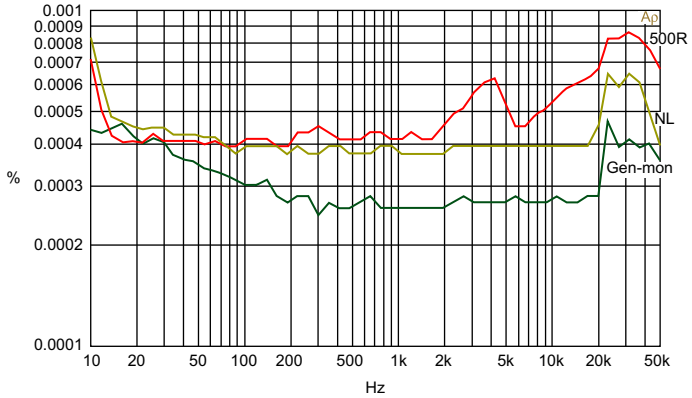


Figure 4.23: The LM4562 in series-feedback mode, with 1 kΩ, 2k2 feedback resistors giving a gain of 3.2×. No load (NL) and 500 Ω load. Output 10 Vrms, ±18 V supply rails

more resistant to common-mode distortion than the 5532. At 10 Vrms and 10 kHz, with a 10 kΩ source resistance the 5532 generates 0.0014% THD (see Figure 4.6) but the LM4562 gives only 0.00046% under the same conditions. I strongly suspect that the LM4562 has a more sophisticated input stage than the 5532, probably incorporating cascoding to minimize the effects of common-mode voltages.

Note that only the rising curves to the right represent actual distortion. The raised levels of the horizontal traces at the LF end are due to Johnson noise from the extra series resistance.

It has taken an unbelievably long time – nearly 30 years – for a better audio op-amp than the 5532 to come along, but at last it has happened. The LM4562 is superior in just about every

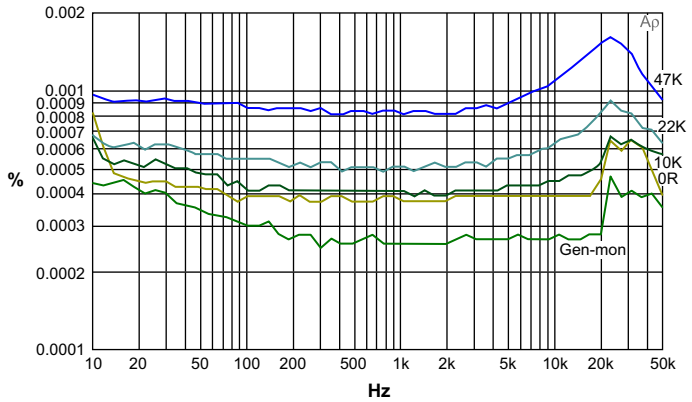


Figure 4.24: The LM4562 in series-feedback mode, gain 3.2×, with varying extra source resistance in the input path. The extra distortion is much lower than for the 5532. Output 10 Vrms, ±18 V supply rails

www.raguvaran.puzl.com

parameter, but it has much higher current noise. At present it also has a much higher price, but hopefully that will change.

The AD797 Op-Amp

The AD797 (Analog Devices) is a single op-amp with very low voltage noise and distortion. It appears to have been developed primarily for the cost-no-object application of submarine sonar, but it works very effectively with normal audio – if you can afford to use it. The cost is something like 20 times that of a 5532. No dual version is available, so the cost ratio per op-amp section is 40 times.

This is a remarkably quiet device in terms of voltage noise, but current noise is correspondingly high due to the high currents in the input devices. Early versions appeared to be rather difficult to stabilize at HF, but the current product is no harder to apply than the 5532. Possibly there has been a design tweak, or on the other hand my impression may be wholly mistaken.

The AD797 incorporates an ingenious feature for internal distortion cancellation. This is described on the manufacturer's data sheet. Figure 4.25 shows that it works effectively.

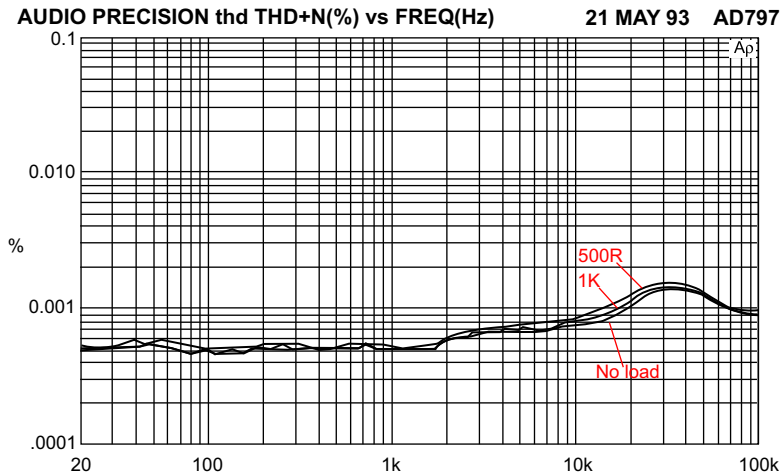


Figure 4.25: AD797 THD into loads down to 500 Ω, at 7.75 Vrms. Output is virtually indistinguishable from input. Series feedback, but no CM problems. Gain = 3.2×

The OP27 Op-Amp

The OP27 from Analog Devices is a bipolar-input single op-amp primarily designed for low noise and DC precision. It was not intended for audio use, but in spite of this it is frequently recommended for such applications as RIAA and tape head preamps. This is unfortunate, because while at first sight it appears that the OP27 is quieter than the 5534/5532, as the e_n is 3.2 nV/$\sqrt{\text{Hz}}$ compared with 4 nV/$\sqrt{\text{Hz}}$ for the 5534, in practice it is usually slightly noisier.

www.raguvaran.puzl.com

This is because the OP27 is in fact optimized for DC characteristics, and so has input bias-current cancellation circuitry that generates common-mode noise. When the impedances on the two inputs are very different – which is the case in RIAA preamps – the CM noise does not cancel, and this appears to degrade the overall noise performance significantly.

For a bipolar input op-amp, there appears to be a high-level common-mode input distortion, enough to bury the output distortion caused by loading (see Figures 4.26 and 4.27). It is likely that this too is related to the bias-cancellation circuitry, as it does not occur in the 5532.

The maximum slew rate is low compared with other op-amps, being typically 2.8 V/µs. However, this is not the problem it may appear. This slew rate would allow a maximum amplitude at 20 kHz of 16 Vrms, if the supply rails permitted it. I have never encountered any particular difficulties with decoupling or stability of the OP27.

The OP270 Op-Amp

The OP270 from Analog Devices is a dual op-amp, intended as a 'very-low-noise precision operational amplifier', in other words combining low noise with great DC accuracy. The input offset voltage is an impressive 75 µV maximum. It has bipolar inputs with a bias-current cancellation system; the presence of this is shown by the 15 nA bias current spec, which is 30 times less than the 500 nA taken by the 5534, which lacks this feature. It will degrade the noise performance with unequal source resistances, as it does in the OP27. The input transistors are protected by back-to-back diodes.

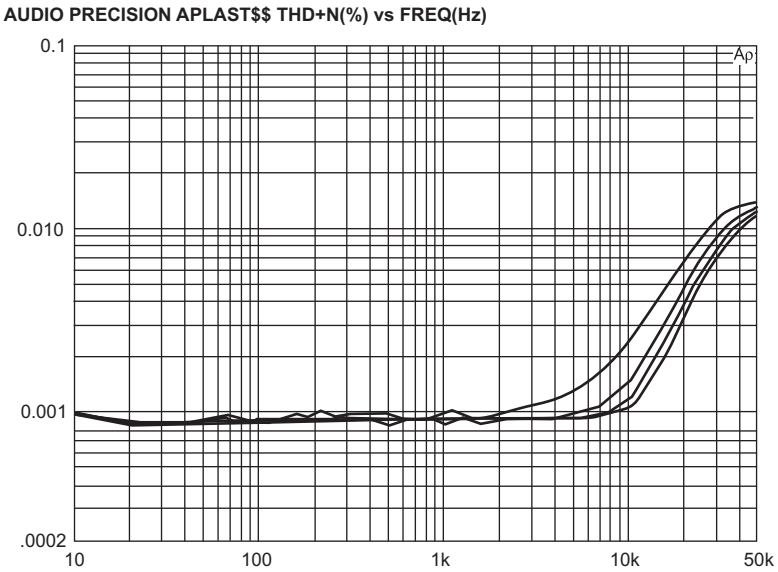


Figure 4.26: OP27 THD in shunt-feedback mode with varying loads. This op-amp accepts even heavy (1 kΩ) loading gracefully

www.raguvaran.puzl.com

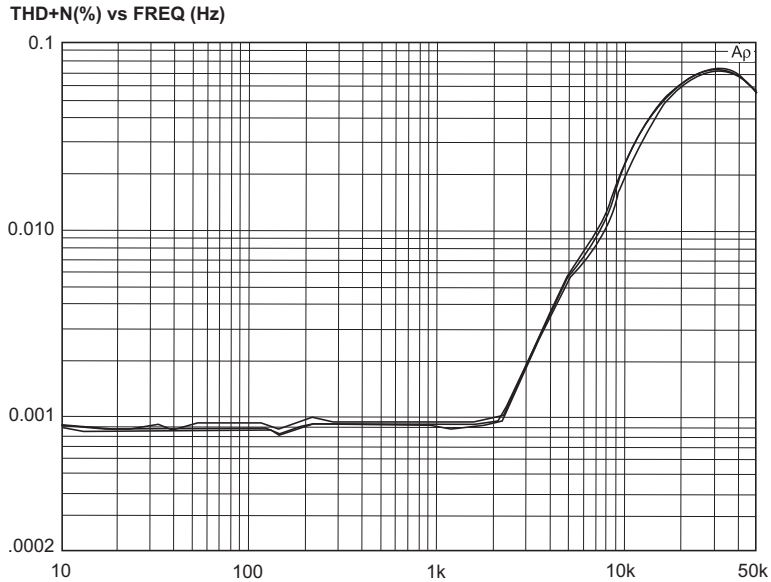


Figure 4.27: OP27 THD in series-feedback mode. The common-mode input distortion completely obscures the output distortion

The OP270 distortion performance suffers badly when driving even modest loads. See Figures 4.28 and 4.29. The slew rate is a rather limited 2.4 V/µs, which is only just enough for a full output swing at 20 kHz. Note also that this is an expensive op-amp, costing something like 25 times as much as a 5532; precision costs money. Unless you have a real need for DC accuracy, this part is not recommended.

The OP275 Op-Amp

The Analog Devices OP275 is one of the few op-amps specifically marketed as an audio device. Its most interesting characteristic is the Butler input stage, which combines bipolar and JFET devices. The idea is that the bipolars give accuracy and low noise, while the JFETs give speed and 'the sound quality of JFETs'. That final phrase is not a happy thing to see on a data sheet from a major manufacturer; the sound of JFETs (if any) would be the sound of high distortion. Just give us the facts, please.

The OP275 is a dual op-amp; no single version is available. It is quite expensive, about six times the price of a 5532, and its performance in most respects is inferior. It is noisier, has higher distortion, and does not like heavy loads. See Figures 4.30 and 4.31. The CM range is only about two-thirds of the voltage between the supply rails, and I_{bias} is high due to the BJT part of the input stage. Unless you think there is something magical about the BJT/JFET input stage – and I am quite sure there is not – it is probably best avoided.

www.raguvaran.puzl.com

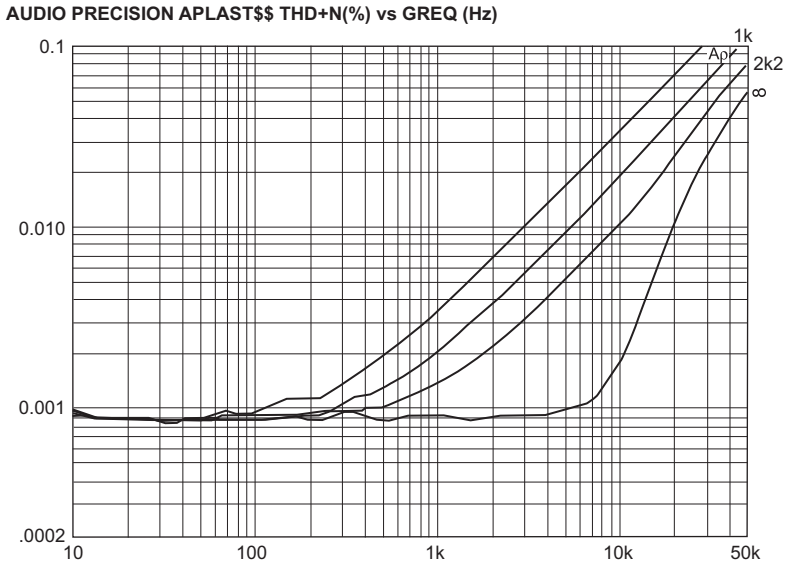


Figure 4.28: OP270 THD in shunt-feedback mode. Linearity is severely degraded even with a 2k2 load

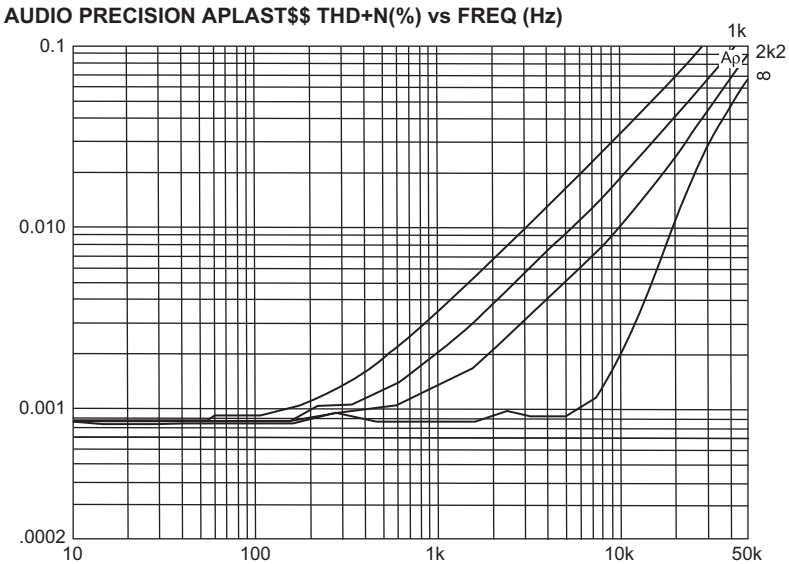


Figure 4.29: OP270 THD in series-feedback mode. This looks the same as in Figure 4.28 so CM input distortion appears to be absent

The THD at 10 kHz with a 600 Ω load is 0.0025% for shunt and 0.009% for series feedback; there is significant CM distortion in the input stage, which is almost certainly coming from the JFETs. (I appreciate the output levels are not the same but I think this only accounts for a small part of the THD difference.) Far from adding magical properties to the input stage, the JFETs seem to be just making it worse.

www.raguvaran.puzl.com

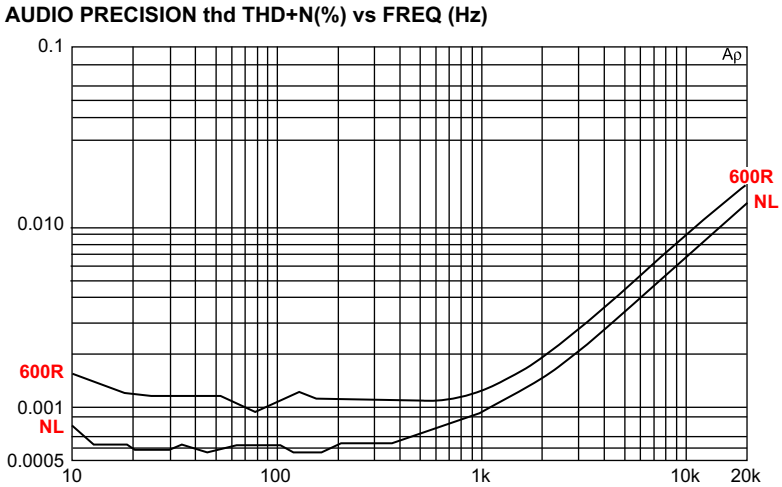


Figure 4.30: An OP275 driving 7.75 Vrms into no load and 600 Ω. THD below 1 kHz is definitely non-zero with the 600 Ω load. Series feedback, gain 3.2×

Op-Amps Surveyed: JFET Input Types

Op-amps with JFET inputs tend to have higher voltage noise and lower current noise than BJT-input types, and therefore give a better noise performance with high source resistances. Their very low bias currents often allow circuitry to be simplified.

The TL072 Op-Amp

The TL072 is one of the most popular op-amps, having very-high-impedance inputs, with effectively zero bias and offset currents. The JFET input devices give their best noise performance at medium impedances, in the range 1–10 kΩ. The TL072 has a modest power consumption at typically 1.4 mA per op-amp section, which is significantly less than with the 5532. The slew rate is higher than for the 5532, at 13 V/µs against 9 V/µs. The TL072 is a dual op-amp. There is a single version called the TL071, which has offset null pins.

However, the TL072 is not THD free in the way the 5532 is. In audio usage, distortion depends primarily upon how heavily the output is loaded. The maximum loading is a trade-off between quality and circuit economy, and I would put 2 kΩ as the lower limit. This op-amp is not the first choice for audio use unless the near-zero bias currents (which allow circuit economies by making blocking capacitors unnecessary), the low price, or the modest power consumption are dominant factors.

It is a quirk of this device that the input common-mode range does not extend all the way between the rails. If the common-mode voltage gets to within a couple of volts of the V− rail, the op-amp suffers phase reversal and the inputs swap their polarities. There may be really horrible clipping, where the output hits the bottom rail and then shoots up to hit the top one, or the stage may simply latch up until the power is turned off.

www.raguvaran.puzl.com

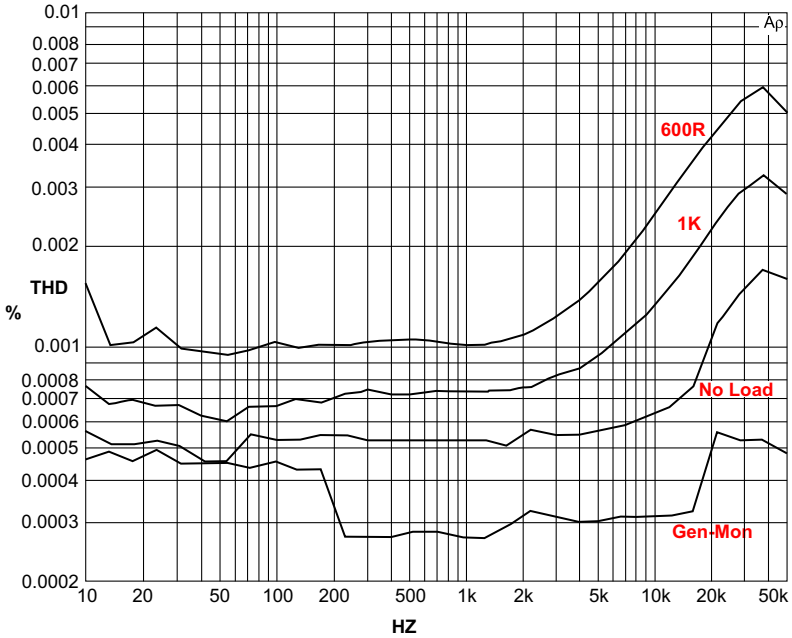


Figure 4.31: OP275 driving 5 Vrms into 1 k and 600 Hz Ω. Shunt feedback, gain 2.2×, but note noise gain was set to 3.2× as for the series case. The 'Gen-Mon' trace shows the distortion of the AP System 2 generator; the steps at 200 Hz and 20 kHz are artefacts generated by internal range switching

TL072s are relatively relaxed about supply-rail decoupling, though they will sometimes show very visible oscillation if they are at the end of long thin supply tracks. One or two rail-to-rail decoupling capacitors (e.g. 100 nF) per few centimeters is usually sufficient to deal with this, but normal practice is to not take chances, and allow one capacitor per package as with other op-amps.

Because of common-mode distortion, a TL072 in shunt configuration is always more linear. In particular compare the results for 3k3 load in Figures 4.32 and 4.33. At heavier loadings the difference is barely visible because most of the distortion is coming from the output stage.

TL072/71 op-amps are prone to HF oscillation if faced with significant capacitance to ground on the output pin; this is particularly likely when they are used as unity-gain buffers with 100% feedback. A few inches of track can sometimes be enough. This can be cured by an isolating resistor, in the 47–75 Ω range, in series with the output, placed at the op-amp end of the track.

The TL052 Op-Amp

The TL052 from Texas Instruments was designed to be an enhancement of the TL072, and so is naturally compared with it. Most of the improvements are in the DC specifications. The

www.raguvaran.puzl.com

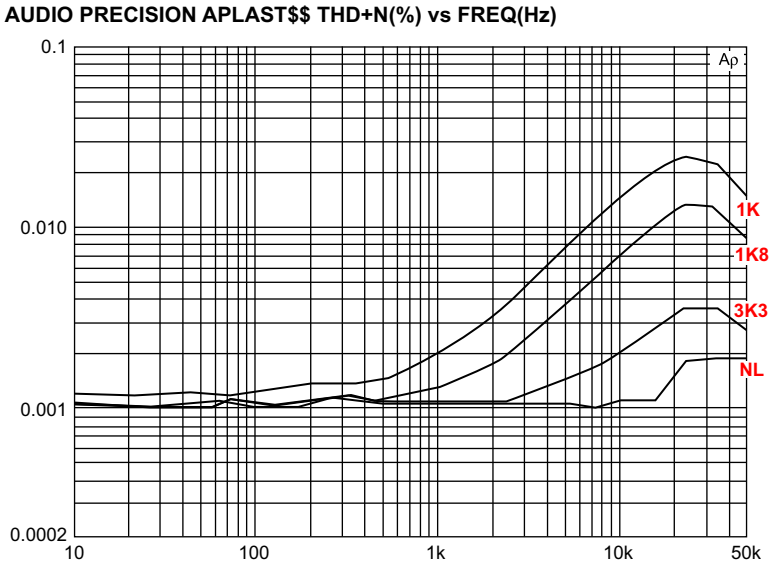


Figure 4.32: Distortion versus loading for the TL072, with various loads. Shunt-feedback configuration eliminates CM input distortion. Output level 3 Vrms, gain 3.2×, rails ±15 V. No output load except for the feedback resistor. The no-load plot is indistinguishable from that of the test-gear alone. Distortion always gets worse as the loading increases. This factor, together with the closed-loop NFB factor, determines the THD

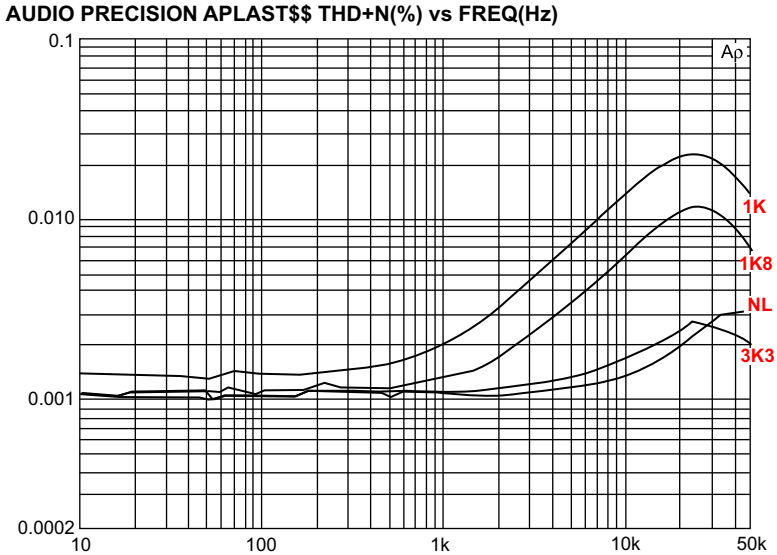


Figure 4.33: Distortion versus loading for the TL072, with various loads. Series-feedback configuration, output level 3 Vrms, gain 3.2×, rails ±15 V. Distortion at 10 kHz with no load is 0.0015% compared with 0.0010% for the shunt configuration. This is due to the 1 Vrms CM signal on the inputs

www.raguvaran.puzl.com

offset voltage is 0.65 mV typical, 1.5 mV max, compared with the TL072's 3 mV typical, 10 mV max. It has half the bias current of the TL072. This is very praiseworthy, but rarely of much relevance to audio.

The distortion, however, *is* important, and this is worse rather than better. THD performance is rather disappointing. The unloaded THD is low, as shown in Figure 4.34, in series-feedback mode. As usual, practical distortion depends very much on how heavily the output is loaded. Figure 4.35 shows that it deteriorates badly for loads of less than 4k7.

The slew rate is higher than for the TL072 (18 against 13 V/µs) but the lower figure is more than adequate for a full-range output at 20 kHz, so this enhancement is of limited interest. The power consumption is higher, typically 2.3 mA per op-amp section, which is almost twice that of the TL072. Like the TL072, the TL052 is relatively relaxed about supply-rail decoupling. At the time of writing (2009) the TL052 costs at least twice as much as the TL072.

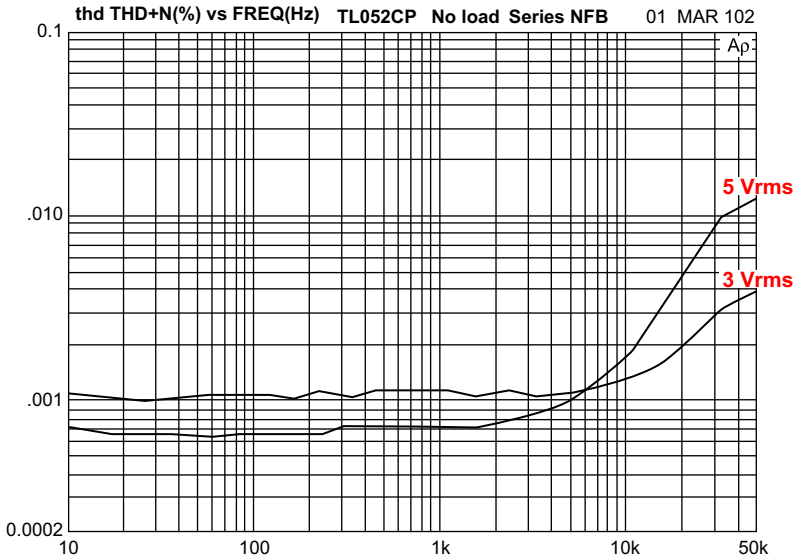


Figure 4.34: Distortion versus frequency at two output levels for the TL052CP, with no load. Series feedback

The OPA2134 Op-Amp

The OPA2134 is a Burr-Brown product, the dual version of the OPA134. The manufacturer claims it has superior sound quality, due to its JFET input stage. Regrettably, but not surprisingly, no evidence is given to back up this assertion. The input noise voltage is 8 nV/$\sqrt{\text{Hz}}$, almost twice that of the 5532. The slew rate is typically ±20 V/µs, which is ample. The OPA2134 does not appear to be optimized for DC precision, the typical offset voltage being

www.raguvaran.puzl.com

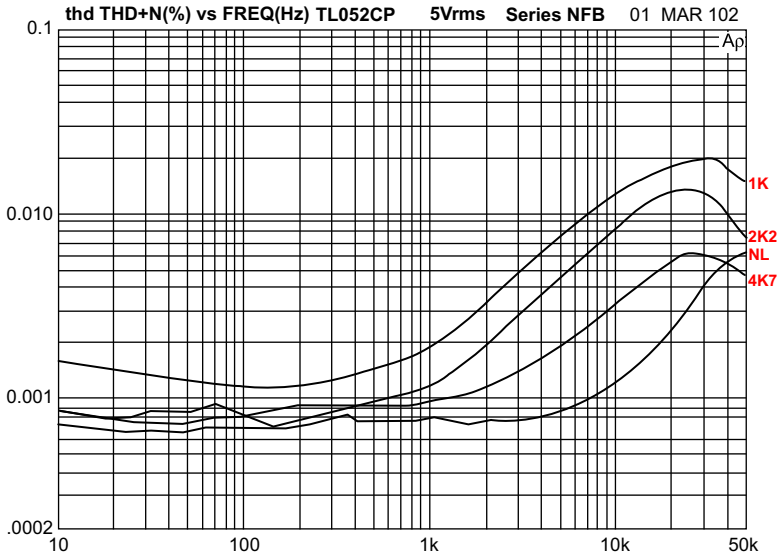


Figure 4.35: Distortion of the TL052 at 5 Vrms output with various loads. At 1 kΩ and 2k2 loading the residual is all crossover distortion at 1 kHz. Gain 3.2×, non-inverting (series feedback)

±1 mV, but this is usually good enough for audio work. I have used it many times as a DC servo in power amplifiers, the low bias currents allowing high resistor values and correspondingly small capacitors.

The OPA2134 does not show phase reversal anywhere in the common-mode range, which immediately marks it as superior to the TL072.

The two THD plots in Figures 4.36 and 4.37 show the device working at a gain of 3× in both shunt and series-feedback modes. It is obvious that a problem emerges in the series plot, where the THD is higher by about three times at 5 Vrms and 10 kHz. This distortion increases with level, which immediately suggests common-mode distortion in the input stage. Distortion increases with even moderate loading, see Figure 4.38.

This is a relatively modern and sophisticated op-amp. When you need JFET inputs (usually because significant input bias currents would be a problem) this definitely beats the TL072; it is, however, four to five times more expensive.

The OPA604 Op-Amp

The OPA604 from Burr-Brown is a single JFET-input op-amp, which claims to be specially designed to give low distortion. The simplified internal circuit diagram in the data sheet includes an enigmatic box intriguingly labeled ‘Distortion Rejection Circuitry’. This apparently ‘linearizes the open-loop response and increases voltage gain’, but no details as to how are

www.raguvaran.puzl.com

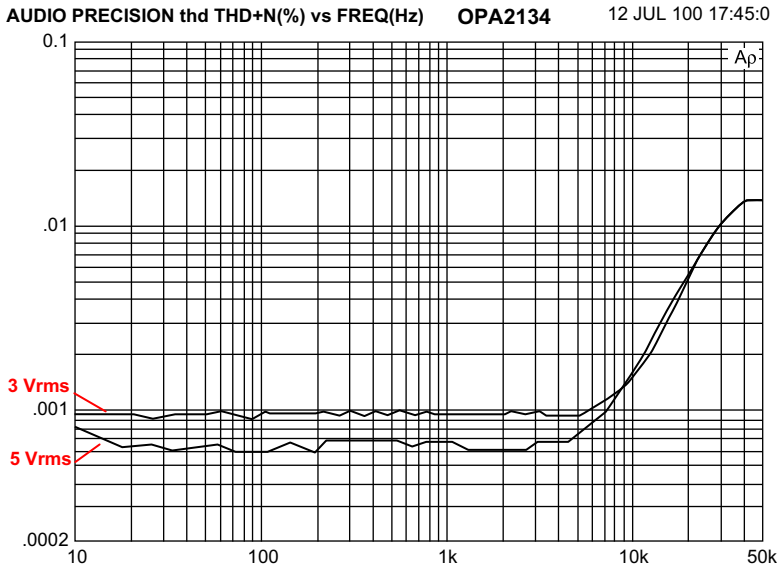


Figure 4.36: The OPA2134 working in shunt-feedback mode. The THD is below the noise until frequency reaches 10 kHz; it appears to be lower at 5 Vrms simply because the noise floor is relatively lower.

given; whatever is in there appears to have been patented so it ought to be possible to track it down. However, despite this, the distortion is not very low even with no load (see Figure 4.39), and is markedly inferior to the 5532's. The OPA604 is not optimized for DC precision, the typical offset voltage being ±1 mV. The OPA2604 is the dual version, which omits the offset null pins.

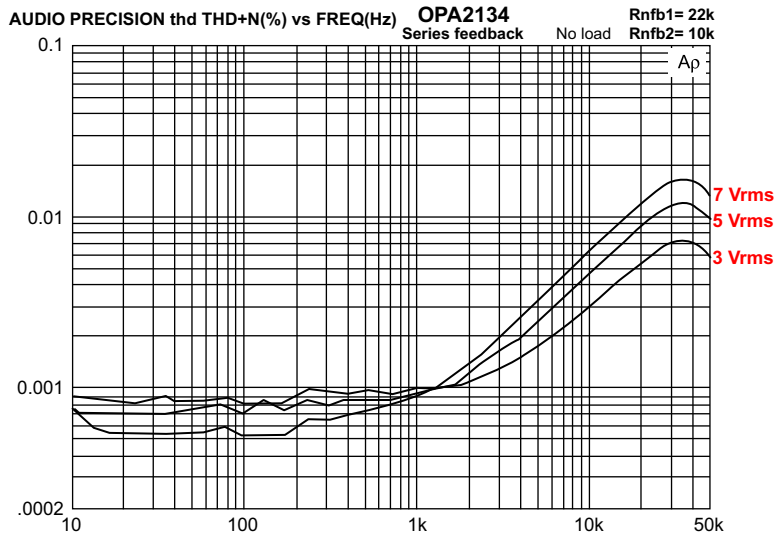


Figure 4.37: The OPA2134 in series-feedback mode. Note much higher distortion at HF

www.raguvaran.puzl.com

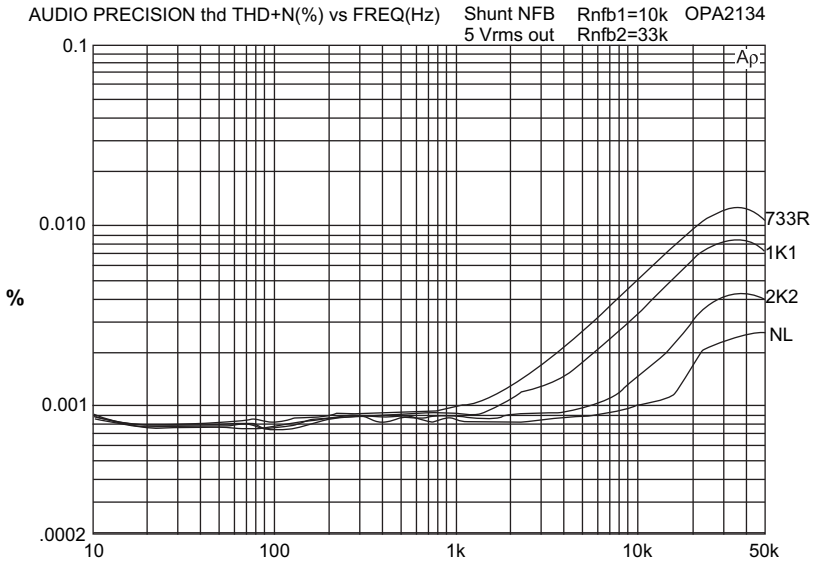


Figure 4.38: The OPA2134 in shunt-feedback mode (to remove input CM distortion) and with varying loads on the output. As usual, more loading makes linearity worse. Output 5 Vrms, gain = 3.3×

The data sheet includes a discussion that attempts to show that JFET inputs produce a more pleasant type of distortion than BJT inputs. This unaccountably omits the fact that the much higher transconductance of BJTs means that they can be linearized by emitter degeneration so that they produce far less distortion of whatever type than a JFET input [6]. Given that the OPA604 costs five times as much as a 5532, it is not very clear under what circumstances this op-amp would be a good choice.

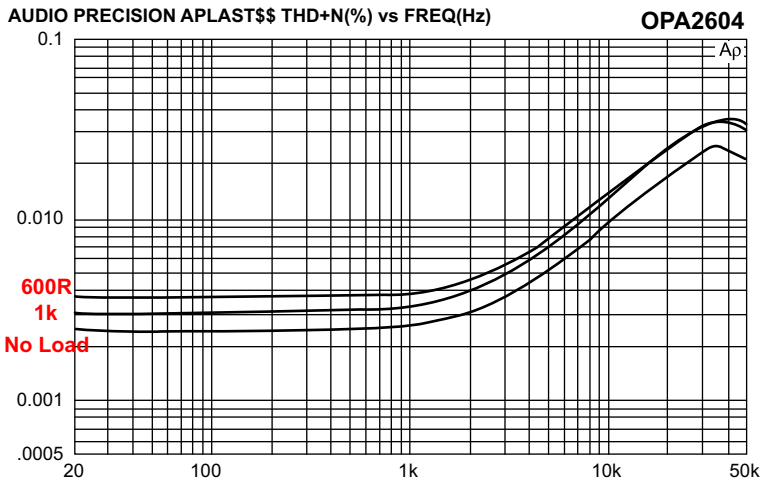


Figure 4.39: An OPA2604 driving various loads at 7.75 Vrms. Series feedback, gain = 3.2×

www.raguvaran.puzl.com

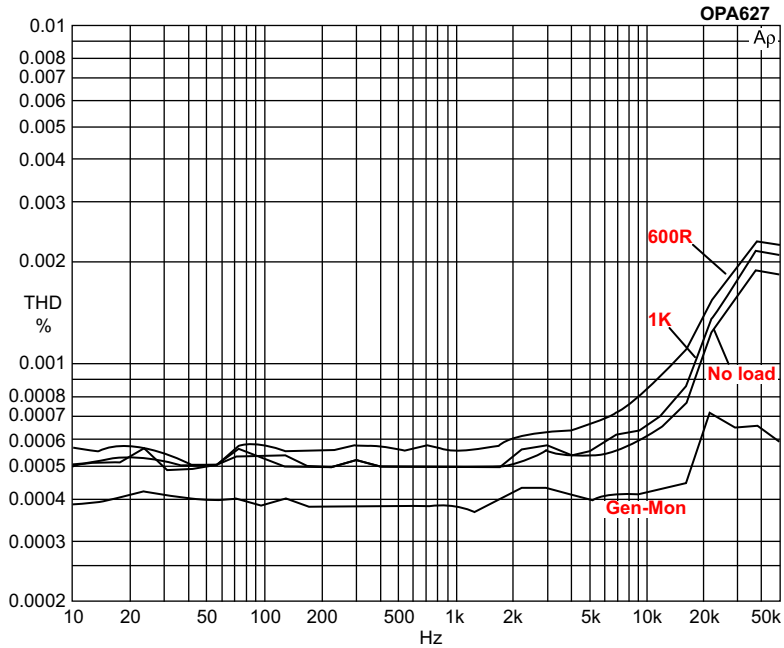


Figure 4.40: OPA627 driving the usual loads at 5 Vrms. Series feedback, gain = 3.2×. 'Gen-Mon' is the test-gear output

The OPA627 Op-Amp

The OPA627 from Burr-Brown is a laser-trimmed JFET-input op-amp with excellent DC precision, the input offset voltage being typically ±100 μV. The distortion is very low, even into a 600 Ω load, though it is increased by the usual common-mode distortion when series feedback is used.

The OPA627 is a single op-amp and no dual version is available. The OPA637 is a decompensated version only stable for closed-loop gains of 5 or more. This op-amp makes a brilliant DC servo for power amplifiers, if you can afford it; it costs about 50 times as much as a 5532, which is 100 times more per op-amp section, and about 20 times more per op-amp than the OPA2134, which is my usual choice for DC servo work.

The current noise i_n is very low, the lowest of any op-amp examined in this book, apparently due to the use of Difet (dielectrically isolated JFET) input devices, and so it will give a good noise performance with high source resistances. Voltage noise is also very respectable at 5.2 nV/$\sqrt{\text{Hz}}$, only fractionally more than that of the 5532.

The series-feedback case barely has more distortion than the shunt one, and only at the extreme HF end. It appears that the Difet input technology also works well to prevent input non-linearity and CM distortion. See Figures 4.40 and 4.41.

www.raguvaran.puzl.com

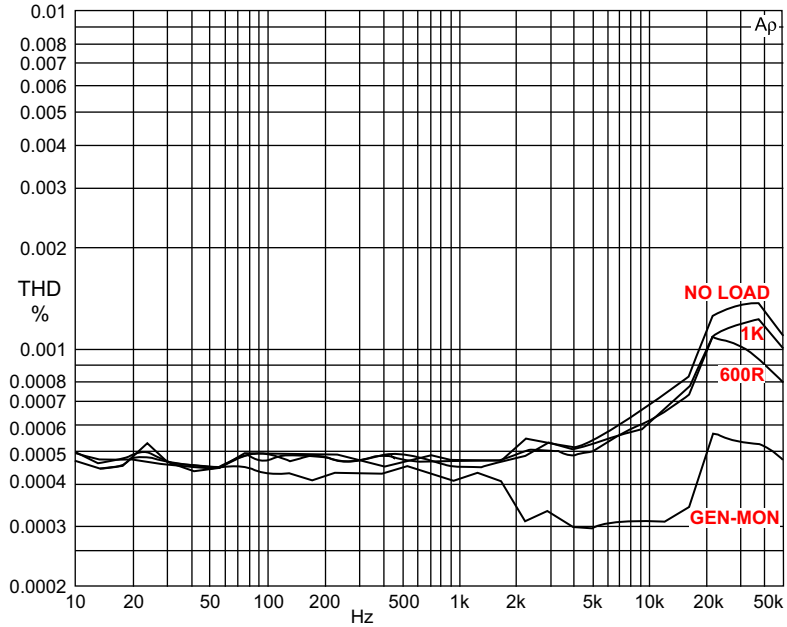


Figure 4.41: OPA627 driving the usual loads at 5 Vrms. Shunt feedback, gain = 2.2× but noise gain = 3.2×. 'Gen-Mon' trace shows the distortion produced by the AP System 2 generator alone

References

[1] A. Blumlein, UK patent 482,470, 1936.
[2] W. Jung (Ed.), Op-Amp Applications Handbook, Newnes, 2006 (Chapter 8).
[3] D. Self, Audio Power Amplifier Design Handbook, fifth ed, Focal Press, 2009, pp. 186–189.
[4] D. Self, Audio Power Amplifier Design Handbook, fifth ed, Focal Press, 2009, p. 96.
[5] W. Jung (Ed.), Op-Amp Applications Handbook, Newnes, 2006, p. 399 (Chapter 5).
[6] D. Self, Audio Power Amplifier Design Handbook, fifth ed, Focal Press, 2009, p. 380.

www.raguvaran.puzl.com

CHAPTER 5

Filters

Analog filter design is an enormous subject, and it is of course quite impossible to cover even its audio aspects in a single chapter. There are many excellent and comprehensive textbooks on filters [1–3] and there would be no point in trying to create another one here. This chapter instead aims to give information on audio applications not found in the standard textbooks.

Filter design is at the root highly mathematical, and it is no accident that all of the common filter characteristics such as Bessel, Butterworth, and Chebyshev are named after mathematicians. Here, however, I am going to avoid the complexities of pole and zero placement etc., and concentrate on practical filter designs that can be adapted for use at different frequencies by a simple process of scaling component values. Because filters are widely used in different kinds of audio system, I make frequent references to information about filters in other chapters.

Filters are either passive or active. Passive or LCR filters use resistors, inductors, and capacitors only. Active filters use resistors, capacitors, and gain elements such as op-amps; active filter technology is usually adopted with the specific intent of avoiding inductors and their well-known limitations. Nevertheless, there are some applications where LCR filters are essential.

Passive Filters

Passive filters do not use active electronics, and this is a crucial advantage in some applications. They are not subject to slew-rate limiting, semiconductor non-linearity, or errors due to falling open-loop gain, and this makes them the best technology for roofing filters. A roofing filter is one that stops out-of-band frequencies before they reach the first stage of electronics, and so prevents RF demodulation and slew limiting. A classic application is in the measurement of Class-D power amplifiers, which emit copious quantities of RF that will greatly upset audio measuring equipment. The answer, as described by Bruce Hofer [4], is a passive LCR roofing filter. There are many excellent text books that describe LCR filter design [1,2], and I am not going into it here.

Small Signal Audio Design; ISBN: 9780240521770
Copyright © 2010 Elsevier Ltd; All rights of reproduction, in any form, reserved.

www.raguvaran.puzl.com

Active Filters

Active filters do not normally use inductors as such, though configurations such as gyrators that explicitly model the action of an inductor are sometimes used. The active element need not be an op-amp; the Sallen-and-Key configuration requires only a voltage-follower, which in some cases can be a simple BJT emitter-follower. Op-amps are usually used nowadays, however. The rest of this chapter deals only with active filters.

Low-Pass Filters

Probably the most common use of low-pass filters is at the output of DACs to remove the high-frequency spurii that remain after oversampling (see Chapter 21 for more on this). They are also used to explicitly define the upper limit of the audio bandwidth in a system at, say, 50 kHz (see the example of a record-cutting amplifier in Chapter 7), though more often this is done by the casual accumulation of a lot of first-order roll-offs in succeeding stages. This is not a duplication of input RF filtering which, as described in Chapter 14, must be passive and positioned before the incoming signals encounter any electronics which can demodulate RF. Low-pass filters are also used in public address (PA) systems to protect power amplifiers and loudspeakers against ultrasonic oscillation in the system.

In what might be called the First Age of Vinyl, a fully equipped preamplifier would certainly have a switchable low-pass filter, called, with brutal frankness, the 'scratch' filter. This, having a slope of 12 or 18 dB/octave, faster than the tone-control stage, and rolling off at a higher frequency around 5–10 kHz, was aimed at suppressing, or at any rate dulling, record surface noise and the inevitable ticks and clicks. Interestingly, preamplifiers today, in the Second Age of Vinyl, rarely have this facility, probably because it requires you to face up to the fact that the reproduction of music from mechanical grooves cut into vinyl is really not very satisfactory. Low-pass filters are of course essential to electronic crossovers for loudspeaker systems.

High-Pass Filters

High-pass filters are widely used. Preamplifiers for vinyl disk usage are commonly fitted with subsonic filtering, often below 10 Hz, to keep disturbances due to record warps and ripples from reaching the loudspeakers. Subsonic noise may affect the linearity of the speaker for the worse as the bass unit is often moved through a substantial part of its mechanical travel; this is particularly true for reflex designs with no cone loading at very low frequencies. If properly designed, this sort of filtering is considered inaudible by most people. In the First Age of Vinyl, preamplifiers sometimes were fitted with a 'rumble filter', which began operations at a higher frequency, typically 35 or 40 Hz, to deal with more severe disk problems, and would not be considered inaudible by anyone. Phono subsonic filtering is comprehensively dealt with in Chapter 7 on moving-magnet preamplifiers.

www.raguvaran.puzl.com

Mixer input channels very often have a switchable high-pass filter at a higher frequency again, usually 100 Hz. This is intended to deal with low-frequency proximity effects with microphones and general environmental rumblings. The slope required to do this effectively is at least 12 dB/octave, putting it outside the capabilities of the EQ section. Some mixer high-pass filters are third-order (18 dB/octave) and fourth-order (24 dB/octave) ones have been used occasionally. Once again, high-pass filters are used in electronic crossovers.

Combined Low-Pass and High-Pass Filters

When both subsonic and ultrasonic filters are required they can sometimes be economically combined into one stage using only one op-amp, to give audio band definition. Filter combination is usually only practicable when the two filter frequencies are widely separated (more on this in Chapter 7).

Bandpass Filters

Bandpass filters are principally used in mixing consoles and stand-alone equalizers (see Chapter 10). The Q required rarely exceeds a value of 5, which can be implemented with relatively simple active filters, such as the multiple-feedback type. Higher Q values or independent control of all the resonance parameters requires the use of the more complex biquad or state-variable filters. Bandpass and notch filters are said to be 'tuneable' if their center frequency can be altered relative easily, say by changing only one component value.

Notch Filters

Notch filters are mainly used in equalizers to deal with narrow peaks in the acoustic response of performance spaces, and in some electronic crossovers, for example [5]. They can also be used to remove a single interfering frequency; once, a long time ago, I was involved with a product that had a slide projector in close proximity to a cassette player. Hi-fi was not the aim, but even so the enormous magnetic field from the projector transformer induced an unacceptable amount of hum into the cassette tape head. Mu-metal only helped a bit, and the fix was a filter that introduced notches at 50 and 100 Hz, working on the '1-bandpass' principle, of which more later.

All-Pass Filters

All-pass filters are so called because they have a flat frequency response, and so pass all frequencies equally. Their point is that they have a phase-shift that *does* vary with frequency,

www.raguvaran.puzl.com

and this is often used for delay correction in electronic crossovers. You may occasionally see a reference to an all-stop filter, which has infinite rejection at all frequencies. This is a filter designer's joke.

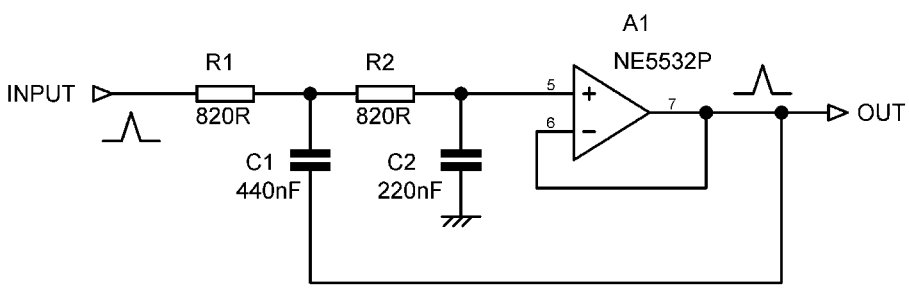
Filter Characteristics

The simple second-order bandpass responses are basically all the same, being completely defined by center frequency, Q, and gain. High-pass and low-pass filter characteristics are much more variable and are selected as a compromise between the need for a rapid roll-off, flatness in the passband, and a clean transient response. The Butterworth (maximally flat) characteristic is the most popular for many applications. Filters with passband ripple, such as the Chebyshev or the Elliptical types, have not found favor for in-band filtering, such as in electronic crossovers, but were once widely used for applications like ninth-order anti-aliasing filters; such filters have mercifully been made obsolete by oversampling. The Bessel characteristic gives a maximally flat group delay (maximally linear phase response) across the passband, and so preserves the waveform of filtered signals, but it has a much slower roll-off than that of the Butterworth.

Sallen-and-Key Filters

The Sallen-and-Key filter configuration was introduced by R.P. Sallen and E.L. Key of the MIT Lincoln Laboratory as long ago as 1955 [6]. It became popular as the only active element required is a unity-gain buffer, so in the days before op-amps were cheap it could be effectively implemented with a simple emitter-follower.

Figure 5.1 shows a second-order low-pass Sallen-and-Key filter with a −3 dB frequency of 624 Hz and a Q of 0.707, together with the pleasingly simple design equations for cut-off (−3 dB)



$$f_0 = \frac{1}{2\pi R\sqrt{C_1 C_2}} \qquad Q = \frac{1}{2}\sqrt{\frac{C_1}{C_2}}$$

Figure 5.1: The classic second-order low-pass Sallen-and-Key filter. Cut-off frequency is 624 Hz, $Q = 0.707$ (critically damped)

www.raguvaran.puzl.com

frequency f_0 and Q. It was part of a fourth-order Linkwitz–Riley electronic crossover for a three-way loudspeaker [7]. The main difference you will notice from textbook filters is that the resistor values are rather low and the capacitor values correspondingly high. This is an example of low-impedance design, where low resistor values minimize Johnson noise and reduce the effect of op-amp current noise and common-mode distortion. The measured noise output is −117.4 dBu. This is after correction by subtracting the test-gear noise floor.

It is important to remember that a Q of 1 does not give the maximally flat Butterworth response; you must use 0.707. Sallen-and-Key filters with a Q of 0.5 are used in second-order Linkwitz–Riley crossovers, but these are not favored because the 12 dB/octave roll-off of the high-pass filter is not steep enough to reduce the excursion of a driver when a flat frequency response is obtained [7].

Sallen-and-Key low-pass filters have a lurking problem. When implemented with op-amps, the response does not carry on falling for ever at the filter slope – instead it reaches a minimum and starts to come back up at 6 dB/octave. This is because the filter action relies on C1 seeing a low impedance to ground, and the impedance of the op-amp output rises with frequency due to falling open-loop gain and hence falling negative feedback. When the circuit of Figure 5.1 is built using a TL072, the maximum attenuation is −57 dB at 21 kHz, rising again and flattening out at −15 dB at 5 MHz. This type of filter should not be used to reject frequencies well above the audio band; a low-pass version of the multiple-feedback filter is preferred.

Low-pass filters used to define the top limit of the audio bandwidth are typically second order with roll-off rates of 12 dB/octave; third-order 18 dB/octave filters are rather rarer, probably because there seems to be a general feeling that phase changes are more audible at the top end of the audio spectrum than the bottom. Either the Butterworth (maximally flat frequency response) or Bessel type (maximally flat group delay) can be used. It is unlikely that there is any real audible difference between the two types of filter in this application, as most of the action occurs above 20 kHz, but using the Bessel alignment does require compromises in the effectiveness of the filtering because of its slow roll-off. I will demonstrate.

The standard Butterworth filter in Figure 5.2(a) has its −3 dB point set to 50 kHz, and this gives a loss of only 0.08 dB at 20 kHz, so there is minimal intrusion into the audio band

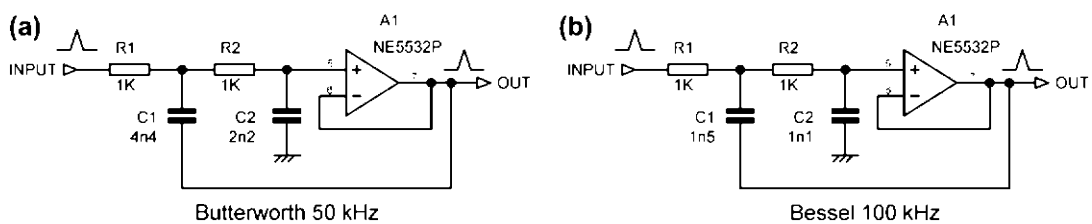


Figure 5.2: Second-order Sallen-and-Key low-pass circuits for ultrasonic filtering: (a) Butterworth; (b) Bessel. Both have a loss of less than 0.2 dB at 20 kHz

www.raguvaran.puzl.com

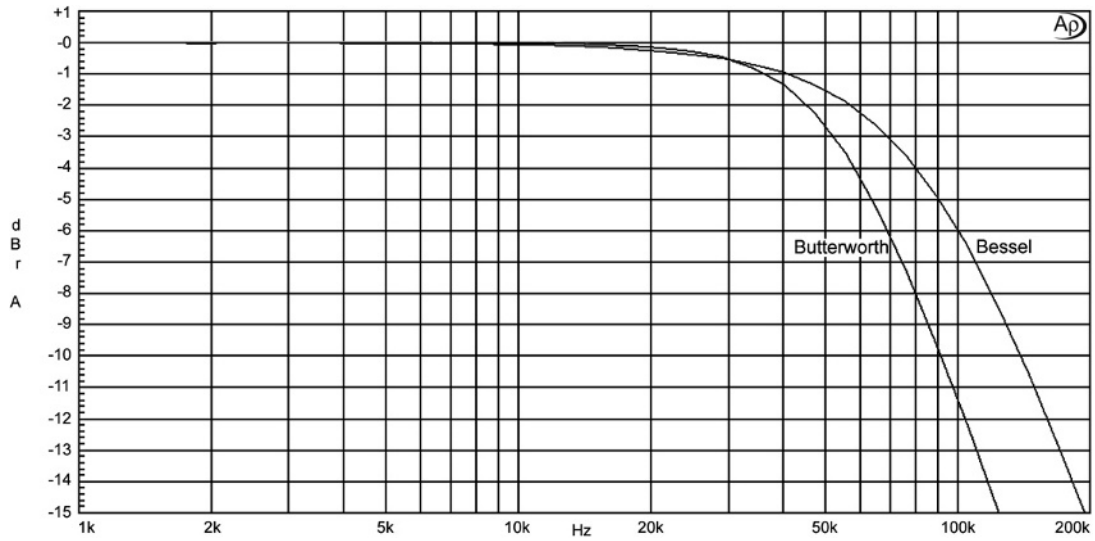


Figure 5.3: Frequency response of a 50 kHz Butterworth and a 72 kHz Bessel filter, as in Figure 5.2

(see Figure 5.3). The response is a useful −11.6 dB at 100 kHz and an authoritative −24.9 dB at 200 kHz. C1 is made up of two 2n2 capacitors in parallel.

But let us suppose we are concerned about linear phase at high frequencies and we decide to use a Bessel filter. The only circuit change is that C1 is now 1.335 times as big as C2 instead of 2 times, but the response is very different. If we design for −3 dB at 50 kHz again, we find that the response is −0.47 dB at 20 kHz – a lot worse than −0.08 dB, and not exactly a stunning figure for your spec sheet. If we decide we can live with −0.2 dB at 20 kHz then the Bessel filter has to be designed for −3 dB at 72 kHz; this is the design shown in Figure 5.2(b). Due to the inherently slower roll-off, the response is only down to −5.6 dB at 100 kHz, and −14.9 dB at 200 kHz, as seen in Figure 5.3; the latter figure is 10 dB worse than for the Butterworth. The measured noise output for both versions is −114.7 dBu (corrected).

If we want to keep the 20 kHz loss to 0.1 dB, the Bessel filter has to be designed for −3 dB at 100 kHz, and the response is now only −10.4 dB down at 200 kHz, more than 14 dB less effective than the Butterworth's. These results are summarized in Table 5.1.

Discussions on filters always remark that the Bessel alignment has a slower roll-off, but often fail to emphasize that it is a *much* slower roll-off. You should think hard before you decide to go for the Bessel option in this sort of application.

It is always worth checking how the input impedance of a filter loads the previous stage. In this case, the input impedance is high in the passband, but above the roll-off point it falls until it reaches the value of R1, which here is 1 kΩ. This is because at high frequencies C1 is not

www.raguvaran.puzl.com

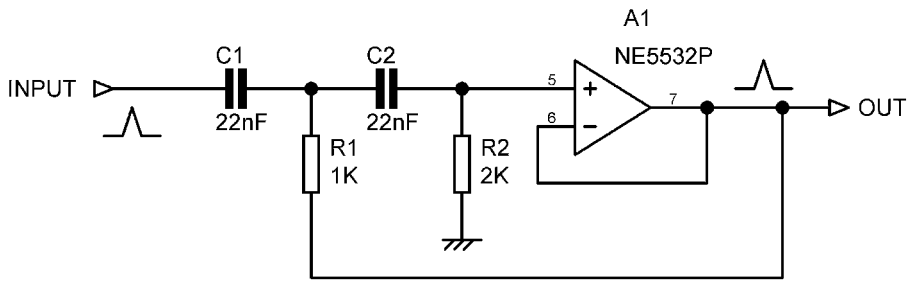
TABLE 5.1 The frequency response (dB) of various ultrasonic filter options

Frequency (kHz)	Butterworth 50 kHz	Bessel 50 kHz	Bessel 72 kHz	Bessel 100 kHz
20	−0.08	−0.47	−0.2	−0.1
100	−11.6	−10.0	−5.6	−3.0
200	−24.9	−20.9	−14.9	−10.4

bootstrapped, and the input goes through R1 and C1 to the low-impedance op-amp output, which is effectively at ground. Fortunately this low impedance only occurs at high frequencies, where one hopes the level of the signals to be filtered out will be low.

Another important consideration with low-pass filters is the balance between the R and C values in terms of noise performance. R1 and R2 are in series with the input and their Johnson noise will be added directly to the signal. Here the two 1 kΩ resistors generate −119.2 dBu of noise (22 kHz bandwidth, 25°C). The obvious conclusion is that R1 and R2 should be made as low in value as possible without causing excess loading (and 1 kΩ is not a bad compromise), with C1, C2 scaled to maintain the desired roll-off frequency. The need for specific capacitor ratios creates problems as capacitors are available in a much more limited range of values than resistors, usually the E6 series, running 10, 15, 22, 33, 47, 68. C1 or C2 often has to be made up of two capacitors in parallel.

Figure 5.4 shows a second-order high-pass Sallen-and-Key filter with a −3 dB frequency of 5115 Hz and a Q of 0.707, with its design equations; this was another part of the electronic crossover. The capacitors are equal while the resistors must have a ratio of 2. The measured noise output of this filter is −115.2 dBu (corrected).



$$f_0 = \frac{1}{2\pi C\sqrt{R_1 R_2}} \qquad Q = \frac{1}{2}\sqrt{\frac{R_2}{R_1}}$$

Figure 5.4: The classic second-order high-pass Sallen-and-Key filter. Cut-off frequency is 5115 Hz, Q = 0.707 (critically damped)

A variation on the low-pass Sallen-and-Key filter that can avoid capacitor ratio difficulties is shown in Figure 5.5. The unity-gain buffer is replaced with a voltage-gain stage; the gain set by

www.raguvaran.puzl.com

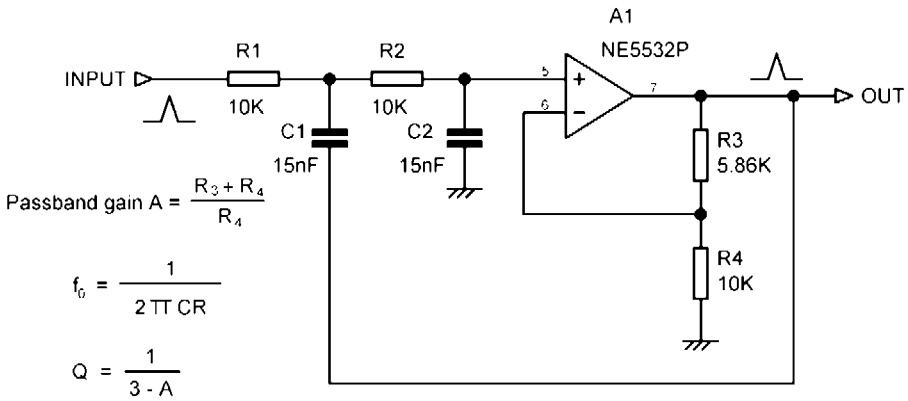


Figure 5.5: Equal-value second-order low-pass Butterworth filter with a cut-off frequency of 1061 Hz. Gain must be 1.586 times for maximally flat response ($Q = 0.707$)

R3 and R4 must be 1.586 times (+4.00 dB) for a Q of 0.707. This allows C1 and C2 to be the same value. An equal-resistor-value high-pass filter can be made in exactly the same way.

To get a faster roll-off, we can use a third-order filter, which is a first-order filter (i.e. a simple RC time constant) cascaded with a second-order filter that has a Q of 1.0, causing a response peak. This peak, combined with the slow first-order roll-off, gives a flat passband and then a steep roll-off. Figure 5.6(a) shows a third-order Butterworth filter with a −3 dB frequency of 100 Hz, as might be used at the front end of a mixer channel. It is built the obvious way, with

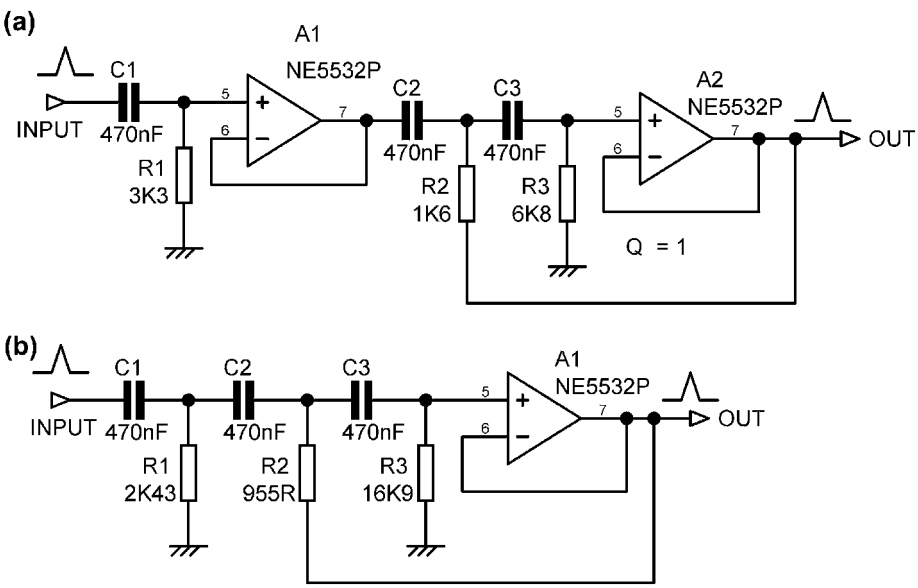


Figure 5.6: Two third-order Butterworth high-pass filters with a cut-off frequency of 100 Hz

www.raguvaran.puzl.com

a first-order filter R1, C1 followed by a unity-gain buffer to give low-impedance drive to the following second-order filter. R2 and R3 now have a ratio of 4 to obtain a Q of 1.0. E24 resistor values are shown and in this case they give an accurate response.

A more economical way to make a third-order filter is shown in Figure 5.6(b), which saves an op-amp section. The resistor values shown are the nearest E96 values to the mathematically exact numbers, and give an extremely accurate response. The nearest E24 values are R1 = 2k4, R2 = 910 Ω, and R3 = 16 kΩ, giving a very small peaking of 0.06 dB at 233 Hz. Errors due to the capacitor tolerances are likely to be larger than this.

Fourth-order filters are the steepest in normal use. They are made by cascading two second-order filters with Q values of 0.54 and 1.31. They can also be made in the same way as Figure 5.6(b), but if so tend to be rather sensitive to component tolerances.

Distortion in Sallen-and-Key Filters

When they have a signal voltage across them, many capacitor types generate distortion. This unwelcome phenomenon is described in Chapter 2. It afflicts not only all electrolytic capacitors, but also some types of non-electrolytic. If the electrolytics are being used as coupling capacitors, then the cure is simply to make them so large that they have a negligible signal voltage across them at the lowest frequency of interest; less than 80 mVrms is a reasonable criterion. This means they may have to be 10 times the value required for a satisfactory frequency response.

However, when non-electrolytics are used to set time-constants in filters they obviously must have substantial signal voltages across them and this fix is not usable. The problem is not a marginal one – the amounts of distortion produced can be surprisingly high. Figure 5.7 shows the frequency response of a conventional second-order Sallen-and-Key high-pass filter as seen in Figure 5.4, with a −3 dB frequency of 520 Hz. C1, C2 were 220 nF, 100 V polyester capacitors, with R1 = 1 kΩ and R2 = 2 kΩ. The op-amp was a 5532. The distortion performance is shown by the upper trace in Figure 5.8; above 1 kHz the distortion comes from the op-amp alone and is very low. However, you can see it rising rapidly below 1 kHz as the filter begins to act, and it has reached 0.015% by 100 Hz, completely overshadowing the op-amp distortion; it is basically third order. The input level was 10 Vrms, which is about as much as you are likely to encounter in an op-amp system. The output from the filter has dropped to −28 dB by 100 Hz, and so the amplitude of the harmonics generated is correspondingly lower, but it is still not a very happy outcome.

As explained in Chapter 2, polypropylene capacitors exhibit negligible distortion compared with polyester, and the lower trace in Figure 5.8 shows the improvement on substituting 220 nF, 250 V polypropylene capacitors. The THD residual below 500 Hz is now pure noise,

www.raguvaran.puzl.com

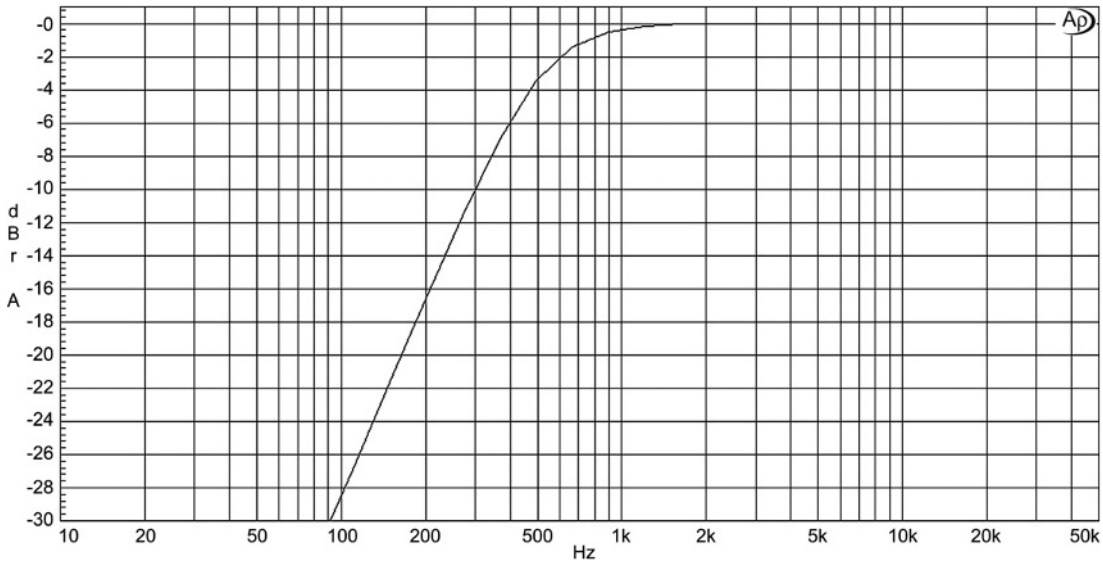


Figure 5.7: The frequency response of the second-order 520 Hz high-pass filter

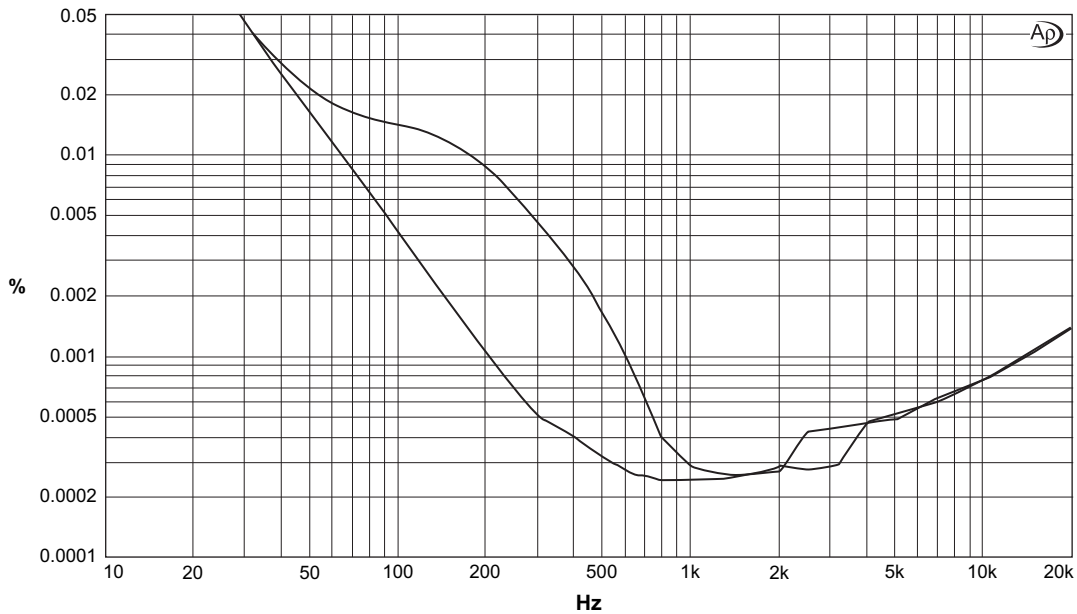


Figure 5.8: THD plot from the second-order 520 Hz high-pass filter; input level 10 Vrms. The upper trace shows distortion from polyester capacitors; the lower trace, with polypropylene capacitors, shows noise only

www.raguvaran.puzl.com

and the trace is only rising at 12 dB/octave because circuit noise is constant but the filter output is falling. The important factor is the dielectric, not the voltage rating; 63 V polypropylene capacitors are also free from distortion. The only downside is that polypropylene capacitors are larger for a given *CV* product and more expensive.

Multiple-Feedback Bandpass Filters

When a bandpass filter of modest Q is required, the multiple-feedback or Rauch type shown in Figure 5.9 has many advantages. The capacitors are equal and so can be made any preferred value. The op-amp is working with shunt feedback and so has no common-mode voltage on the inputs, which avoids one source of distortion. It does, however, phase-invert, which can be inconvenient.

The filter response is defined by three parameters – the center frequency f_0, the Q, and the passband gain (i.e. the gain at the response peak) A. The filter in Figure 5.9 was designed for $f_0 = 250$ Hz, $Q = 2$, and $A = 1$ using the equations given, and the usual awkward resistor values emerged. The resistors are the nearest E96 value, and the simulated results come out as $f_0 = 251$ Hz, $Q = 1.99$, and $A = 1.0024$, which, as they say, is good enough for rock'n'roll.

The Q of the filter can be quickly checked from the response curve as Q is equal to the center frequency divided by the -3 dB bandwidth, i.e. the frequency difference between the two -3 dB points on either side of the peak. This configuration is not suitable for Q values greater than about 10, as the filter characteristics become unduly sensitive to component tolerances. If independent control of f_0 and Q is required the state-variable filter should be used instead.

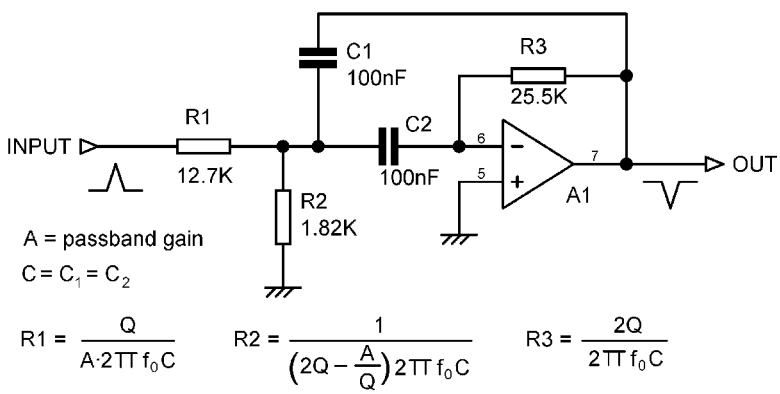


Figure 5.9: A bandpass multiple-feedback filter with $f_0 = 250$ Hz, $Q = 2$, and a gain of 1

www.raguvaran.puzl.com

Similar configurations can be used for low-pass and high-pass filters. The low-pass version does not depend on a low op-amp output impedance to maintain stop-band attenuation at high frequencies, and avoids the oh-no-it's-coming-back-up-again behavior of Sallen-and-Key low-pass filters.

Making Notch Filters

There are many ways to dig a deep notch in your frequency response. The width of a notch is described by its Q; exactly as for a resonance peak, Q is equal to the center frequency divided by the −3 dB bandwidth, i.e. the frequency difference between the −3 dB points either side of the notch. Q has no relation to the depth of the notch.

The best-known notch filter is the twin-T notch network shown in Figure 5.10(a), invented in 1934 by Herbert Augustadt [8]. The notch depth is infinite with exactly matched components, but with ordinary ones it is unlikely to be deeper than 40 dB. It requires ratios of 2 in component values, which with the exception of 1k/2k does not fit in well with preferred values, and when used alone has a Q of only 0.25. It is therefore normally used with positive feedback via an op-amp buffer A2, as shown. The proportion of feedback K and hence the Q enhancement is set by R4 and R5, which here give a Q of 1. A great drawback is that the notch frequency can only be altered by changing three components, so it is not considered tuneable.

Another way of making notch filters is the '1-bandpass' principle mentioned earlier. The input goes through a bandpass filter, typically the multiple-feedback type described earlier, and is then subtracted from the original signal. The accuracy of the cancellation and hence the notch depth is critically dependent on the mid-band gain of the bandpass filter. Figure 5.10(b) shows an example that gives a notch at 50 Hz with a Q of 2.85. The subtraction is performed by A2, as the output of the filter is phase-inverted. The multiple-feedback filter is designed for unity passband gain, but the use of E24 values as shown means that the actual gain is 0.97, limiting the notch depth to −32 dB. The value of R6 can be tweaked to deepen the notch; the nearest E96 value is 10.2 kΩ, which gives a depth of −45 dB. The final output is inconveniently phase-inverted in the passband.

A notch filter that deserves to be better known is the Bainter filter [9,10] shown in Figure 5.10(c). It is non-inverting in the passband, and two out of three op-amps are working at virtual earth and will give no trouble with common-mode distortion. An important advantage is that the notch depth does not depend on the matching of components, but only on the open-loop gain of the op-amps, being roughly proportional to it. With TL072-type op-amps the depth is from −40 to −50 dB. The values shown give a notch at 700 Hz with a Q of 1.29. The design equations can be found in Ref. [10].

www.raguvaran.puzl.com

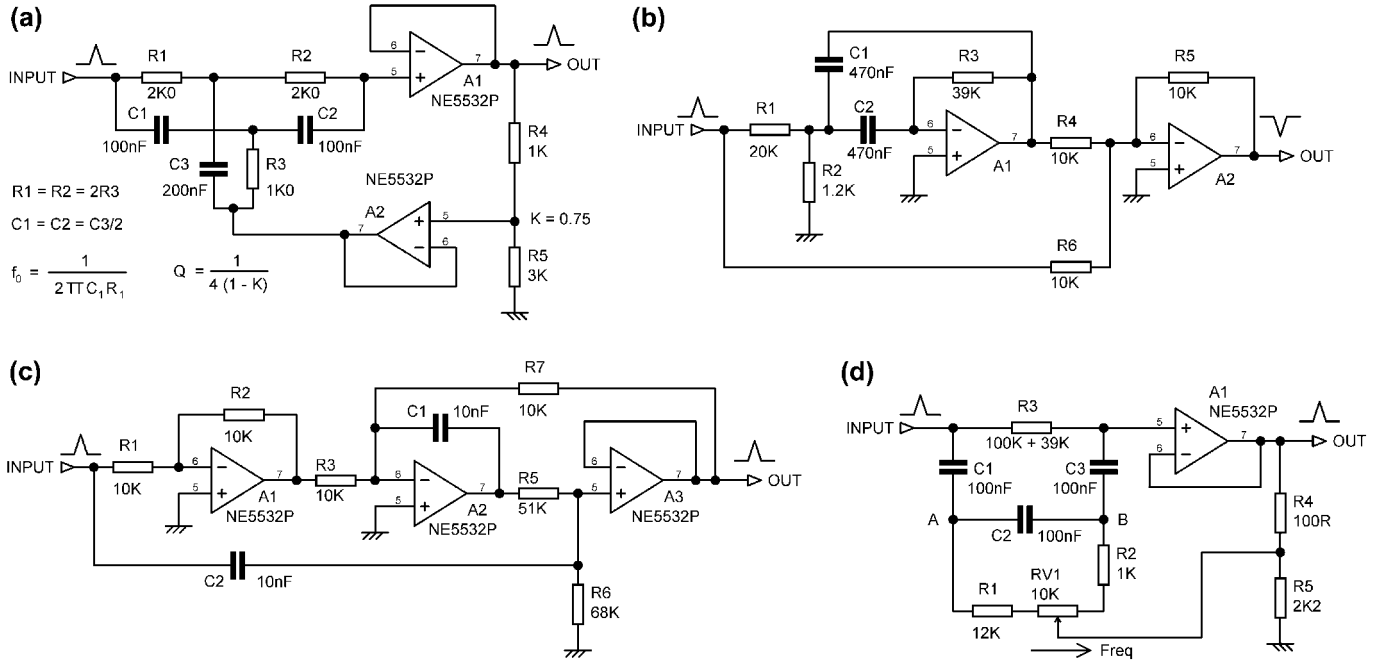


Figure 5.10: Notch filters. (a) Twin-T with positive feedback, notch at 795 Hz and $Q = 1$. (b) '1-bandpass' filter with notch at 50 Hz, $Q = 2.85$. (c) Bainter filter with notch at 50 Hz, $Q = 2.85$. (d) Bridged-differentiator notch filter tuneable at 80–180 Hz

www.raguvaran.puzl.com

A property of this filter that does not seem to appear in the textbooks is that if R1 and R4 are altered together, i.e. having the same values, then the notch frequency is tuneable with a good depth maintained, but the Q does change proportionally to frequency. To get a standard notch with equal gain either side of the crevasse R3 must equal R4. R4 greater than R3 gives a low-pass notch, while R3 greater than R4 gives a high-pass notch; these responses are useful in crossover design [5].

The Bainter filter is usually shown with equal values for C1 and C2. This leads to values for R5 and R6 that are a good deal higher than other circuit resistances and this will impair the noise performance. I suggest that C2 is made 10 times C1, i.e. 100 nF, and R5 and R6 are reduced by 10 times to 5.1 and 6.8 kΩ; the response is unaltered and the Johnson noise much reduced.

But what about a notch filter that can be tuned with one control? Figure 5.10(d) shows a bridged-differentiator notch filter tuneable from 80 to 180 Hz by RV1. R3 must theoretically be six times the total resistance between A and B, which here is 138 kΩ, but 139 kΩ gives a deeper notch, about −27 dB across the tuning range. The downside is that Q varies with frequency from 3.9 at 80 Hz to 1.4 at 180 Hz.

Another interesting notch filter to look up is the Boctor, which uses only one op-amp [11,12]. Both the biquad and state-variable filters can be configured to give notch outputs.

When simulating notch filters, assessing the notch depth can be tricky. You need a lot of frequency steps to ensure you really have hit bottom with one of them. For example, in one run, 50 steps/decade showed a −20 dB notch, but upping it to 500 steps/decade revealed it was really −31 dB deep. In most cases having a stupendously deep notch is pointless. If you are trying to remove an unwanted signal then it only has to alter in frequency by a tiny amount and you are on the side of the notch rather than the bottom and the attenuation is much reduced. The exception to this is the THD analyzer, where a very deep notch (120 dB or more) is needed to reject the fundamental so very low levels of harmonics can be measured. This is achieved by continuously servo-tuning the notch so it is kept exactly on the incoming frequency.

Differential Filters

At first an active filter that is also a differential amplifier, and thereby carries out an accurate subtraction, sounds like a very exotic creature. Actually they are quite common, and are normally based on the multiple-feedback filter described above. Their main application is DAC output filtering in CD payers and the like; more sophisticated versions are useful for data acquisition in difficult environments. Differential filters are dealt with in the section of Chapter 21 on interfacing with DAC outputs.

www.raguvaran.puzl.com

References

[1] Williams, Taylor, Electronic Filter Design Handbook, fourth ed, McGraw-Hill, 2006.
[2] Van Valkenburg, Analog Filter Design, Holt-Saunders International Editions, 1982.
[3] H.M. Berlin, Design of Active Filters with Experiments, Blacksburg, 1978, p. 85.
[4] B. Hofer, Switch-mode amplifier performance measurements, Electronics World (September 2005) p. 30.
[5] W. Hardman, Precise active crossover, Electronics World (August 2009) p. 652.
[6] R.P. Sallen, E.L. Key, A practical method of designing RC active filters, IRE Transactions on Circuit Theory 2 (1) (1955) pp. 74–85.
[7] S. Linkwitz, Active crossover networks for non-coincident drivers, J. Audio Eng. Soc (January/February 1976).
[8] H. Augustadt, Electric filter, US patent 2,106,785, February 1938, assigned to Bell Telephone Labs.
[9] J.R. Bainter, Active filter has stable notch, and response can be regulated, Electronics (October 1975) pp. 115–117.
[10] Van Valkenburg, Analog Filter Design, Holt-Saunders International Editions, 1982, p. 346.
[11] S.A. Boctor, Single amplifier functionally tuneable low-pass notch filter, IEEE Trans. Circuits Systems CAS-22 (1975) pp. 875–881.
[12] Van Valkenburg, Analog Filter Design, Holt-Saunders International Editions, 1982, p. 349.

www.raguvaran.puzl.com

This page intentionally left blank

www.raguvaran.puzl.com

CHAPTER 6

Preamplifier Architectures

Some sort of preamplifier or control unit is required in all hi-fi systems, even if its only function is to select the source and set the volume. You could even argue that the source-selection switch could be done away with, if you are prepared to plug and unplug those rhodium-plated connectors, leaving a 'preamplifier' that basically consists solely of a volume-control potentiometer in a box. Mind you don't catch your sackcloth on those heat-sink corners, and try to keep the ashes away from the turntable.

I am assuming here that a selector switch will be required, and that gives us the 'passive preamplifier' (oxymoron alert!) in Figure 6.1(a).

Passive Preamplifiers

A device may have only one component but it does not follow that it is easy to design, even though the only parameter to decide is the resistance of the volume pot. Any piece of equipment that embodies its internal contradictions in its very name needs to be treated with caution. The pot resistance of a 'passive preamplifier' cannot be too high because the output impedance, maximal at one-quarter the track resistance when volume is set to −6 dB, will cause an HF roll-off in conjunction with the connecting cable capacitance. It also makes life difficult for those designing RF filters on the inputs of the equipment being driven, as described in Chapter 14.

On the other hand, if the volume pot resistance is too low the source equipment will suffer excessive loading. If the source is valve equipment, which does not respond well to even moderate loading, the problem starts to look insoluble.

If, however, we can assume that our source equipment has a reasonable drive capability, we can use a 10 kΩ pot. Its maximum output impedance (at −6 dB) will then be 2.5 kΩ. The capacitance of most audio cable is 50–150 pF/meter, so with a 2.5 kΩ source impedance and 100 pF/meter cable, a maximum length of 5 meters is permissible before the HF loss hits the magic figure of −0.1 dB at 20 kHz. A very rapid survey of current (2009) 'passive preamplifiers' confirmed that 10 kΩ seems to be the most popular value. One model had a 20 kΩ potentiometer, and another had a 100 kΩ pot, which with its 25 kΩ source impedance would hardly allow any cable at all. I suspect that the only reason such a pot can appear

Small Signal Audio Design; ISBN: 9780240521770
Copyright © 2010 Elsevier Ltd; All rights of reproduction, in any form, reserved.

www.raguvaran.puzl.com

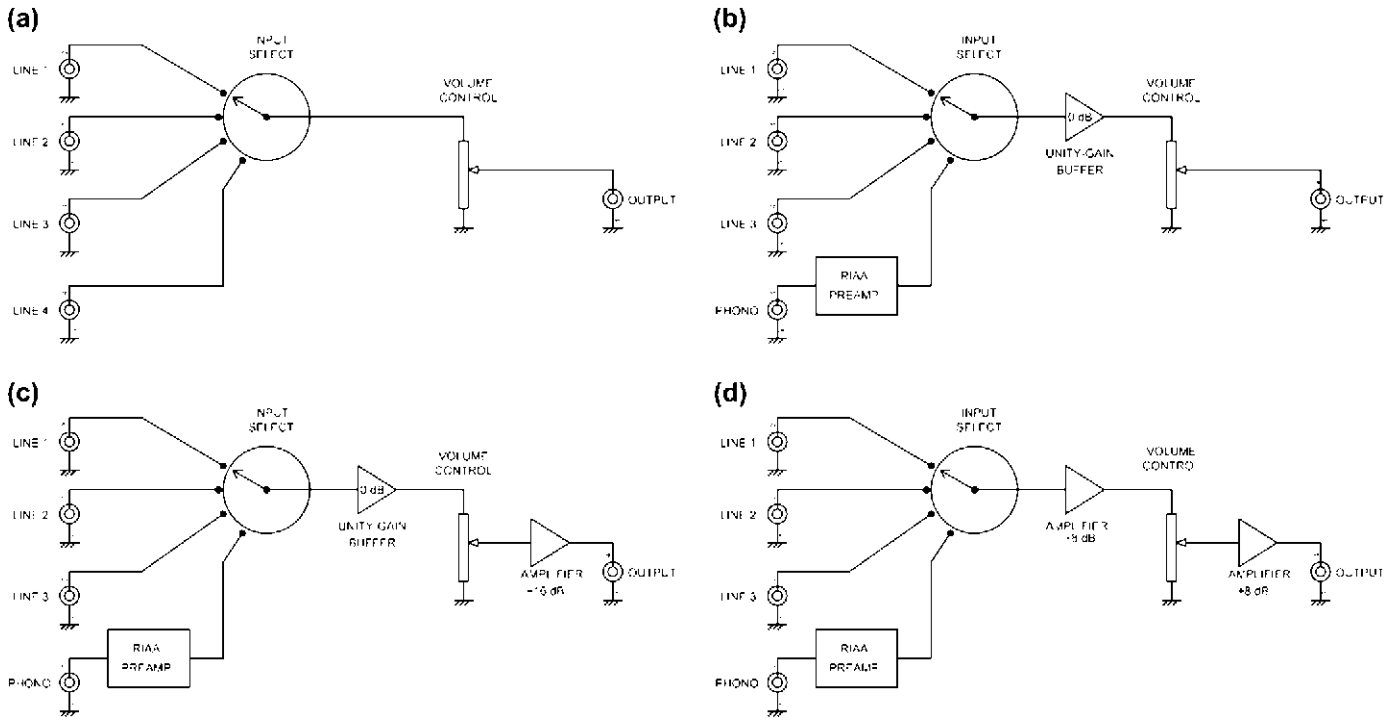


Figure 6.1: Preamplifier evolution. (a) Passive preamplifier. (b) Input buffer and phono amplifier added. (c) Amplification after volume control added. (d) Amplification split into two stages, before and after volume control

www.raguvaran.puzl.com

acceptable is because in normal use it is set well below −6 dB. For example, if a 100 kΩ pot is set to −15 dB, the output impedance is reduced to 14 kΩ, which would allow just under a meter of 100 pF/meter cable to be used with HF loss still limited to −0.1 dB at 20 kHz. A high output impedance also makes an interconnection more susceptible to interference unless it is totally screened at every point on its length.

Consider also that many power amplifiers have RC filters at the input, not so much for EMC immunity, but more as a gesture against what used to be called 'transient intermodulation distortion' (TID); but this is actually just old-fashioned slew limiting and highly unlikely in practice. These can add extra shunt capacitance to the input ranging from 100 to 1000 pF, apparently having been designed on assumption of near-zero source impedance, and this can cause serious HF roll-offs.

There is at least one passive preamplifier on the market that controls volume by changing the taps on the secondary of a transformer; this should give much lower output impedances. There is more on that approach in Chapter 9 on volume controls.

Active Preamplifiers

Once we permit ourselves active electronics, things get much easier. If a unity-gain buffer stage is added after the selector switch, as in Figure 6.1(b), the volume pot resistance can be reduced to much less than 10 kΩ, while presenting a high impedance to the sources. If a 5532 is used there is no technical reason why the pot could not be as low as 1 kΩ, which will give a much more usable maximum output impedance of only 250 Ω, and also reduce Johnson noise by 10 dB. A phono preamplifier has also been added. Now we've paid for a power supply, it might as well supply something else.

This still leaves us with an 'amplifier' that has only unity line gain. Normally only CD players, which have an output of 2 Vrms, can fully drive a power amplifier without additional gain, and there are some high-power amplifiers that require more than this for full output. iPods appear to have a maximum output of 1.2 Vrms. Output levels for tuners, phono amps, and so on vary but may be as low as 150 mVrms, while power amplifiers rarely have sensitivities lower than 500 mV. Clearly some gain would be a good thing, so one option is adding a gain stage after the volume control, as in Figure 6.1(c). The output level can be increased and the output impedance kept down to 100 Ω or lower.

This amplifier stage introduces its own difficulties. If its nominal output level with the volume control fully up is taken as 1 Vrms for 150 mV in, which will let us drive most power amps to full output from most sources most of the time, we will need a gain of 6.7 times or 16.5 dB. If we decide to increase the nominal output level to 2 Vrms, to be sure of driving most if not all exotica to its limits, we need 22.5 dB. The problem is that the gain stage is amplifying its own noise at all volume settings, and amplifying a proportion of the Johnson noise of the pot

www.raguvaran.puzl.com

whenever the wiper is off the zero stop. The noise performance will therefore deteriorate markedly at low volume levels, which are the ones most used.

Amplification and the gain-distribution problem

One answer to this difficulty is to take the total gain and split it so there is some before and some after the volume control, so there is less gain amplifying the noise at low volume settings. One version of this is shown in Figure 6.1(d). The question is: how much gain before and how much after? This is inevitably a compromise, and it might be called the gain-distribution problem. Putting more of the total gain before the volume control reduces the headroom as there is no way to reduce the signal level, while putting more after increases the noise output at low volume settings.

If you are exclusively using sources with a predictable output, of which the 2 Vrms from a CD player will be the maximum, the overload situation is well defined, and if we assume that the pre-volume gain stage is capable of at least 8 Vrms out, so long as the pre-volume control gain is less than 4 times there will never be a clipping problem. However, phono cartridges, particularly moving-coil ones, which have a very wide range of sensitivities, produce much less predictable outputs after fixed-gain preamplification, and it is a judgement call as to how much safety margin is desirable.

As an aside, it's worth bearing in mind that even putting a unity-gain buffer before the volume control, which we did as the first step in preamp evolution, does place a constraint on the signal levels that can be handled, albeit at rather a high level of 8–10 Vrms depending on the supply rails in use. The only source likely to be capable of putting out such levels is a mixing console with the group faders fully advanced. There is also the ultimate constraint that a volume control pot can only handle so much power, and the manufacturers' ratings are surprisingly low, sometimes only 50 mW. This means that a 10 kΩ pot would be limited to 22 Vrms across it, and if you are planning to use lower resistance pots than this to reduce noise, their power rating needs to be kept very much in mind.

Whenever a compromise appears in engineering, you can bet that someone will try to find a way round it and get the best of both worlds. What can be done about the gain-distribution dilemma?

One possibility is the use of a special low-noise amplifier after the volume control, combined with a low-resistance volume pot as suggested above. This could be done either by a discrete device and op-amp hybrid stage, or by using a multiple op-amp array, as described in Chapter 1. It is doubtful if it is possible to obtain more than a 10 dB noise improvement by these means, but it would be an interesting project.

Another possible solution is the use of double gain controls. There is an input gain control before any amplification stage, which is used to set the internal level appropriately, thus

www.raguvaran.puzl.com

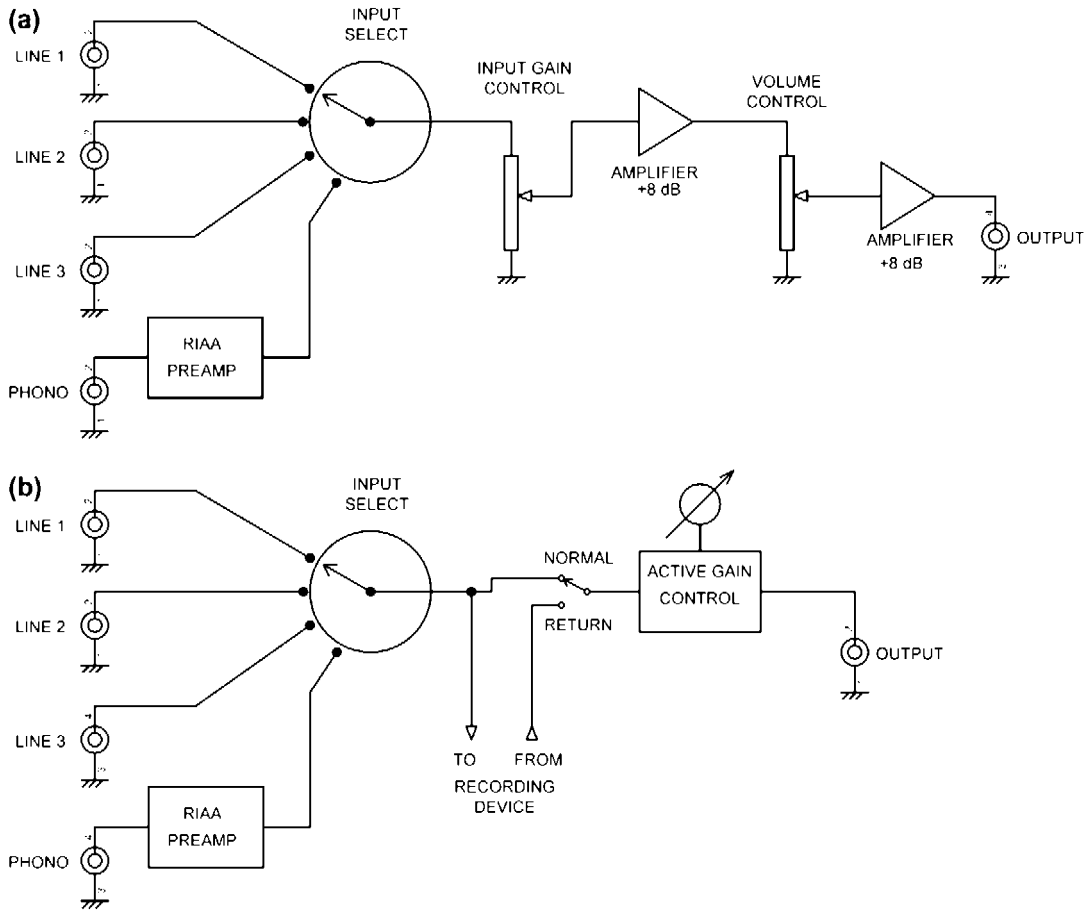


Figure 6.2: More preamp architectures. (a) With input gain control and output volume control. (b) With recording output and return input, and an active gain control

avoiding overload, and after the active stages there is an output volume control, which gives the much-desired silence at zero volume (see Figure 6.2(a)). The input gain controls can be separate for each channel, so they double as a balance facility; this approach was used on the Radford HD250 amplifier, and also in one of my early preamplifier designs [1]. This helps to offset the cost of the extra pot. However, having two gain controls is operationally rather awkward, and however attenuation and fixed amplification are arranged, there are always going to be some trade-offs between noise and headroom. It could also be argued that this scheme does not make a lot of sense unless some means of metering the signal level after the input gain control is provided.

If the input and output gain controls are ganged together, to improve ease of operation at the expense of flexibility, this is sometimes called a distributed gain control.

www.raguvaran.puzl.com

Active Gain Controls

The noise/headroom compromise is completely avoided by replacing the combination of volume control and amplifier with an active gain control, i.e. an amplifier stage whose gain is variable from near-zero to the required maximum (see Figure 6.2(b)). We get lower noise at gain settings below maximum, and we can increase that maximum gain so even the least sensitive power amplifiers can be fully driven, without impairing the noise performance at lower settings. We also get the ability to generate a quasi-logarithmic law from a linear pot, which gives excellent channel balance as it depends only on mechanical alignment. The only snag is that most active gain controls phase-invert. The technology is dealt with fully in Chapter 9.

Recording Facilities

If a preamplifier is going to be used for recording, the minimum requirement is an output taken from before the volume control, and there is usually also a dedicated input for a signal coming back from the recorder that can be switched to for checking purposes. Back when recording was done on tape, the return signal could be taken from the replay heads and gave assurance that recording had actually happened. Now that recording is done on hard-disk machines or PCs the return signal normally only assures you that the signal has actually got there and back.

Much ingenuity used to be expended in designing switching systems so you could listen to one source while recording another, though it is rather doubtful how many people actually wanted to do this; it demands very high standards of crosstalk inside the preamplifier to make sure that the signal being recorded is not contaminated by another source.

Tone Controls

Let us now consider adding tone controls. While it is currently unfashionable to have tone controls on a preamplifier, this seems to be starting to change. I think they are absolutely necessary, and it is a startling situation when, as frequently happens, anxious inquirers to hi-fi advice columns are advised to change their loudspeakers to correct excess or lack of bass or treble. This is an extremely expensive alternative to tone controls.

There are many possible types, as described in Chapter 10, but one thing they have in common is that they must be fed from a low-impedance source to give the correct boost/cut figures and predictable EQ curves. Another vital point is that most types, including the famous Baxandall configuration, phase-invert. Since there is now pretty much a consensus that all audio equipment should maintain absolute phase polarity for all input and outputs, this can be highly inconvenient.

www.raguvaran.puzl.com

However, this phase inversion can very conveniently be undone by the use of an active gain control, which also uses shunt feedback and so also phase-inverts. The tone control can be placed before or after the active gain control, but if placed afterwards it generates noise that cannot be turned down. Putting it before the active gain control reduces headroom if boost is in use, but if we assume the maximum boost used is +10 dB, the preamp inputs will not overload before 3 Vrms is applied, and domestic equipment can rarely generate such levels. It therefore seems best to put the tone control before the active gain control. This is what I did in my most recent preamplifier designs [2,3].

References

[1] D. Self, An advanced preamplifier, Wireless World (November 1976).
[2] D. Self, A precision preamplifier, Wireless World (October 1983).
[3] D. Self, Precision Preamplifier 96, Electronics World (July/August and September 1996).

www.raguvaran.puzl.com

This page intentionally left blank

www.raguvaran.puzl.com

CHAPTER 7

Moving-Magnet Disc Inputs

The Vinyl Medium

The vinyl disc as a medium for music delivery in its present form dates back to 1948, when Columbia introduced microgroove 33⅓ rpm LP records. These were followed soon after by microgroove 45 rpm records from RCA Victor. Stereo vinyl did not appear until 1958. The introduction of Varigroove technology, which adjusts groove spacing to suit the amplitude of the groove vibrations, using an extra look-ahead tape head to see what the future holds, allowed increases in groove packing density. This density rarely exceeded 100 grooves per inch in the 78 rpm format, but with Varigroove 180–360 grooves/inch could be used at 33⅓ rpm.

While microgroove technology was unquestionably a considerable improvement on 78 rpm records, any technology that is 50 years old is likely to show definite limitations compared with contemporary standards, and indeed it does. Compared with modern digital formats, vinyl has a restricted dynamic range, poor linearity (especially at the end of a side) and is very vulnerable to permanent and irritating damage in the form of scratches. Even with the greatest care, scratches are likely to be inflicted when the record is removed from its sleeve. This action also generates significant static charges, which attract dust and lint to the record surface. If not carefully removed this dirt builds up on the stylus and not only degrades the reproduction of high-frequency information today, but may also damage it in the future if it provokes mistracking.

Vinyl discs do not shatter under impact like the 78 shellac discs, but they are subject to warping by heat, improper storage, or poor manufacturing quality control. Possibly the worst feature of vinyl is that the stored material is degraded every time the disc is played, as the delicate high-frequency groove modulations are worn away by the stylus. When a good turntable with a properly balanced tone-arm and correctly set up low-mass stylus is used this wear process is relatively slow, but it nevertheless proceeds inexorably.

However, for reasons that have very little to do with logic or common sense, vinyl is still very much alive. Even if it is accepted that as a music-delivery medium it is technically as obsolete as wax cylinders, there remain many sizeable album collections that it is impractical to replace

Small Signal Audio Design; ISBN: 9780240521770
Copyright © 2010 Elsevier Ltd; All rights of reproduction, in any form, reserved.

www.raguvaran.puzl.com

with CDs and would take an interminable time to transfer to the digital domain. I have one of them. Disc inputs must therefore remain part of the audio designer's repertoire for the foreseeable future, and the design of the specialized electronics to get the best from the vinyl medium is still very relevant.

Spurious Signals

It is not easy to find dependable statistics on the dynamic range of vinyl, but there seems to be general agreement that it is in the range 50–80 dB, the 50 dB coming from the standard quality discs and the 80 dB representing direct-cut discs produced with quality as the prime aim. My own view is that 80 dB is rather optimistic.

The most audible spurious noise coming from vinyl is that in the mid frequencies, stemming from the inescapable fact that the music is read by a stylus sliding along a groove of finite smoothness. There is nothing that the designer of audio electronics can do about this. Scratches create clicks that have a large high-frequency content, and it has been shown that they can easily exceed the level of the audio [1]. It is important that such clicks do not cause slew limiting or other forms of overload, as this makes their subjective impact worse.

The signal from a record deck also includes copious amounts of low-frequency noise, which is often called rumble; it is typically below 30 Hz. This can come from several sources:

1. Mechanical is noise generated by the motor bearings and picked up by the stylus/arm combination. This tends to be at the upper end of the low-frequency domain, extending up to 30 Hz or thereabouts. This is a matter for the mechanical designer of the turntable, as it clearly cannot be filtered out without removing the lower part of the audio spectrum.

2. Room vibrations will be picked up if the turntable and arm system is not well isolated from the floor. This is a particular problem in older houses where the wooden floors are not built to modern standards of rigidity, and have a perceptible bounce to them. Mounting the turntable shelf to the wall usually gives a major improvement. Subsonic filtering is effective in removing this.

3. Low-frequency noise from disc imperfections. This is the worst cause of disturbances. They can extend as low as 0.55 Hz, the frequency at which a $33\frac{1}{3}$ rpm disc rotates on the turntable, and are due to large-scale disc warps. Warping can also produce ripples in the surface, generating spurious subsonic signals up to a few hertz at surprisingly high levels. These can be further amplified by a poorly controlled resonance of the cartridge compliance and the pickup arm mass. When woofer speaker cones can be seen wobbling – and bass reflex designs with no cone loading at very low frequencies are the worst for this – disc warps are usually the cause. Subsonic filtering is again effective in removing this.

www.raguvaran.puzl.com

(As an aside, I have heard it convincingly argued that bass reflex designs have only achieved their current popularity because of the advent of the CD player, with its greater bass signal extension, but lack of subsonic output.)

Some fascinating data on the subsonic output from vinyl was given by Tomlinson Holman [2] and shows that the highest warp signals occur in the 2–4 Hz region, being some 8 dB less at 10 Hz. By matching these signals with a wide variety of cartridge–arm combinations, he concluded that to accommodate the very worst cases, a preamplifier should be able to accept not less than 35 mVrms in the 3–4 Hz region. This is a rather demanding requirement, driven by some truly diabolical cartridge–arm set-ups that accentuated subsonic frequencies by up to 24 dB.

Since the subsonic content generated by room vibrations and disc imperfections tends to cause vertical movements of the stylus, the resulting electrical output will be out of phase in the left and right channels. The use of a central mono subwoofer system that sums the two channels will provide partial cancellation, reducing the amount of rumble that is reproduced. It is, however, still important to ensure that subsonic signals do not reach the left and right speakers.

Maximum Signal Levels on Vinyl

There are some definite limits to the signal level possible on a vinyl disc, and they impose maxima on the signal that a cartridge and its associated electronics will be expected to reproduce. The exact values of these limits may not be precisely defined, but the way they work sets the ways in which maximum levels vary with frequency, and this is of great importance.

There are no variable gain controls on RIAA inputs, because implementing an uneven but very precisely controlled frequency response and a suitably good noise performance is quite hard enough without adding variable gain as a feature. No doubt it could be done, but it would not be easy, and the general consensus is that it is not necessary. The overload margin, or headroom, is therefore of considerable importance, and it is very much a case of the more the merrier when it comes to the numbers game of specmanship. The issue can get a bit involved, as a situation with frequency-dependent vinyl limitations and frequency-dependent gain is often further complicated by a heavy frequency-dependent load in the shape of the feedback network, which can put its own limit on amplifier output at high frequencies. Let us first look at the limits on the signal levels that stylus-in-vinyl technology can deliver. In the diagrams that follow the response curves have been simplified to the straight-line asymptotes.

Figure 7.1(a) shows the physical groove amplitudes that can be put on to a disc. From subsonic up to about 1 kHz, groove amplitude is the constraint. If the sideways excursion is too great,

www.raguvaran.puzl.com

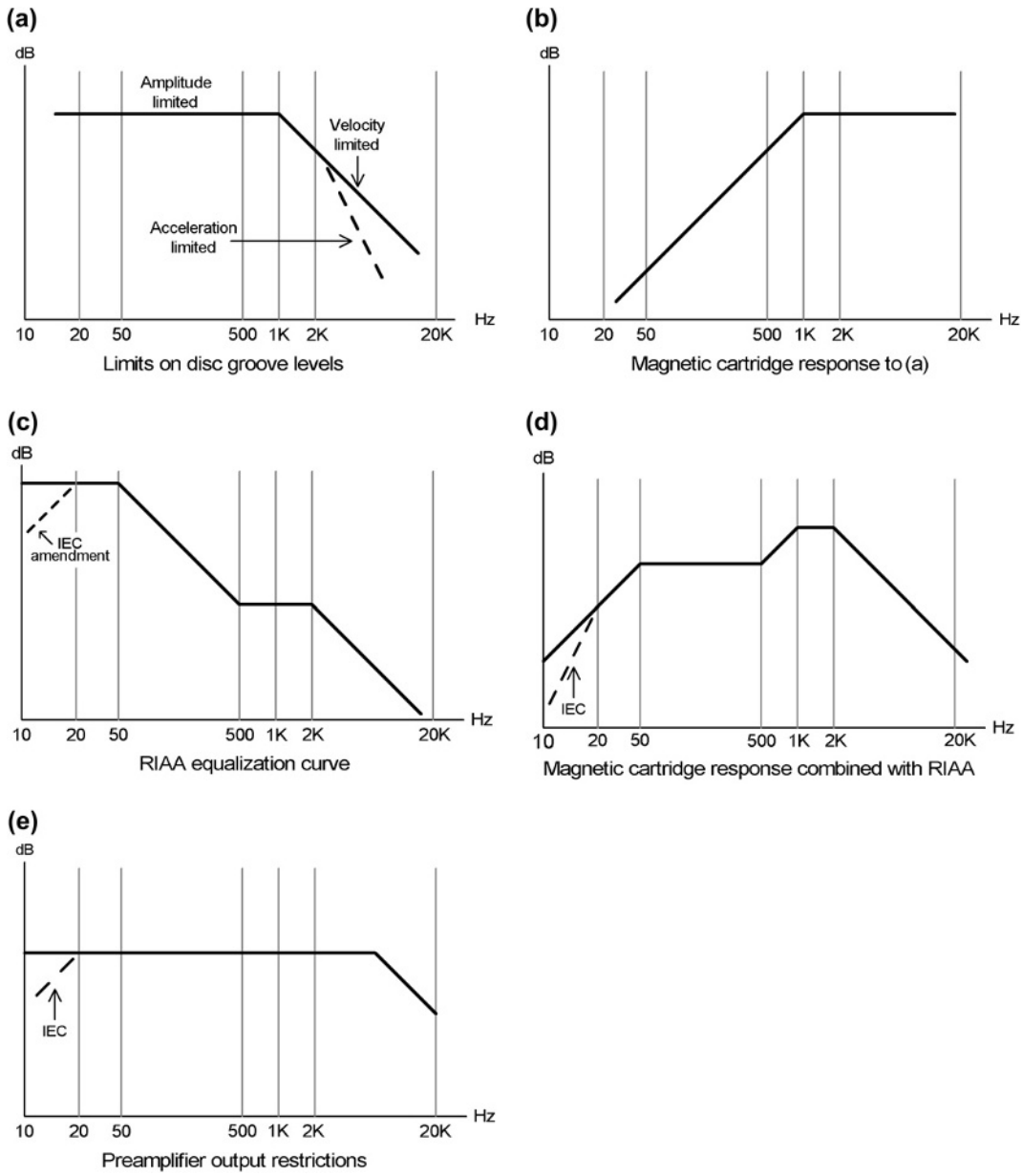


Figure 7.1: (a) The levels on a vinyl disc. (b) The cartridge response combined with the disc levels. (c) The RIAA curve. (d) The RIAA combined with curve (b). (e) Possible preamplifier output restrictions

the groove spacing will need to be increased to prevent one groove breaking into another, and playing time will be reduced. Well before actual breakthrough occurs, the cutter can distort the groove it has cut on the previous revolution, leading to 'pre-echo' in quiet sections, where

www.raguvaran.puzl.com

a faint version of the music you are about to hear is produced. Time travel may be fine in science fiction but it does not enhance the musical experience. The ultimate limit to groove amplitude is set by mechanical stops in the cutter head.

There is an extra limitation on groove amplitude; out-of-phase signals cause vertical motion of the cutter, and if this becomes excessive it can cause it to cut either too deeply into the disc medium and dig into the aluminum substrate, or lose contact with the disc altogether. An excessive vertical component can also upset the playback process, especially when low tracking forces are used; in the worst case the stylus can be thrown out of the groove completely. To control this problem the stereo signal is passed through a matrix that isolates the L–R vertical signal, which is then amplitude limited. This potentially reduces the perceived stereo separation at low frequencies, but there appears to be a general consensus that the effect is not audible. The most important factor in controlling out-of-phase signals is the panning of bass instruments (which create the largest cutter amplitudes) to the center of the stereo stage. This approach is still advantageous with digital media as it means that there are two channels of amplification to reproduce the bass information rather than one.

From about 1 kHz up to the ultrasonic regions, the limit is groove velocity rather than amplitude. If the disc cutter head tries to move sideways too quickly compared with its relative forward motion, the back facets of the cutter destroy the groove that has just been cut by its forward edges.

On disc replay, there is a third restriction – that of stylus acceleration, or, to put it another way, groove curvature. This sets a limit on how well a stylus of a given size and shape can track the groove. Allowing for this at cutting time places an extra limitation on signal level, shown by the dashed line in Figure 7.1(a). The severity of this restriction depends on the stylus shape; an old-fashioned spherical type with a tip diameter of 0.7 mil. requires a roll-off of maximum levels from 2 kHz, while a (relatively) modern elliptical type with 0.2 mil. effective diameter postpones the problem to about 8 kHz. The limit, however, still remains.

Thus, disc-cutting and playback technology put at least three limits on the maximum signal level. This is not as bad a problem as it might be, because the distribution of amplitude with frequency for music is not flat with frequency; there is always more energy at LF than HF. This is especially true of the regrettable phenomenon known as rap music. For some reason there seems to be very little literature on the distribution of musical energy versus frequency, but a rough rule is that levels can be expected to be fairly constant up to 1 kHz and then fall by something like 10 dB/octave. The end result is that despite the limits on disc levels at HF, it is still possible to apply a considerable amount of HF boost to reduce surface noise problems. At the same time the LF levels are reduced to keep groove amplitude under control. Both functions are implemented by applying the inverse of the familiar RIAA replay equalization at cutting time.

www.raguvaran.puzl.com

Equalization and Its Discontents

Both moving-magnet (MM) and moving-coil (MC) cartridges operate by the relative motion of conductors and magnetic field, so the voltage produced is proportional to rate of change of flux. The cartridge is therefore sensitive to groove velocity rather than groove amplitude, and so its sensitivity is expressed as *X* mV per cm/s. This velocity sensitivity gives a frequency response rising steadily at 6 dB/octave across the whole audio band for a groove of constant amplitude. Therefore a maximal signal on the disc, as in Figure 7.1(a), would give a cartridge output like that shown in Figure 7.1(b), which is simply Figure 7.1(a) tilted upwards at 6 dB/octave. From here on the acceleration limits are omitted for greater clarity.

The RIAA replay equalization curve is shown in Figure 7.1(c). It has three corners in its response curve, with frequencies at 50.05 Hz, 500.5 Hz, and 2.122 kHz, which are set by three time-constants of 3180, 318, and 75 µs. The RIAA curve was of USA origin but was adopted internationally with surprising speed, probably because everyone concerned was heartily sick of the ragbag of equalization curves that existed previously. It became part of the IEC 98 standard, first published in 1964, and is now enshrined in IEC 60098, 'Analogue Audio Disk Records and Reproducing Equipment'.

Note the flat shelf between 500 Hz and 2 kHz. It may occur to you that a constant upward slope across the audio band would have been simpler, and requiring fewer precision components to accurately replicate. But such a response would require 60 dB more gain at 20 Hz than at 20 kHz, equivalent to 1000 times. The minimum open-loop gain at 20 Hz would have to be 70 dB (3000 times) to allow even a minimal 10 dB of feedback at that frequency, and implementing that with a simple two-transistor preamplifier stage would have been difficult if not impossible (must try it some time). The 500 Hz – 2 kHz shelf in the RIAA curve reduces the 20 Hz – 20 kHz gain difference to only 40 dB, making a two-transistor preamplifier stage practical. One has to conclude that the people who established the RIAA curve knew what they were doing.

Figure 7.1(c) shows in dotted lines an extra response corner at 20.02 Hz, corresponding to a time-constant of 7950 µs. This extra roll-off is called the 'IEC Amendment' and it was added to what was then IEC 98 in 1976. Its apparent intention was to reduce the subsonic output from the preamplifier, but its introduction is something of a mystery. It was certainly not asked for by either manufacturers or their customers, and it was unpopular with both, with some manufacturers simply refusing to implement it. It still attracts negative comments today. On one hand, it was pointed out that as an anti-rumble measure it was ineffective, as its slow first-order roll-off meant that the extra attenuation at 13 Hz, a typical cartridge–arm resonance frequency, was a feeble −5.3 dB; however, at 4 Hz, a typical disc warp frequency, it did give a somewhat more useful −14.2 dB, reducing the unwanted frequencies to a quarter of their original amplitude. On the other hand, there were loud complaints that the extra unwanted replay time-constant caused significant frequency response errors at the low end of the audio band, namely −3.0 dB at 20 Hz and −1.0 dB at 40 Hz. Some of the more sophisticated

www.raguvaran.puzl.com

equipment allows the Amendment to be switched in or out; a current example is the Audiolab 8000PPA phono preamplifier.

Another possible problem with the IEC Amendment is that it was almost certainly intended to be implemented by restricting the value of the capacitor at the bottom of a series feedback arm, i.e. C0 in Figure 7.2. While electrolytic capacitors nowadays (2009) have relatively tight tolerances of ±20%, in 1976 you would have been more likely to encounter −20% + 50%, the asymmetry reflecting the assumption that electrolytics would be used for non-critical coupling or decoupling purposes where too little capacitance might cause a problem, but more than expected would be fine. This meant that there could be significant errors in the LF response.

And another problem… is the non-linearity of electrolytic capacitors when they are asked to form part of a time-constant. This is described in detail in Chapter 2. Since the MM preamps of the 1970s tended to have poor linearity at LF anyway, because the need for bass boost meant a reduction in the LF negative-feedback factor, introducing another potential source of distortion was not exactly an inspired move. There is little doubt that even a simple second-order filter subsonic, switchable in and out, would be a better approach to controlling subsonic disturbances. If a Butterworth (maximally flat) alignment was used, with a −3 dB point at 20 Hz, this would only attenuate by 0.3 dB at 40 Hz, but would give a more useful −8.2 dB at 13 Hz and a thoroughly effective −28 dB at 4 Hz. It should be said, however, that not all commentators are convinced that the more rapid LF phase changes that result are wholly inaudible. Subsonic filters are examined more closely at the end of this chapter.

To get back to the levels coming from the cartridge, when the RIAA equalization of Figure 7.1(c) is applied to the cartridge output of Figure 7.1(b), the result looks like Figure 7.1(d), with the maximum amplitudes occurring around 1–2 kHz. This in agreement with Holman's data [2].

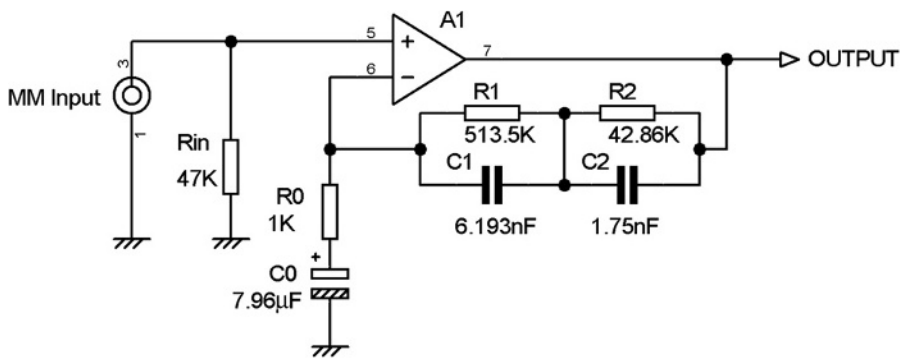


Figure 7.2: Series-feedback RIAA equalization, with the IEC Amendment implemented by C0. Component values for a gain at 1 kHz of 35.0 dB. RIAA accuracy is within ±0.1 dB from 20 Hz to 20 kHz

www.raguvaran.puzl.com

Figure 7.1(e) shows some possible output level restrictions that might affect Figure 7.1(d). If the IEC Amendment is implemented after the first stage, there is a possibility of overload at low frequencies that does not exist if the Amendment is implemented in the feedback loop by restricting C0. At the high end, the output may be limited by problems driving the RIAA feedback network, which falls in impedance as frequency rises. More on this later.

The 'Neumann Pole'

The RIAA curve is only defined to 20 kHz, but by implication carries on down at 6 dB/octave forever. This implies a recording characteristic rising at 6 dB/octave forever, which could clearly endanger the cutting head if ultrasonic signals were allowed through. From 1995 a belief began to circulate that record lathes incorporated an extra unofficial pole at 3.18 μs (50.0 kHz) to limit HF gain. This would cause a loss of 0.17 dB at 10 kHz and 0.64 dB at 20 kHz, and would require compensation if an accurate replay response was to be obtained. The name of Neumann became attached to this concept simply because they are the best-known manufacturers of record lathes.

The main problem with this story is that it is not true. The most popular cutting amplifier is the Neumann SAL 74B, which has no such pole. For protection against ultrasonics and RF it has instead a rather more effective second-order low-pass filter with a corner frequency of 49.9 kHz and a Q of 0.72 [3], giving a Butterworth (maximally flat) response rolling off at 12 dB/octave. Combined with the RIAA equalization this gives a 6 dB/octave roll-off above 50 kHz. The loss from this filter at 20 kHz is less than −0.1 dB, so there is little point in trying to compensate for it, particularly because other cutting amplifiers are unlikely to have identical filters.

Implementing RIAA Equalization

It can be firmly stated from the start that the best way to implement RIAA equalization is the traditional series-feedback method, as in Figure 7.2. This stage is designed for a gain of 35.0 dB at 1 kHz, and RIAA accuracy is within ±0.1 dB from 20 Hz to 20 kHz. With a nominal 5 mVrms input at 1 kHz the output will be 280 mV. The IEC Amendment is implemented by making C0 a mere 7.96 μF. You will note with apprehension that only one of the components, R0, is a standard value, and that is because it was used as the input to the RIAA design calculations that defined the overall RIAA network impedance. This is always the case for accurate RIAA networks. Here, even if we assume that capacitors of the exact value could be obtained, and we use the nearest E96 resistor values, systematic errors of up to 0.06 dB will be introduced. Not a long way adrift, it's true, but if we are aiming for an accuracy of ±0.1 dB it's not a good start. If E24 resistors are the best available the errors grow to a maximum of 0.12 dB, and don't forget that we have not considered tolerances – we are assuming the values

www.raguvaran.puzl.com

are exact. If we resort to the nearest E12 value (which really shouldn't be necessary these days), then the errors exceed 0.7 dB at the HF end. And what about those capacitors?

The answer is of course that by using multiple components in parallel or series we can get pretty much what value we like, and it is perhaps surprising that this approach is not adopted more often. The reason is probably cost – a couple of extra resistors are no big deal but multiple capacitors make more of an impact on the costing sheet. More on this important topic later in this chapter.

Be aware that the circuit of Figure 7.2 is not optimized, but attempts to represent a 'typical' design. R0 could and should be made lower in value to reduce its Johnson noise contribution.

The only drawback to the series configuration is what might be called the unity-gain problem. While the RIAA equalization curve is not specified above 20 kHz, the implication is clear that it will go on falling indefinitely at 6 dB/octave. A series-feedback stage cannot have a gain of less than unity, so at some point the curve will begin to level out and eventually become flat at unity gain – in other words, there is a zero in the response. Figure 7.3 shows the various poles and zero frequencies of the circuit in Figure 7.2, with their associated time-constants. T3, T4, and T5, are the time-constants that define the basic RIAA curve. T2 is the extra time-constant for the IEC Amendment, and T1 shows where its effect ceases at very low frequencies when the gain is approaching unity. At the high end, the final zero is at frequency f6, with associated time-constant T6, and because the gain was chosen to be +35 dB at 1 kHz it is quite a long way from 20 kHz and has very little effect at this frequency, giving an excess gain of only 0.10 dB. This quite quickly dies away to nothing as frequency falls below 20 kHz.

However, if the gain of the stage is set low to maximize the input overload margin, the 6 dB/octave fall tends to level out at unity early enough to cause significant errors in the audio

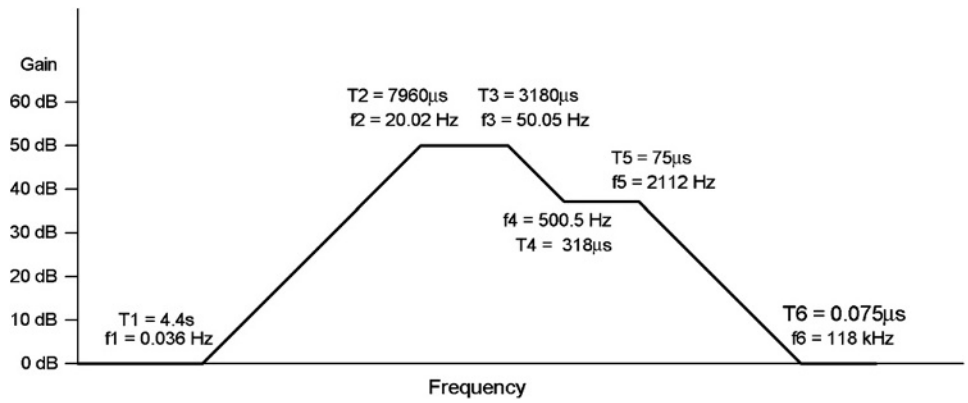


Figure 7.3: The practical response for series-feedback RIAA equalization, including the IEC Amendment, which gives an extra roll-off at 20.02 Hz

www.raguvaran.puzl.com

band. Adding an HF correction pole (i.e. low-pass time-constant) just after the input stage makes the simulated and measured frequency response exactly correct. It is not a question of bodging the response to make it roughly right. If the correction pole frequency is correctly chosen the roll-off cancels *exactly* with the 'roll-up' of the final zero.

The technique is demonstrated in Figure 7.4, where several changes have been made. The overall impedance of the RIAA network has been reduced by making R0 220 Ω, to reduce Johnson noise from the resistors; we still end up with some very awkward values. The IEC Amendment is no longer implemented in this stage; if it was then the correct value of C0 would be 36.18 μF, and instead it has been made 220 μF so that its associated −3 dB roll-off does not occur until 3.29 Hz. Even this wide spacing introduces an unwanted 0.1 dB loss at 20 Hz, and perfectionists will want to use 470 μF here, which reduces the error to 0.06 dB. Most importantly, the gain has been reduced to +30 dB at 1 kHz to get more overload margin. With a nominal 5 mVrms input at 1 kHz the output will be 160 mV. The result is that the final zero f6 in Figure 7.3 is now at 66.4 kHz, much closer in, and it introduces an excess gain at 20 kHz of 0.38 dB, which is too much to ignore if you are aiming to make high-class gear. The correcting pole R4, C4 is therefore added, which solves the problem completely. Since there are only two components, and no interaction with other parts of the circuit, we have complete freedom in choosing C3 so we use a standard E3 value and then get the pole frequency exactly right by using two resistors in series – 470 and 68 Ω. Since these components are only doing a little fine-tuning at the top of the frequency range, the tolerance requirements are somewhat relaxed compared with the main RIAA network. The design considerations are (a) that the resistive section R4 should be as low as possible in value to minimize Johnson noise, and on the other hand (b) that the shunt capacitor C4 should not be large enough to load the op-amp output excessively at 20 kHz. At this level of accuracy, the finite gain open-loop gain of even a 5534 at HF has a slight effect, and the frequency of the HF pole has been trimmed to compensate for this.

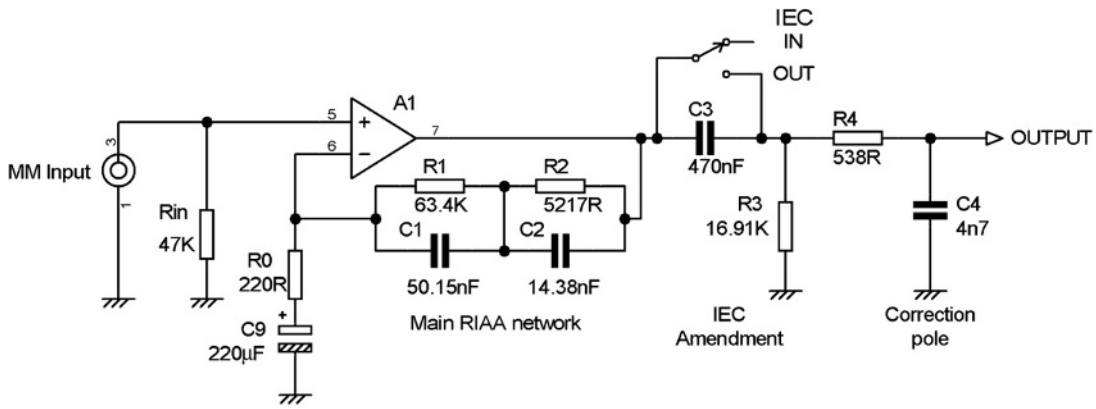


Figure 7.4: Series-feedback RIAA equalization reconfigured for 30.0 dB gain at 1 kHz, and a lower impedance RIAA network. The switchable IEC Amendment is implemented by C3, R3. An HF correction pole R4,C4 is added to keep RIAA accuracy within ±0.1 dB, 20 Hz to 20 kHz

www.raguvaran.puzl.com

But what about the IEC Amendment? In several of my designs it has been integrated into the subsonic filter that immediately follows the RIAA preamplifier; this gives economy of component use but means that it is not really practicable to make it switchable in and out. To do this a separate 7950 µs time-constant is required after the preamplifier, as shown in Figure 7.4, where R3, C3 give the required −3 dB roll-off at 20.02 Hz. Once again we can use a standard E3 capacitor value and 470 nF has been chosen here, and once again an unhelpful resistor value results – in this case 16.91 kΩ. However, with E24 values, this can be implemented exactly as 16 kΩ + 910 Ω. The switch as shown will not be entirely click-free because of the offset voltage at A1 output, but that is relatively unimportant as it will probably only be operated a few times in the life of the equipment.

The shunt-feedback equivalent of Figures 7.2 and 7.3 is shown in Figure 7.5. It has occasionally been advocated because it avoids the unity-gain problem, but it has the crippling disadvantage that with a real cartridge load, with its substantial inductance, it is about 14 dB noisier than the series RIAA configuration [4]. A great deal of grievous twaddle has been talked about RIAA equalization and transient response, in perverse attempts to render the shunt RIAA configuration acceptable despite its serious noise disadvantage. A series-feedback disc stage cannot make its gain fall below 1, as described above, while the shunt-feedback version can; however, an extra output pole solves that problem completely. Shunt feedback eliminates any possibility of common-mode distortion, but then at the signal levels we are dealing with that is not a problem, at least with bipolar input op-amps. A further disadvantage is that a shunt-feedback RIAA stage gives a phase inversion that can be highly inconvenient if you are concerned to preserve absolute phase.

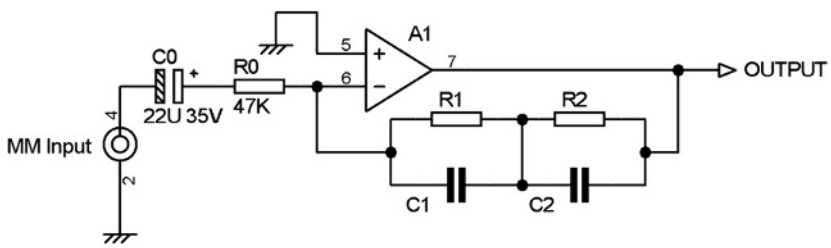


Figure 7.5: Shunt-feedback RIAA configuration. This is 14 dB noisier than the series-feedback version

Passive and Semi-Passive RIAA Equalization

For many years, series-feedback RIAA preamplifiers as described above were virtually universal, it being accepted by all that they gave the best noise and overload performance. However, human nature being what it is, some people will always want to do things the hard way, and this is exemplified by the fashion for passive (actually, semi-passive is more accurate) RIAA equalization. The basic notion is to split the RIAA equalization into separate stages, and

www.raguvaran.puzl.com

I have a dark and abiding suspicion that this approach may be popular simply because it makes the design of accurate RIAA equalization much easier, as all you have to do is calculate simple time-constants instead of grappling with foot-long equations. There is a price, and a heavy one: the overload and/or noise performance is inevitably compromised.

Clearly a completely passive RIAA stage is a daft idea because a lot of gain is required somewhere to get the 5 mV cartridge signal up to a usable amplitude. The nearest you can get is the scheme shown in Figure 7.6 (a), where the amplification and the equalization are wholly separate, with no frequency-dependent feedback used at all. R2, R3 and C1 implement T3 and T4, while C2 implements T5. There is no inconvenient T6 because the response carries on falling indefinitely with frequency. This network clearly gives its maximum gain at 20 Hz, and at 1 kHz it attenuates by about 20 dB. Therefore, if we want the modest +30 dB gain at 1 kHz used in the previous example, the A1 stage must have a gain of no less than 50 dB. A 5 mVrms, 1 kHz input would therefore result in 1.58 V at the output of A1. This is only 16 dB below clipping, assuming we are using the usual sort of op-amps, and an overload margin of 16 dB is much too small to be usable. It is obviously impossible to drive anything like a volume-control or tone-control stage from the passive network, so the buffer stage A2 is shown to emphasize that extra electronics are required with this approach.

The only answer is to split the gain so that the A1 stage has perhaps 30 dB, while A2 after the passive RIAA network makes up the loss with 20 dB more gain. Sadly, this second stage of amplification must introduce extra noise, and there is always the point that you now have to put the signal through two amplifiers instead of one, so there is the potential for increased distortion.

The most popular architecture that separates the high and low RIAA sections is seen in Figure 7.6 (b). Here there is an active LF RIAA stage using feedback to implement T3 and T4 with R1, C1, R2, followed by R3, C2, which give a passive HF cut for T5. This is what can only be called a semi-passive configuration. The values shown give an RIAA curve correct to within 0.04 dB from 20 Hz to 20 kHz. Note that because of the lack of time-constant interaction, we can choose standard values for both capacitors, but we are still left with awkward resistor values.

As before, amplification followed by attenuation means a headroom bottleneck, and this passive HF roll-off is no exception. Signals direct from disc have their highest amplitudes at high frequencies, so both these configurations give poor HF headroom, overload occurring at A1 output before passive HF cut can reduce the level. Figure 7.7 shows how the level at A1 output (trace B) is higher at HF than the output signal (trace A). The difference is trace C, the headroom loss; from 1 dB at 1 kHz this rises to 14 dB at 10 kHz and continues to increase in the ultrasonic region. The passive circuit was driven from an inverse RIAA network, so a totally accurate disc stage would give a straight line just below the +30 dB mark.

www.raguvaran.puzl.com

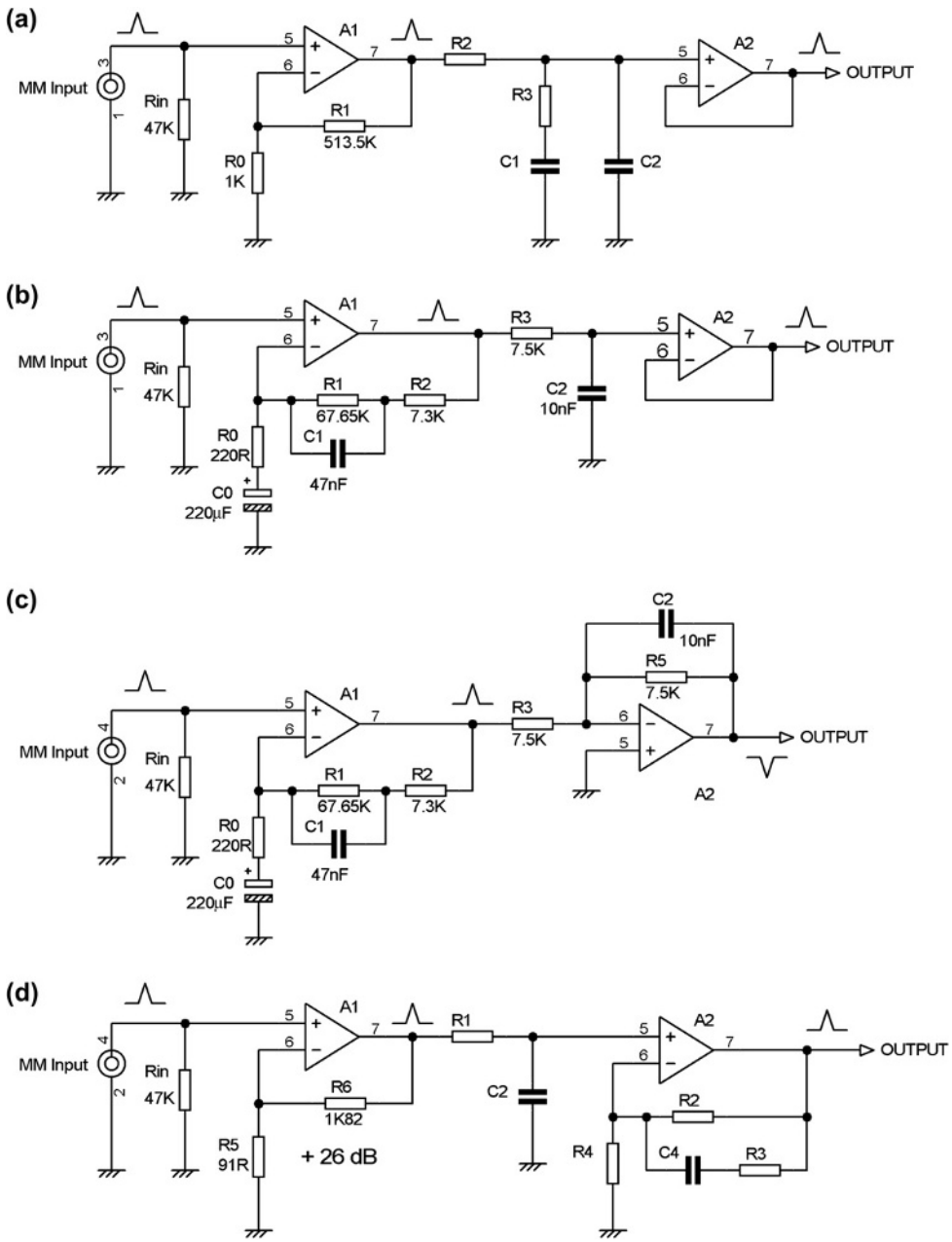


Figure 7.6: Passive and semi-passive RIAA configurations

A related problem in this semi-passive configuration is that the op-amp A1 must handle a signal with much more HF content than the op-amp in the series-feedback stage, worsening any difficulties with slew limiting and HF distortion. It uses two amplifier stages rather than one, and more precision components, because of the extra resistor. Another difficulty is that

www.raguvaran.puzl.com

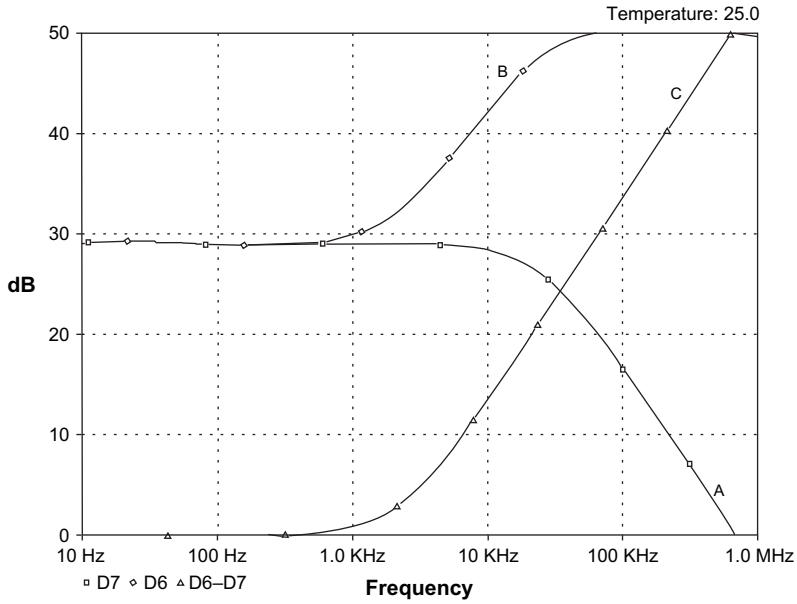


Figure 7.7: Headroom loss with passive RIAA equalization. The signal level at A1 (trace B) is greater than at A2 (trace A) so clipping occurs there first. Trace C shows the headroom loss, which reaches 18 dB at 20 kHz

A1 is more likely to run out of open-loop gain at HF, as the response plateaus above 1 kHz rather than being steadily reduced by increasing negative feedback. Once again a buffer stage A2 is required to isolate the final time-constant from loading.

Another method of equalization is a variation on the circuit just described, shown in Figure 7.6 (c), where the T5 roll-off is done by feedback via R4, C2 rather than by passive attenuation. This is not really passive in any way, as the equalization is done in two active stages, but it does share the feature of splitting up the time-constants for easier design. As with the previous circuit, A1 is running under unfavorable conditions as it has to handle a larger HF content than in the series-feedback version, and there is now an inconvenient phase reversal. The values shown give the same gain and RIAA accuracy as the previous circuit, though in this case the value of R3 can be scaled if more gain is required.

There are many other alternative arrangements that can be used for passive or semi-passive equalization. There may be a flat input stage followed by a passive HF cut and then another stage to give the LF boost, as in Figure 7.6 (d), which has even more headroom problems and uses yet more bits. In contrast, the 'all-in-one-go' series-feedback configuration avoids unnecessary headroom restrictions and has the minimum number of stages. Passive RIAA is not an attractive option.

www.raguvaran.puzl.com

Calculating the RIAA Equalization Components

Calculating the values required for the series-feedback configuration is not straightforward. You absolutely cannot take Figure 7.2 and calculate the time-constants of R2, C2 and R3, C3 as if they were independent of each other; the answers will be wrong. Empirical approaches (cut-and-try) are possible if no great accuracy is required, but attempting to reach even ±0.2 dB by this route is tedious, frustrating, and generally bad for your mental health.

The definitive paper on this subject is by Stanley Lipshitz [5]. This heroic work covers both series and shunt configurations and much more besides, including the effects of inadequate open-loop gain. It is relatively straightforward to build a spreadsheet using the Lipshitz equations that allows extremely accurate RIAA networks to be designed in a second or two; the greatest difficulty is that some of the equations are long and complicated – we're talking real turn-the-paper-sideways algebra here – and some very careful typing is required.

My spreadsheet model takes the desired gain at 1 kHz and the value of R1, which sets the overall impedance levels of the RIAA network. In my preamplifier designs the IEC Amendment is definitely *not* implemented by restricting the value of C0; this component is made large enough to have no significant effect in the audio band, and the Amendment roll-off is realized in the next stage.

RIAA Equalization Component Tolerances

Having calculated the component values we want, we must face the fact that real components have tolerances on their value, and we need to assess what RIAA accuracy is possible without spending a fortune on precision parts; ±1% is the best tolerance readily available for resistors and capacitors, so at first it appears anything better than ±0.1 dB accuracy is out of the question. This is not so; let us consider the circuit in Figure 7.4. The component–sensitivity plots in Figures 7.8 and 7.9 show the effect of 1% deviations in the value of R1 and R2; the response errors never exceed ±0.05 dB, as there are always at least two components contributing to the RIAA response. The rapid fall-off at low frequencies is due to the subsonic filter, which was included in the simulations as it implements the IEC Amendment.

The RIAA capacitor sensitivity is shown in Figures 7.10 and 7.11, and it can be seen that tighter tolerances are needed for C1 and C2 than for R1 and R2 to achieve the same ±0.05 dB accuracy, as the capacitors have more effect on the response than the resistors.

Finding affordable close-tolerance capacitors is not easy; the best solution still seems to be axial polystyrene types, which are freely available at a 1% tolerance. These capacitor ranges usually only go up to 10 nF, so some paralleling is required, and in fact is highly desirable. This is because the sum of multiple capacitors is more accurate than a single component of the same tolerance.

www.raguvaran.puzl.com

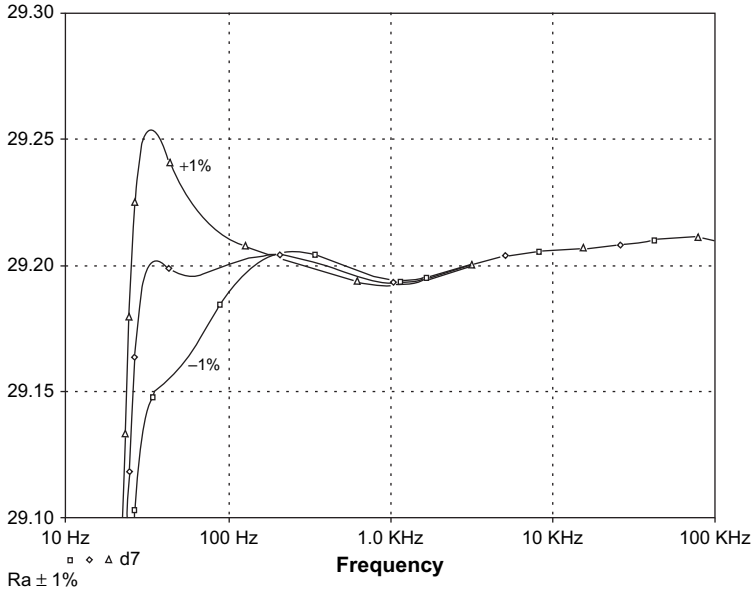


Figure 7.8: The effect on RIAA accuracy of a ±1% change in R1. The worst error is 0.05 dB, and the variation is negligible above 100 Hz

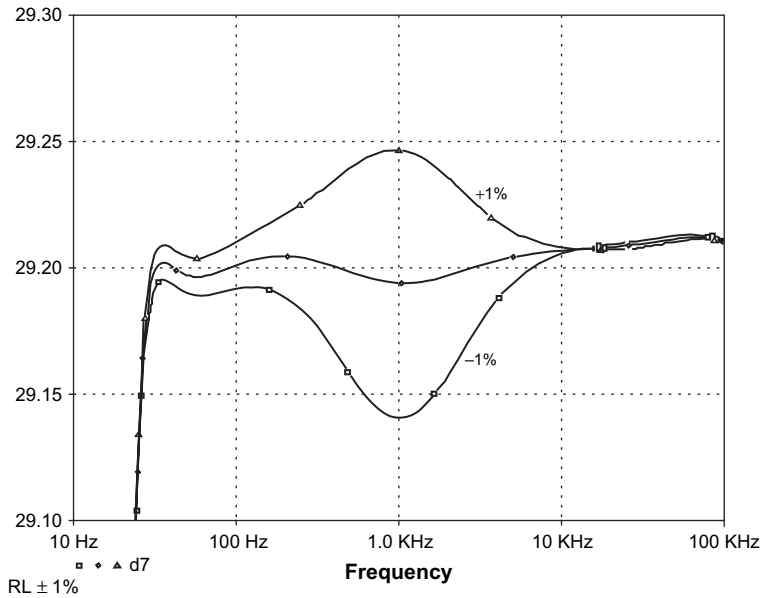


Figure 7.9: The effect on RIAA accuracy of a ±1% change in R2. The worst error is ±0.05 dB around 1 kHz

www.raguvaran.puzl.com

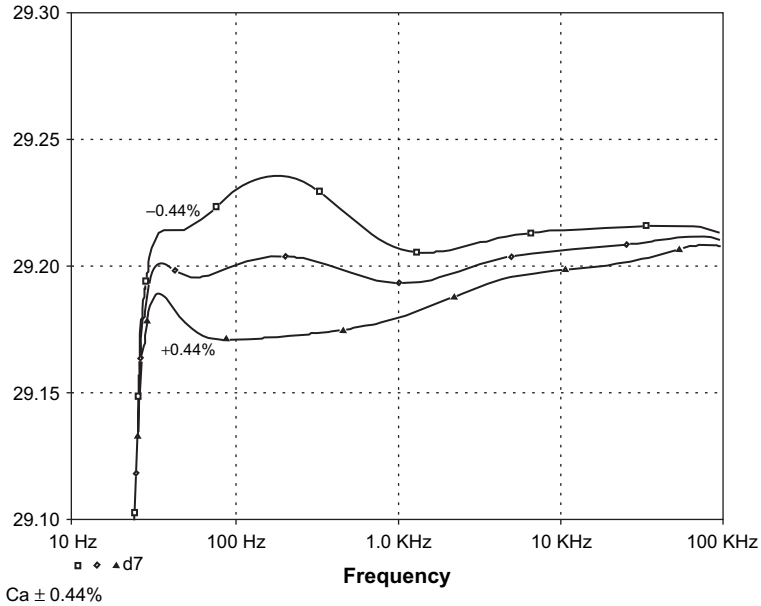


Figure 7.10: The effect on RIAA accuracy of a ±0.44% change in C1. The effect is less than ±0.05 dB at low frequencies. There is a small and diminishing effect above 1 kHz

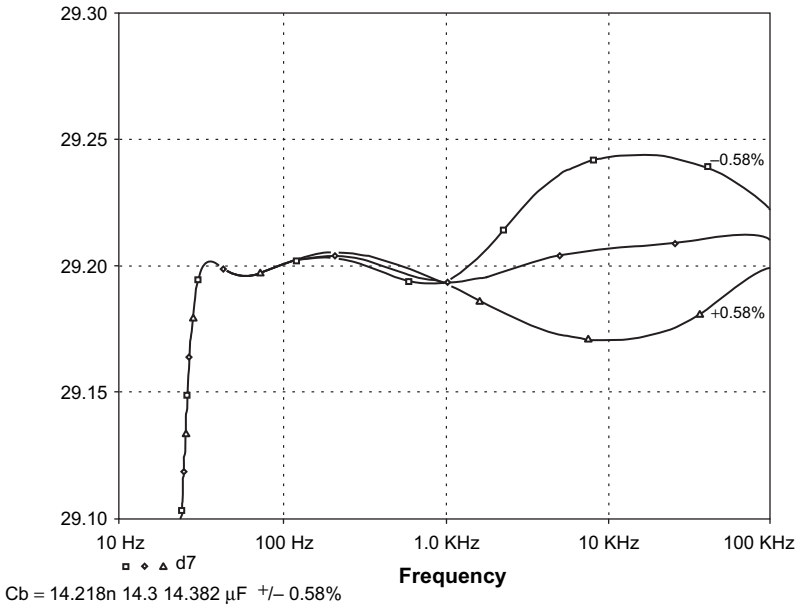


Figure 7.11: The effect on RIAA accuracy of a ±0.58% change in C2. The error is less than ±0.05 dB in the top four octaves, and the variation is negligible below 1 kHz

www.raguvaran.puzl.com

This statistical trick works because the variance of equal summed components is the sum of the individual variances. Thus, if we have four 10 nF capacitors in parallel, the standard deviation (the square root of the variance) increases only by the square root of 4, that is 2, while the total capacitance has increased 4 times; therefore we have inexpensively made an otherwise costly 0.5% close-tolerance 40 nF capacitor. There is a happy analog here with the use of multiple amplifiers to reduce electrical noise. Here we are using essentially the same method to reduce 'statistical noise'. The application of this principle can be seen in Figure 7.17 below. The same technique could be used with resistors if necessary.

The figures given above assume that the component values vary with a normal (Gaussian) distribution. If the 1% components were obtained by selection from a population that varied over say 5%, the distribution would be much flatter than Gaussian and the accuracy improvement much reduced. Having said that, this appears not to be the way that polystyrene capacitors at least are manufactured, and the expected improvement really is attained in practice. A very good discussion of managing statistical variation to enhance accuracy is given in Ref. [6]. This deals with the addition of tolerances in optical instruments, but the principles are the same.

Simulating Inverse RIAA Equalization

SPICE simulation is well suited to the task of checking that the RIAA component values chosen are accurate. The best way to do this is to build an inverse RIAA simulation that can be used to feed the RIAA preamplifier under test. This is much, much simpler than designing the preamplifier RIAA network because the time-constants can be completely decoupled from each other by using unity-gain buffers with zero-impedance outputs. The required response can be implemented in many ways, but my version is shown in Figure 7.12. The component values have nothing to do with practical circuitry and are chosen simply for ease of calculation. The first network implements the 7960 µs time-constant of the notorious IEC Amendment, and can be omitted if so desired. Since this first network is the inverse of a bass roll-off, its output must continue to rise indefinitely at 6 dB/octave as frequency falls, and it is therefore

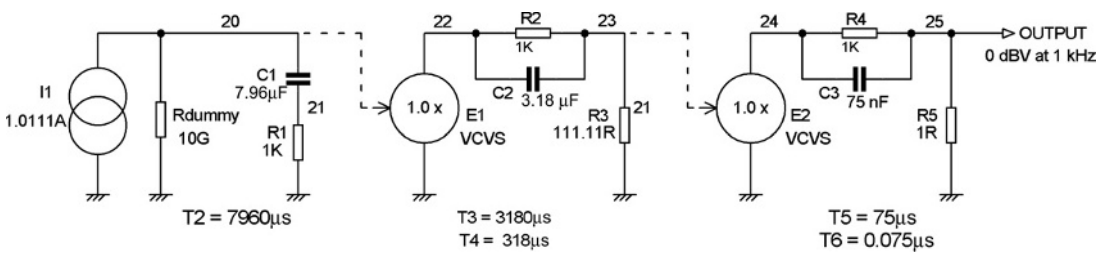


Figure 7.12: Inverse RIAA network for SPICE simulation

www.raguvaran.puzl.com

implemented with a current source, so that as the impedance of C1 rises the output voltage at node 20 rises indefinitely. The apparently odd value of 1.011 A for the current source is in fact cunningly chosen to give a final output of 0 dBV at 1 kHz, which simplifies SPICE output plotting. The 10 GΩ resistor R_{dummy} is required as SPICE otherwise considers node 20 to be at an undefined DC level, and objects strongly.

The voltage at node 20 controls the output of the voltage-controlled voltage source (VCVS) E1, which has its gain set to unity. It has zero output impedance and so acts as a mathematically perfect buffer. E is the conventional designator for a VCVS in SPICE.

E1 then drives the network R2, C3, R3, which implements the 3180 and 318 μs time-constants. E2 acts as another perfect buffer for the voltage at node 23, and drives R4, C3, R5, which implement the 75 μs time-constant. The very low value for R5 allows the output to go on rising at 6 dB/octave to well beyond 20 kHz; the response does not level out until the T6 zero at 2.12 MHz is reached.

Physical Inverse RIAA Equalization

Building a sufficiently accurate inverse RIAA network for precision measurements is not to be entered upon lightly or unadvisedly. The component values will need to have an accuracy a good deal better than 1%, and this makes sourcing components difficult and expensive. A much better alternative is to use a test system such as those by Audio Precision that allow an equalization file to modify the generator output level during a frequency sweep.

Overload Margins

Since RIAA preamplifiers rarely have variable gain, it is important that they have an adequate margin against overload. The nominal output of a magnetic cartridge is 5 mVrms at 1 kHz, but much greater levels are experienced with high output cartridges that can track a high-level recording. In a classic paper, Tomlinson Holman [7] gives 95 mVrms at 1 kHz as an absolute worst case. If we assume op-amp preamplifier technology with an output limit of 10 Vrms, this means that the gain at 1 kHz cannot exceed 40 dB if this input level is to be accepted without clipping. The nominal output is then 500 mVrms at 1 kHz. The 1 kHz overload margin is thus 95/5 = 25.6 dB, which is not generally considered to be better than mediocre. Reducing the gain of the preamplifier stage increases the overload margin at the cost of reducing the nominal output, so that the 35 dB gain (1 kHz) stage of Figure 7.2 gives an output of 280 mV, an overload level of 178 mVrms, and a margin of 31 dB, which might be called good. The 30 dB (1 kHz) stage in Figure 7.4 gives a nominal 158 mVout and has an overload level of 316 mVrms and a margin of 36 dB, which is definitely excellent, giving 10 dB more headroom. There is a question as to whether 160 mV is adequate for driving external equipment; if not, then another stage with variable gain may be required. Some gains and overload levels are shown in Table 7.1.

www.raguvaran.puzl.com

TABLE 7.1 Preamplifier gains, and output levels and overload margins for a nominal 5 mVrms input and maximum output of 10 Vrms (all at 1 kHz)

Gain (dB)	Gain (×)	Output (mVrms)	Input overload (mVrms)	Overload margin (dB)
45	178	889	56	21
40	100	500	100	26
35	56.2	281	178	31
30	31.6	158	316	36
25	17.8	89	562	41
20	10.0	50	1000	46

A complication is that the preamp stage output capability may vary with frequency. In the discrete preamplifier stages described below, most have an emitter-follower output, which is much better at sourcing current than sinking it. An RIAA feedback network, particularly one designed with a relatively low impedance to reduce noise, presents a heavy load at high frequencies because of the capacitors, and this heavy loading was often a major cause of distortion and headroom limitation in discrete RIAA stages that had emitter-follower outputs with highly asymmetrical drive capabilities, with the 20 kHz output capability, and thus overload margin, often reduced by 6 dB or more. Replacing the emitter resistor in the emitter-follower with a current source much reduces the problem, and the slight extra complication of using a push–pull Class-A output can bring it down to negligible proportions (for more details see Chapter 3 on discrete design). Op-amps such as the TL072 also struggled to drive RIAA networks at HF, and it was the advent of the 5532 op-amp, with its excellent load-driving capabilities, that finally solved *that* problem.

Further headroom restrictions may occur when not all of the RIAA equalization is implemented in one feedback loop. Putting the IEC Amendment roll-off after the preamplifier stage, as in Figure 7.4 above, means that very low frequencies are amplified by 3 dB more at 20 Hz than they otherwise would be, and this is then undone by the later roll-off. This sort of audio impropriety always carries a penalty in headroom as the signal will clip before it is attenuated, and the overload margin at 20 Hz is reduced by 3.0 dB. The effect reduces quickly as frequency increases, being 1.6 dB at 30 Hz and only 1.0 dB at 40 Hz. Whether this loss of overload margin is more important than providing an accurate IEC Amendment response is a judgement call, but in my experience it creates no trace of any problem in a stage with a gain of 30 dB (1 kHz). Passive-equalization input architectures that put flat amplification before an RIAA stage suffer much more severely from this kind of headroom restriction, and it is quite common to encounter preamplifiers that claim to be high-end, with a high-end price tag, but only have an overload margin of 20–22 dB. Bad show, chaps.

Earlier in this chapter we saw that Tomlinson Holman concluded that to accommodate the worst of the worst cases, a preamplifier should be able to accept not less than 35 mVrms in the

www.raguvaran.puzl.com

3–4 Hz region [2]. If the IEC Amendment is after the preamplifier stage, and C0 is made very large so it has no effect, the RIAA gain in the 2–5 Hz region has flattened out at 19.9 dB, implying that the equivalent overload level at 1 kHz will need to be 346 mVrms. The 30 dB (1 kHz) gain stage of Figure 7.4 has a 1 kHz overload level of 316 mVrms, which is only 0.78 dB below this rather extreme criterion. We are good to go.

At the other end of the spectrum, adding a correction pole after the preamplifier, to correct the RIAA response, also introduces a compromise in the overload margin, though generally a much smaller one. The 30 dB (1 kHz) stage in Figure 7.4 has a mid-band overload margin of 36 dB, which falls to +33 dB at 20 kHz. Only 0.4 dB of this is due to the amplify-then-attenuate action of the correction pole, the rest being due to the heavy capacitative loading on A1 of both the main RIAA feedback path and the pole-correcting RC network. This slight compromise could be eliminated by using an op-amp structure with greater load-driving capabilities, so long as it retains the low noise of a 5534A.

An attempt has been made to show these extra preamp limitations on output level in Figure 7.1(e) above, and comparing Figure 7.1(d), it appears they are almost irrelevant because of the fall-off in possible input levels at each end of the audio band.

To put all this into some sort of perspective, here are the 1 kHz overload margins for a few of my public designs. My first preamplifier, the 'Advanced Preamplifier' [8] achieved +39 dB in 1976, partly by using all-discrete design and ±24 V supply rails. A later discrete design in 1979 [9] gave a tour-de-force +47 dB, accepting over 1.1 Vrms at 1 kHz, but I must confess this was showing off a bit and involved some quite complicated discrete circuitry, including the push–pull Class-A output stages mentioned above. Later designs such as the Precision Preamplifier [10] and its descendant the Precision Preamplifier '96 [11] accepted the limitations of op-amp output voltage in exchange for much greater convenience in most other directions, and still have an excellent overload margin of 36 dB.

Cartridge Impedances

The impedance of the cartridge strongly influences the noise performance of an MM RIAA stage. Manufacturers do not always supply this data, and so I have had to make the best of what is available. Some of the cartridges listed in Table 7.2 are vintage, some more up to date, the collection covering from about 1972 to the present day (2009).

The Shure V15V values are known to be accurate, having been checked by Burkhard Vogel [12]. Resistance ranges from 430 to 1500 Ω, and inductance from 330 to 720 mH.

Cartridge Loading

The standard loading for a moving-magnet cartridge is 47 kΩ in parallel with a certain amount of capacitance, the latter usually being specified by the manufacturer. It is normally in the

www.raguvaran.puzl.com

TABLE 7.2 Some moving-magnet cartridge impedances, both current and historical

Type	Resistance (Ω)	Inductance (mH)
Goldring 1006	Not stated	570
Goldring 1042	Not stated	570
Goldring 2044	Not stated	720
Goldring 2100	Not stated	550
Goldring 2200	Not stated	680
Shure ME75-ED Type 2	610	470
Shure ME95-ED	1500	650
Shure V15V MR	815	330
Shure V15V IV	1380	500
Shure V15V III	1350	500
Shure M44G	650	650
Stanton 5000 AL-II	535	400
CS1 'Carl Cox'	430	400

range 50–200 pF. The capacitance is often the subject of experimentation by enthusiasts, and so switchable capacitors are often provided at the input of high-end preamplifiers, which allow several values to be set up by combinations of switch positions. The exact effect of altering the capacitance depends on the inductance and resistance of the cartridge, but a typical result is shown in Figure 7.13, where increasing the load capacitance lowers the resonance peak frequency and makes it more prominent and less damped. It is important to remember that it is the total capacitance, including that of the connecting leads, which counts.

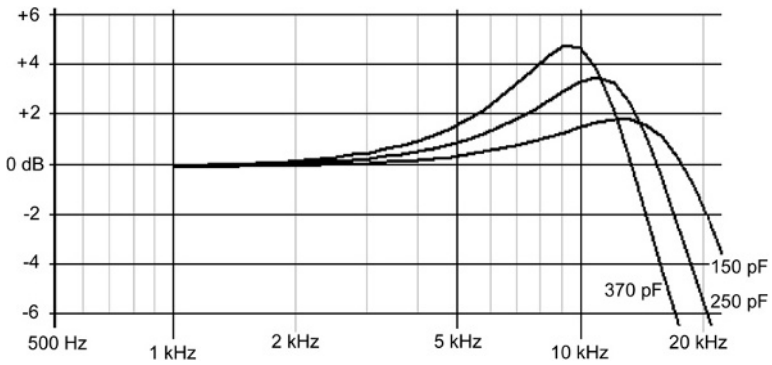


Figure 7.13: The typical effect of changing the loading capacitance on an MM cartridge

Because of the high inductance of an MM cartridge, adjusting the load resistance can also have significant effects on the frequency response, and some preamplifiers allow this too to be

www.raguvaran.puzl.com

adjusted. The only way to assess the effects of these modifications is to measure the output when a special (and expensive) test disc is played.

Cartridge–Preamplifier Interaction

One often hears that there can be problems due to interaction between the impedance of the cartridge and the negative-feedback network. Most commentators are extremely vague as to what this actually means, but according to the work of Tomlinson Holman [7], the factual basis is that it used to be all too easy to design an RIAA stage, if you are using only two or three discrete transistors, in which the NFB factor is falling significantly with frequency in the upper reaches of the audio band, perhaps as a result of using excessive dominant-pole compensation to achieve HF stability. Assuming a series-feedback configuration is being used, this means that the input impedance will fall with frequency, which is equivalent to having a capacitative input impedance. This interacts with the cartridge inductance and can cause a resonant peak in the frequency response, in the same way that cable capacitance or a deliberately added load capacitance can do.

For this reason a flat-response buffer stage between the cartridge and the first stage performing RIAA equalization is sometimes advocated. One design including this feature was the Cambridge Audio P50, which used a Darlington emitter-follower as a buffer; with this approach there is an obvious danger of compromising the noise performance because there is no gain to lift the signal level above the noise floor of the next stage.

Cartridge DC and AC Coupling

Some uninformed commentators have said that there should be no DC-blocking capacitor between the cartridge and the preamplifier. This is insane. The signal currents are tiny (for MM cartridges 5 mV in 47 kΩ = 106 nA, while for MC ones 245 μV in 100 Ω = 2.45 μA – a good deal higher) and even a small DC bias current could interfere with linearity. I am not aware of any published work on how cartridge distortion is affected by DC bias currents, but I think it pretty clear they will not improve things and may make them very much worse. Large currents might upset the magnet strength. Keep DC out of your cartridge.

Discrete MM Disc Input Stages

Discrete moving-magnet (MM) input amplifiers were almost universal until the early 1970s. For a long time op-amps had, quite deservedly, a poor reputation for noise when used in this application.

When the first bipolar transistor MM inputs were designed, active components were still expensive, and adding another transistor to a circuit was something not to be done lightly. The

www.raguvaran.puzl.com

circuitry from that era therefore looks to us very much cut to the bone, and before disrespecting it we need to remember that it was designed under very different economic constraints.

Figure 7.14 shows a typical two-transistor MM input amplifier from the late 1960s. The configuration is generally considered to have been introduced by Dinsdale [13] in 1965, in a classic preamplifier design that was one of the first to deal effectively with the new equalization requirements for microgroove records. It is a two-stage series-feedback amplifier composed of two common-emitter stages. R3 and C1 make up an RF filter; note the high value of the series resistance. This is considerably greater than the DC resistance of most MM cartridges and looks like a mistake. The RIAA network is R6, R7, C4, C5, and has a high impedance to reduce loading on the stage output. Since R6 has the high value of 1M8, the RIAA network cannot be used for the DC feedback that is required to set the quiescent conditions. There is a separate DC feedback network comprising R1, R4, R5, R10, and C3, which establishes the appropriate voltage across R10.

Because of its simplicity, this stage contains compromises. R11 needs to be high in value to maximize open-loop gain, but low to adequately drive the RIAA network and any external loading. RIAA accuracy is poor, with the errors reaching +1.6 dB at 20 Hz and +0.7 dB at 20 kHz; the IEC amendment is not implemented. THD is 0.011% for 1 Vrms out at 1kHz, and about 3 times greater at 40 Hz for 1 Vrms 1 kHz. Linearity can be much improved by raising the supply rail to +24 V and adjusting the DC biasing to suit.

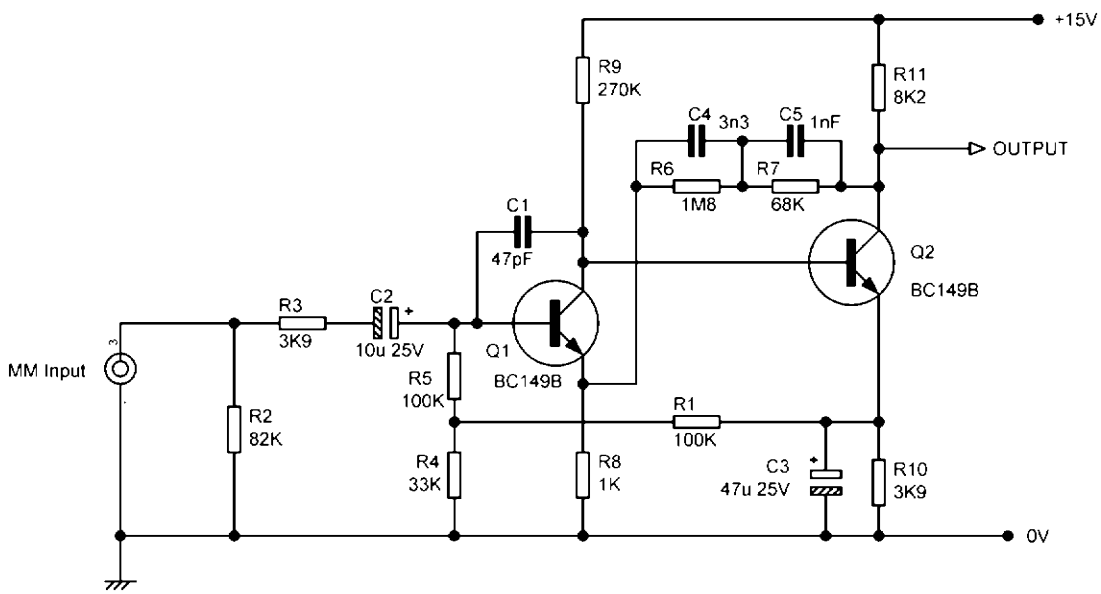


Figure 7.14: A typical two-transistor MM amplifier as commonly used in the 1960s and early 1970s. Gain is +38.9 dB at 1 kHz with values shown

www.raguvaran.puzl.com

A moving-magnet input has to deliver a maximum boost of nearly 20 dB at low frequencies, and this is on top of the gain required to get the desired output level at 1 kHz. If the cartridge output is taken as 5 mVrms at 1 kHz, and the amplifier output as 150 mVrms (which is about as low as you could hope to get away with if you are sending this signal to the outside world), then a total closed-loop gain of 20 + 34 = 54 dB is required at low frequencies. The open-loop gain obviously needs to be considerably higher than this, for a decent feedback factor is required not only to reduce distortion, but also to ensure that the RIAA equalization is accurately rendered by the feedback network. By 1970 it had become clear that the two-transistor configuration was really not up to the job, and more sophisticated circuits were developed. It must now be regarded as of purely historical interest.

Adding an extra transistor improves the possible performance remarkably. The most common three-transistor configuration was introduced by Arthur Bailey in 1966 [14], though his version had a rather awkward level-shift between the first and second transistors. Figure 7.15 shows an improved version from the early 1970s, designed by H.P. Walker [4] and later enhanced by me [8]. It consists of two voltage-amplifier stages as before, but an emitter-follower Q3 is added to buffer the collector of the second transistor from the load of the RIAA network and any external load. This means that the second transistor collector can be operated at a much higher impedance, generating more open-loop gain. The original Walker design had a simple 22 kΩ resistor as a collector load for Q2; when I was using this configuration in Ref. [8] I split this into two 12 kΩ resistors, bootstrapping their central point from the emitter of Q3 as shown in Figure 7.14; this further increased the open-loop gain and reduced the stage

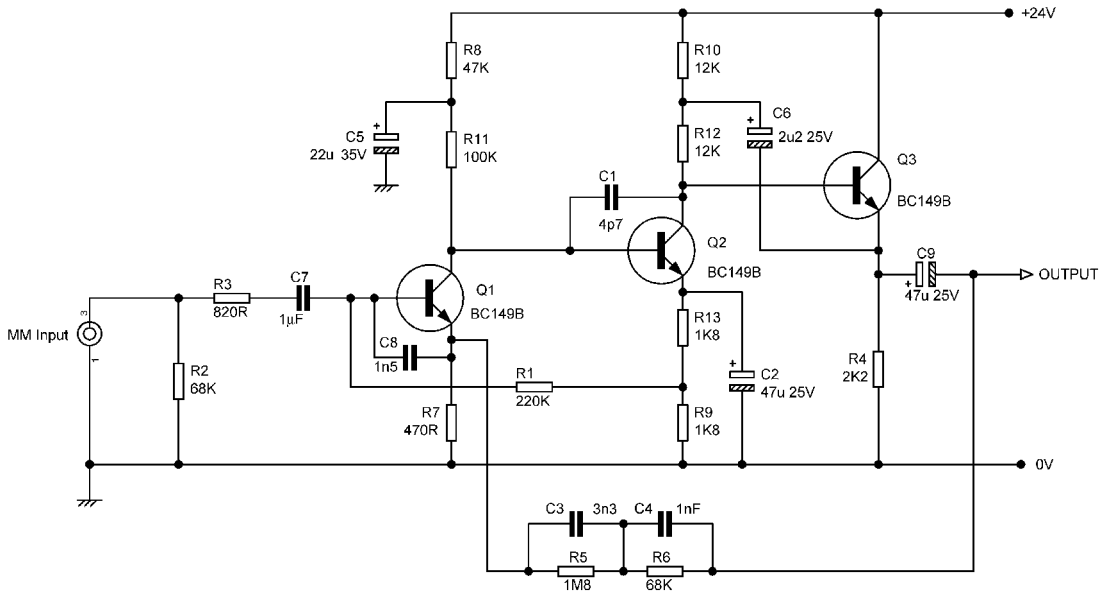


Figure 7.15: A typical three-transistor MM amplifier as commonly used in the 1960s and early 1970s. The original design was by H.P. Walker, the bootstrapping was added by me

www.raguvaran.puzl.com

distortion by a factor of 3. Dominant-pole compensation is applied to Q2 by C1. Once again, the RIAA network has a high impedance and a separate path for DC feedback must be provided by R1; there is in fact no DC feedback at all through the RIAA network as it is connected to the outside of C9. R3 and C8 are the input RF filter. Note the heavy filtering of the supply to the first stage by R8 and C5.

The added emitter-follower, running at a much higher collector current than Q2 (5 mA versus 500 μA), has a greater drive capability, which increases the output voltage swing into a load and so improves the input overload margin. However, a simple emitter-follower output stage has an asymmetrical output current capability, so this is less effective at high frequencies, where the impedance of the RIAA network is falling. Low-gain versions of this circuit may have the overload margin compromised by several decibels at 20 kHz. This can be overcome by making the output stage more sophisticated – replacing the emitter resistor R4 with a current source greatly improves matters, and using a push–pull Class-A output doubles the output current capability again. See Chapter 3 on discrete design for more details.

While this stage is a great improvement on the two-transistor configuration, it is not well adapted to dual supply rails, which would allow a still greater output swing without recourse to discrete supply regulators.

Figure 7.16 demonstrates another way to use three transistors in an RIAA amplifier; this configuration consists of a voltage-amplifier stage, an emitter-follower, and then another voltage-amplifier stage. The first transistor is now a PNP type, its collector bootstrapped for increased gain by connecting the lower end of collector load R6 to the emitter-follower Q2. R2 and C8 make up an RF filter, and once more there is a high value of series resistance. Q2 drives the output transistor Q3, which is now a common-emitter voltage amplifier with collector load R8. Note the asymmetrical supply voltages; the positive rail is 8 V greater than the negative rail, and this is almost certainly intended to increase the positive-swing capabilities of the Q3 stage. Once again note the high impedances in the RIAA network to reduce the loading on the output, and the consequent need for a separate DC feedback path via R10, the AC content being filtered out by C7. Versions of this configuration were used by Pioneer and Sonab.

The record for the highest supply voltage to an RIAA stage was set in 1974 by the Technics SU9600, which employed ±24 V rails and a third rail at a staggering +136 V (this gives a whole new meaning to the phrase 'third-rail electrification'). To the best of my knowledge this record still stands. The general configuration is shown in Figure 7.17; seven transistors are used, in three cascaded differential voltage amplifiers, followed by an emitter-follower output buffer. The final voltage amplifier and the output stage work on asymmetrical supplies, running between +136 and −24 V. The output sits at +56 V to allow a symmetrical output swing, which accounts for the DC-blocking capacitor C11. Note that several resistors around the output stage are high-wattage types. RV1 allowed gain adjustment, while RV2 and RV3 were for setting the DC conditions. My information is that

www.raguvaran.puzl.com

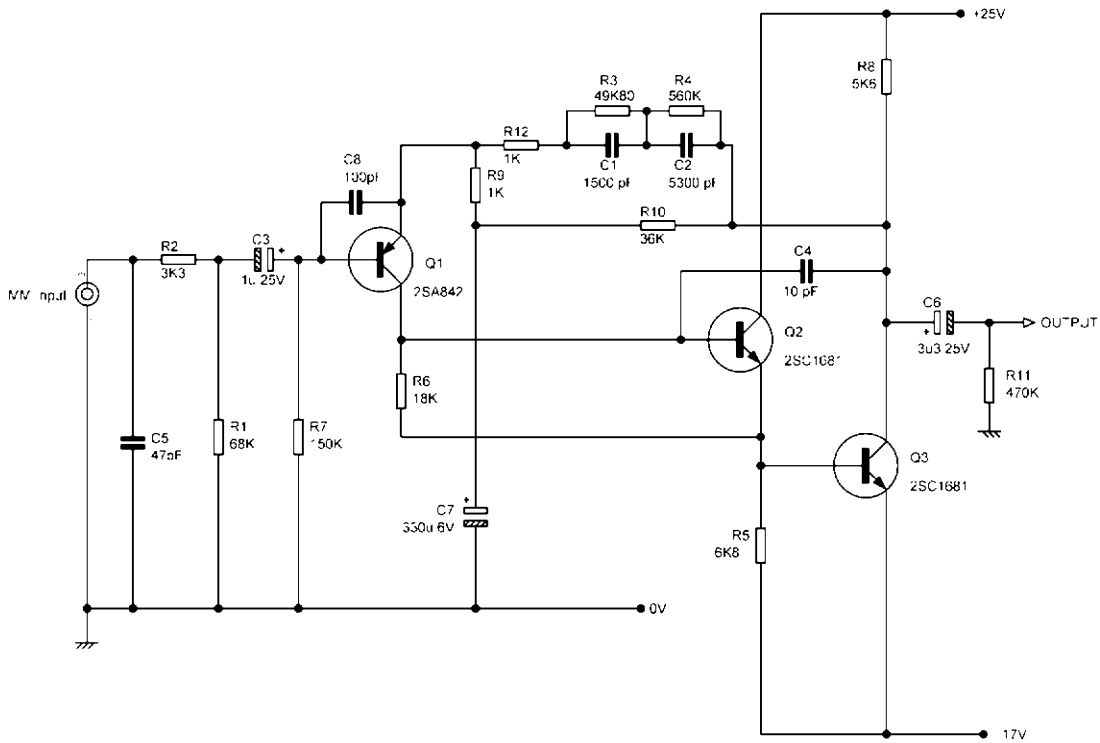


Figure 7.16: Another three-transistor MM amplifier configuration with a quite different structure. This is a simplified version of a circuit used by both Pioneer and Sonab

the maximum input was 900 mV (frequency unstated, but presumably at 1 kHz), the THD was 0.08% (frequency and level unstated), the signal-to-noise ratio was 73 dB with reference to 2 mV, and the RIAA accuracy was ±0.3 dB.

The output stage dissipation is of course enormous for a preamp stage, and the use of a constant-current source or a push–pull Class-A output stage would have allowed this to be much reduced; one can only speculate as to why those techniques were not used. There would have been some fearsome transients at the output on switch-on, and it is notable that an output muting relay was required, probably not so much for reducing audible noise as to give the later stages in the preamplifier a chance of survival.

In the original circuit small capacitors were freely sprinkled over the diagram, leading me to suspect that HF stability was a serious issue during development.

Op-Amp MM Disc Input Stages

The previous section should have convinced you that discrete MM preamplifier circuitry is not that straightforward to design, and there is a lot to be said for using a good op-amp, which if

www.raguvaran.puzl.com

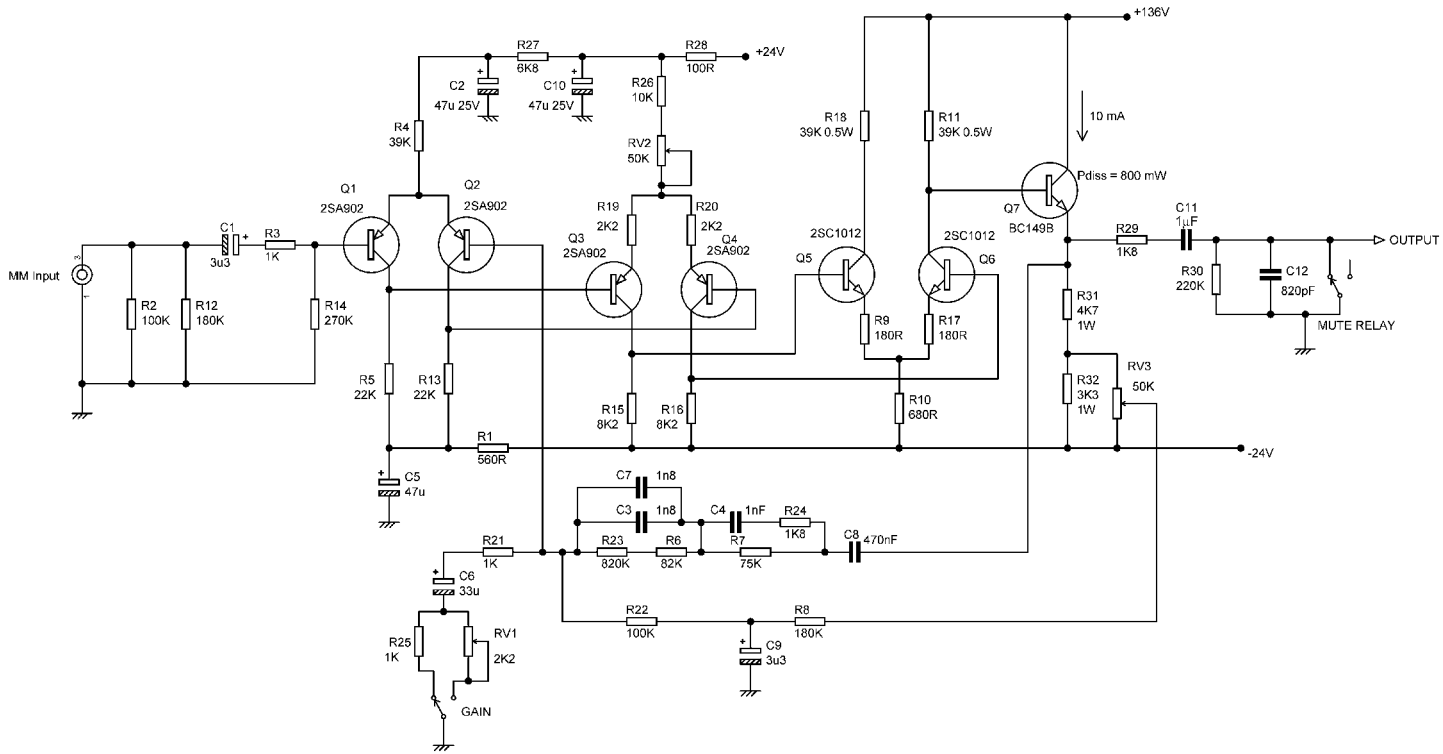


Figure 7.17: A simplified schematic of the Technics SU9600 RIAA stage, with its +136 V supply rail

www.raguvaran.puzl.com

well chosen will have more than enough open-loop gain to implement the RIAA bass boost without introducing detectable distortion. The 5534/5532 op-amps have input noise parameters that are well suited to moving-magnet (MM) cartridges.

Having digested this chapter so far, we are in a position to summarize the requirements for a good RIAA preamplifier. These are as follows:

1. Provide series feedback, as shunt feedback is approximately 14 dB noisier.
2. Ensure correct gain at 1 kHz. This sounds elementary, but setting up the Lipshitz equations is not a negligible task.
3. Ensure accuracy. My 1983 preamplifier was designed for ±0.2 dB accuracy from 20 Hz to 20 kHz, the limit of the test-gear I had at the time. This was tightened to ±0.05 dB without using rare parts in my 1996 preamplifier.
4. Use obtainable components. Resistors will be from the E24 series and capacitors probably E6, so intermediate values must be made by series or parallel combinations.
5. R0 must be as low as possible as its Johnson noise is effectively in series with the input signal. This is particularly important when the MM preamplifier is fed from a low impedance, which typically occurs when it is providing RIAA equalization for the output of an MC preamplifier, rather than accepting input direct from an MM cartridge with its high inductance.
6. The feedback RIAA network impedance to be driven must not be so low as to increase distortion or limit output swing, especially at HF.
7. The resistive path through the feedback arm should ideally have the same DC resistance as input bias resistor R18, to minimize offsets at A1 output. This is a bit of a minor point as the offset would have to be quite large to significantly affect the output voltage swing.

The circuit shown in Figure 7.18 meets all these requirements, and is a practical development of the circuit shown in Figure 7.4. Note that a 5534A is used at the input stage to get the best possible noise performance. A 5534A without external compensation has a minimum stable closed-loop gain of about 3 times; that is close to the gain at 20 kHz here, so a touch of extra compensation is likely be required for reliable stability. The capacitor shown is 4.7 pF.

The resistors can be made more than exact enough by combining two E24 values in series. The RIAA network capacitances are made up of multiple 1% polystyrene capacitors for improved accuracy. Thus, for the five 10 nF capacitors that make up C1, the standard deviation (square root of variance) increases by the square root of 5, while total capacitance has increased five times, and we have inexpensively built an otherwise costly 0.44% close-tolerance 50 nF capacitor. Similarly, C2 is essentially composed of three 4n7 components

www.raguvaran.puzl.com

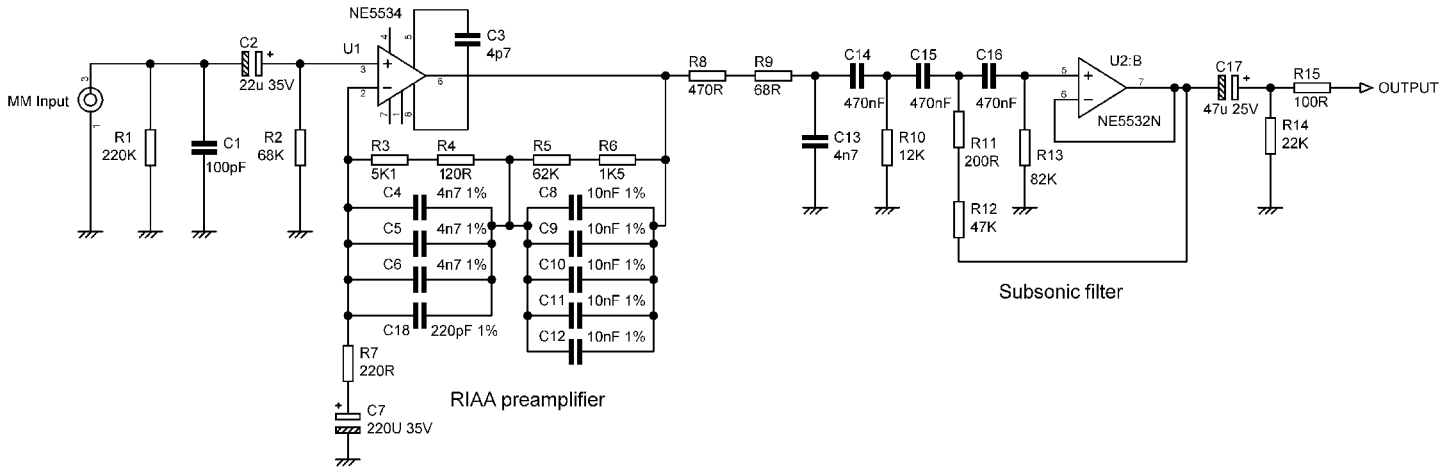


Figure 7.18: An op-amp RIAA preamplifier using a 5534 to get the best possible noise performance

www.raguvaran.puzl.com

and its tolerance is improved by root-three, to 0.58%. An HF correction pole R8, R9, C13 is fitted, and here the resultant loss of HF headroom is only 0.5 dB at 20 kHz, which I think I can live with.

Immediately after the RIAA stage is the subsonic filter. It is shown here because it also implements the IEC Amendment; there is more on such filtering later in this chapter. The filter is a third-order Butterworth HP filter, with its response modified to give a slow initial roll-off that implements the Amendment. This is done by reducing the value of R11 + R12 below that for maximal flatness. The stage also buffers the HF correction pole, and gives the capability to drive a 600 Ω load, if you can find one. A version of this design, using appropriate precision components, is manufactured by the Signal Transfer Company in bare PCB, kit, and fully built and tested formats [15].

Noise in RIAA Preamplifiers

The first priority is to find out the physical limits that set how low the noise can be. With a purely resistive source it is easy to calculate the Johnson noise from the input source resistance. With a noiseless amplifier this would be the equivalent input noise (EIN), but real amplifiers have their own noise, and the amount by which the source/amplifier combination is noisier is the noise figure (NF). I often wonder why NFs are used so little in audio; perhaps they are a bit too revealing. Certainly manufacturers seem to have no interest at all in quoting MM noise specs in a way that would allow easy comparison.

The subject of noise in MM RIAA preamplifiers is considerably complicated by the facts that the signal source is a complex impedance, and the equalization curve is a long way from flat. The best possible EIN for purely resistive sources, such as 200 Ω microphones, is easily calculated to be −129.6 dBu at the usual temperature and bandwidth, but the same calculation for a moving-magnet input is much harder. A highly inductive source – that has no standard value of inductance, or indeed resistance – is combined with the complications of RIAA equalization [5].

The basic noise situation for a series-feedback RIAA stage like that in Figure 7.18 is shown in Figure 7.19. The cartridge is modeled as a resistance R_{gen} in series with a large inductance L_{gen}, and is loaded by the standard 47 kΩ resistor R_{in}. The amplifier block A1 is treated as noiseless, its voltage noise being represented by the voltage generator V_{noise} and its current noise being represented by the current generator I_{noise}. It does not matter which side of V_{noise}, I_{noise} is connected because V_{noise} has no internal resistance. Here V_{noise} represents not only the noise of whatever input device is used in the amplifier, but also noise generated by resistors in the feedback network. The RIAA equalization is done by the noiseless black box after the amplifier.

Figure 7.19 illustrates several important effects that are due to the rising impedance of the cartridge inductance L_{gen} with frequency:

www.raguvaran.puzl.com

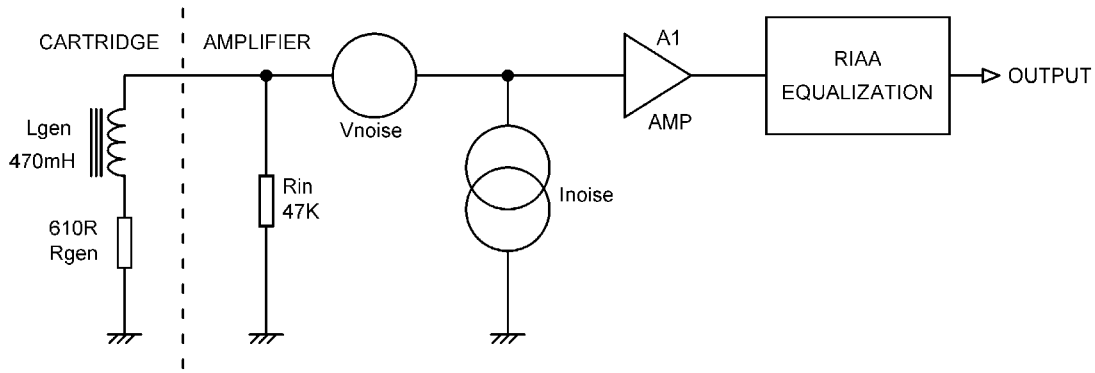


Figure 7.19: A moving-magnet input simplified for noise calculations, with typical cartridge parameter values (Shure ME75-ED2)

1. The impedance seen by the current source I_{noise} increases with frequency. Here it increases from 640 Ω at 75 Hz to 24.9 kΩ at 18 kHz; the increase at the top end is moderated by the shunting effect of R_{in}. This increase has a major effect on the noise behavior. For the lowest noise you must design for a higher impedance than you might think, and van de Gevel [16] quotes 12 kΩ as a suitable value for noise optimization; this assumes A-weighting, inclusion of the IEC amendment, and cartridge parameters of 1000 Ω and 494 mH.

2. The Johnson noise generated by R_{in} is shunted away from the amplifier input by an amount that decreases with frequency. Here the fraction reaching the amplifier rises from 0.014 to 0.53 from 75 Hz to 18 kHz.

3. The proportion of noise from R_{gen} that reaches the amplifier input falls with frequency as the impedance of Z_{gen} increases. Here the fraction reaching the amplifier falls from 0.99 to 0.47 from 75 Hz to 18 kHz.

4. A complication that is not visible in the diagram is that the effective value of R_{gen} is not simply the resistance of the coils. It increases in value with frequency (while still remaining resistive – we are not talking about inductance here) as a consequence of hysteresis and eddy current magnetic losses in the iron on which the coils are wound (see Hallgren [17]). According to van de Gevel [16], this has little effect on noise issues.

On top of this complicated frequency-dependent behavior is overlaid the effect of the RIAA equalization.

Clearly this simple model has some quite complex behavior. It could be analyzed mathematically, using a package such as MathCAD, or it could be simulated by SPICE. The solution I chose is a spreadsheet mathematical model of the cartridge input that I call MAGNOISE, the basic method being as described by Sherwin [18], which divides the audio spectrum into a number of bands so RIAA equalization factors can be applied, and V_{noise}, I_{noise},

www.raguvaran.puzl.com

and R_{gen} can be varied with frequency if desired. I extended the Sherwin method somewhat by using nine octave bands to cover from 50 Hz to 22 kHz, and my latest version adds switchable A-weighting. An advantage of the spreadsheet method is that it is very simple to turn off various noise contributions so you can experiment with noiseless amplifiers or other flights from physical reality. For example, the noise generated by 47 kΩ resistor R_{in} is modeled separately from its loading effects so it can be switched off independently. It is also possible to switch off the bottom three octave bands to make the results comparable with real cartridge measurements that require a steep 400 Hz high-pass filter to control the hum. The results match well with my 5532 and TL072 measurements, and experience shows the model is a usable tool. While it is no substitute for careful measurements, it gives a good physical insight and allows noise comparisons at the LF end, where hum is very difficult to exclude completely.

Table 7.3 shows some interesting cases; output noise is calculated for gain of +29.55 dB at 1 kHz, and signal-to-noise ratio for a 5 mVrms input at 1 kHz. The cartridge parameters were set to 610 Ω + 470 mH, the measured values for the Shure M75ED 2. No weighting is used in this table.

Firstly let us see what the target is; how quiet would the circuit of Figure 7.19 be if we had miraculously noise-free electronics?

TABLE 7.3 RIAA noise results from the MAGNOISE spreadsheet model under differing conditions, in order of merit (Cases 1–3 assume a noiseless amplifier and are purely theoretical)

Case	Amplifier	e_n (nV/$\sqrt{Hz}$)	i_n (pA/$\sqrt{Hz}$)	R_{in} (Ω)	R_0 (Ω)	Noise output (dBu)	S/N ref. 5 mV in (dB)	EIN (dBu)	NF (dB)	Ref. 5534A (dB)
1a	Noiseless amp	0	0	1000M	0	−100.5	−86.2	−130.0	n/a	
1b	Noiseless amp	0	0	10M	0	−100.4	−86.1	−129.9	n/a	
1c	Noiseless amp	0	0	1M	0	−99.6	−85.4	−129.2	n/a	
2	Noiseless amp	0	0	47k	0	−96.1	−81.8	−125.7	0	
3	Noiseless amp	0	0	47k	220	−95.5	−81.2	−125.0	0.7	
4a	2SB737, I_c = 70 μA	1.75	0.386	47k	220	−94.2	−79.9	−123.7	2.0	−1.0
4b	2SB737, I_c = 100 μA	1.47	0.46	47k	220	−94.0	−79.7	−123.5	2.2	−0.8
4c	2SB737, I_c = 200 μA	1.04	0.65	47k	220	−93.2	−78.9	−122.7	3.0	0
5	5534A	3.5	0.4	47k	220	−93.2	−78.9	−122.7	3.0	0
6	5532A	5	0.7	47k	220	−92.5	−78.2	−120.7	5.0	2.0
7	OPA2134	8	0.003	47k	220	−90.2	−76.0	−119.8	5.9	2.9
8	LM4562	2.7	1.6	47k	220	−88.2	−74.0	−117.8	7.9	4.9
9	TL072	18	0.01	47k	220	−86.9	−72.6	−113.9	11.8	8.8

www.raguvaran.puzl.com

Case 1. To begin with we will completely ignore the cartridge loading requirements and set R_{in} to 1000 MΩ, at which value it has no effect. The minimum EIN with these particular cartridge parameters is then −130.0 dBu (Case 1a). This is a very low value, and it is the quietest possible condition. All of this noise comes from R_{gen}, the resistive component of the cartridge impedance. The only way to improve on this would be to select a cartridge with a lower R_{gen} but the same sensitivity, or start pumping liquid nitrogen down the tone-arm. (As an aside, if you *did* cool your cartridge with liquid nitrogen at −196°C, the Johnson noise from R_{gen} would only be reduced by 5.8 dB, and if you are using a 5534A in the preamplifier, as in Case 5 below, the overall improvement would only be 0.7 dB. Also of course, the compliant materials would go solid and the cartridge wouldn't work at all. Hold the cryostats).

With lower, but still high, values of R_{in} the noise increases; with R_{in} set to 10 MΩ (Case 1b) the EIN is −129.9 dBu, a bare 0.1 dB worse. With R_{in} set to 1 MΩ (Case 1c) the EIN is now −129.2 dBu, 0.8 dB worse than the best possible condition.

Case 2. It is, however, a fact of life that MM cartridges need to be properly loaded, and when we set R_{in} to its correct value of 47 kΩ things deteriorate sharply, the EIN rising by 4.3 dB to −125.7 dBu. We are still assuming a noiseless amplifier, and this appears to be the appropriate noise reference for amplifier design, so the NF is 0 dB.

Case 3. We leave the amplifier noise switched off, but add in the Johnson noise from the resistor in the bottom leg of the NFB network (R0 in Figure 7.4, R7 in Figure 7.18) to see if its value of 220 Ω is appropriate. The noise only worsens by 0.7 dB, so it looks like that resistor is not the first thing to worry about. The noise from this resistor is included in all the cases that follow. The NF is now 0.7 dB.

Case 5. We will now take a deep breath and switch on the amplifier noise. We jump to Case 5, where we have a 5534A as the amplifying element, and using the typical 1 kHz specs for the A-suffix part, we get an EIN of −122.7 dBu and an NF of 3.0 dB. This could be regarded as both good and bad. Using thoroughly standard technology, we are within a few decibels of perfection; on the other hand, the opportunities for showing off some virtuoso circuit design appear limited.

Case 6. It is well known that the single 5534 has somewhat better noise specs than the dual 5532, with both e_n and i_n being significantly lower, but does this translate into a significant noise advantage in the RIAA application? Case 6 shows that on plugging in a 5532A the noise output increases by 2.0 dB, the EIN increasing to −120.7 dBu. The NF is now 5.0 dB, which looks a bit less satisfactory. If you want good performance then the inconvenience of a single package and an external compensation capacitor are well worth putting up with – there is no easier way to obtain a 2.0 dB noise improvement than by moving from the 5532A to the 5534A. The rightmost column of Table 7.3 shows how other op-amps compare with the 5534A. If your circuit design ends up with an odd number of half-5532s, a single 5534A can be placed in the MM stage, where its lower noise is best used.

www.raguvaran.puzl.com

Case 7. Ho-hum, I hear you murmur, it all comes back to the 5534, doesn't it? What about all the other op-amps on the market? If for Case 7 we take the FET-input OPA2134, which is a very fine op-amp when DC accuracy and low bias currents are required, we find the e_n is much higher at 8 nV/$\sqrt{\text{Hz}}$, but i_n is much lower at 3 fA/$\sqrt{\text{Hz}}$. It looks like we might be in with a chance, but the greater voltage noise does more harm than the lower current noise does good and the EIN goes up to −119.8 dBu. The OPA2134 is therefore 2.9 dB noisier than the 5534A and 0.9 dB noisier than the 5532A, and it is not cheap. The NF is now 5.9 dB.

Case 8. The LM4562 BJT-input op-amp is the new chip on the block, and a very fine op-amp it is too, giving significant noise improvements over the 5534/5532 when used in low-impedance circuitry. This is because its e_n is lower at 2.7 nV/$\sqrt{\text{Hz}}$. However, the impedances we are dealing with here are not low, and the i_n, at 1.6 pA/$\sqrt{\text{Hz}}$, is four times that of the 5534A, leading us to think it will not do well here. We are sadly correct, with EIN deteriorating to −117.8 dBu and the NF an unimpressive 7.9 dB. The LM4562 is almost 5 dB noisier than the 5534A and at the time of writing is a lot more expensive. I have done some measurements that confirm this, and since I had not at the time done the noise calculations, it came as a rather unwelcome surprise.

Case 9. The TL072 with its FET input has high voltage noise but low current noise. The voltage noise is in fact very high at 18 nV/$\sqrt{\text{Hz}}$, and we can expect a poor performance. We duly get it, with EIN rising to −116.5 dBu and a rather poor NF of 11.8 dB. The TL072 is 8.8 dB noisier than a 5534A, and 6.8 dB noisier than a 5532A. The latter figure is confirmed (within experimental error anyway) by the data listed in the section below on noise measurements.

You may be wondering what has happened to other well-known op-amps, particularly the LT1028 and the OP27. Both are sometimes recommended for audio use, because of their low e_n values, but this ignores some serious problems.

The LT1028 gives a poor performance in MM applications because while it has an appealing low e_n of 0.85 nV/$\sqrt{\text{Hz}}$, its i_n is high at 1 pA/$\sqrt{\text{Hz}}$, as a result of running big input BJTs at high currents. The EIN is by calculation −120.9 dBu, making it a shade quieter than the 5532A. However, read to the end of this section before ordering any.

The OP27 has a low e_n and i_n, and in fact gives a calculated EIN of −123.0 dB, which beats the 5532A noise, but when you measure it in real life it is actually several decibels noisier; I have confirmed this several times. This is due to extra noise generated by bias-current cancellation circuitry. Correlated noise currents are fed into both inputs, and will only cancel if both inputs see the same impedance. In this RIAA application the impedances are wildly different, and the result is greatly increased noise. This problem with the OP27 was originally pointed out to me by Marcel van de Gevel [19].

The LT1028 also has bias-current cancellation circuitry, and the data sheet explicitly states: 'The cancellation circuitry injects two correlated current noise components into the two inputs.' In real life it will be several decibels noisier than the 5532A.

www.raguvaran.puzl.com

We will now go back to Case 4 (I have done it this way to keep the rows in Table 7.3 in order of noise performance), which is that of a single discrete bipolar transistor used as an input device, either as part of a fully discrete RIAA stage or as the front end to an op-amp. If we turn a blind eye to supply difficulties and use the remarkable 2SB737 transistor (with R_b only 2 Ω typical) then some interesting results are possible. We now can decide what collector current to run the device at, so we can to some extent trade off voltage noise against current noise. We know that current noise is important with an MM input, and so we will start with $I_c = 200$ μA, which gives Case 4c in Table 7.3. By coincidence, this gives exactly the same noise results as for the 5534A. Undiscouraged, we drop I_c to 100 μA (Case 4b) and voltage noise increases but current noise decreases, the net result being that things are now 0.8 dB quieter. If we reduce I_c again to 70 μA (Case 4a) we gain another 0.2 dB, and we have an EIN of −123.7 and an NF of only 2.0 dB. Voltage noise is now increasing fast and there is virtually nothing to be gained by reducing the collector current further.

We therefore must conclude that even an exceptionally good discrete transistor with appropriate support circuitry will only gain us a 1.0 dB noise advantage over the 5534A, and it is questionable if the extra complications are worth it.

Noise Results with A-Weighting

Many commentators feel that it is appropriate to use psychoacoustic weighting when studying RIAA preamplifier noise, because the frequency-dependent nature of the noise behavior interacts with the non-flat human hearing response at normal listening levels. A-weighting is universally used. You may therefore be wondering how the unweighted results described above will be affected by the application of A-weighting. Will the order of merit in Table 7.3 be upset?

Table 7.4 compares the results of no weighting and A-weighting on the results in Table 7.3. The apparent performance improves because the low-frequency noise, which has been emphasized by the RIAA equalization, is now discriminated against by the LF roll-off of the A-weighting curve. Note that the reference case for determining the NF is also subjected to A-weighting.

It is clear that A-weighting does not introduce any revolutionary changes into the order of merit of the various devices. The 2SB737 at 200 μA now comes out as 0.3 dB noisier than the 5534A, but this is the only alteration.

RIAA Noise Measurements

In the past, many people who should have known better recommended that MM input noise should be measured with a 1 kΩ load, presumably thinking that this emulates the resistance R_{gen}, which is the only parameter in the cartridge actually generating noise – the inductance is

www.raguvaran.puzl.com

TABLE 7.4 The effect of A-weighting on the calculated noise performances of various amplifiers

	Unweighted EIN (dBu)	Unweighted NF (dB)	A-weighted EIN (dBu)	A-weighted NF (dB)
2SB737 70 μA	−123.7	2.0	−126.0	1.9
2SB737 100 μA	−123.5	2.2	−125.7	2.2
2SB737 200 μA	−122.7	3.0	−124.8	3.1
5534A	−122.7	3.0	−125.1	2.8
5532A	−120.7	5.0	−123.0	4.9
OPA2134	−119.8	5.9	−122.5	5.4
LM4562	−117.8	7.9	−119.7	8.2
TL072	−113.9	11.8	−116.7	11.2

of course noiseless. This overlooks the massive effect that the inductance has in making the impedance seen at the preamp input rise very strongly with frequency, so that at higher frequencies most of the input noise actually comes from the 47 kΩ loading resistance. I am grateful to Marcel van de Gevel for drawing my attention to some of the deeper implications of this point [19].

The importance of using a real cartridge load is demonstrated in Table 7.5, where the noise performances of a TL072 and a 5532 are compared. The TL072 result is 0.8 dB too low, and the 5532 result 4.9 dB too low – a hefty error. In general, results with the 1 kΩ resistor will always be too low, by a variable amount. In this case you still get the right overall answer – i.e. you should use a 5532 for least noise – but the decibel difference between the two has been exaggerated by almost a factor of 2.

These tests were done with an amplifier gain of +29.55 dB at 1 kHz. Bandwidth was 400 Hz–22 kHz to remove hum, rms sensing, no weighting, cartridge parameters were 610 Ω + 470 mH.

The 1 kΩ recommendation was perhaps made because the obvious measurement method of loading the input with an MM cartridge has serious difficulties with hum from the ambient magnetic fields. To get useful results it is essential to enclose the cartridge completely in a mu-metal can – I use one from a redundant microphone transformer and it works very well. I suppose the ideal load would be a toroidal inductor, but it would be an expensive custom part. It is desirable to use complete electrostatic screening of the amplifier itself. If it has a 22 μF input coupling capacitor and the input is short-circuited, the impedance downstream of the capacitor is 145 Ω at 50 Hz, which is enough to make it susceptible to electrostatic hum pickup.

Electronic Cartridge Loading for Lower Noise

Going back to Table 7.3, you will recall that when we were examining the situation with the amplifier and feedback network noise switched off, adding in the Johnson noise from the

www.raguvaran.puzl.com

TABLE 7.5 Measured noise performance of 5532 and TL072 with two different source impedances

Z_{source}	TL072	5532	5532 benefit	5532 EIN
1 kΩ res.	−88.0	−97.2 dBu	+9.8 dB	−126.7 dBu
Shure M75ED 2	−87.2	−92.3 dBu	+5.1 dB	−121.8 dBu

47 kΩ loading resistor R_{in} caused the output noise to rise by 4.3 dB. In real conditions with amplifier noise included the effect is obviously less dramatic, but it is still significant. In the 5534A case the removal of the noise from R_{in} (but *not* the loading effect of R_{in}) reduces the noise output by 1.7 dB. Table 7.6 summarizes the results for various amplifier options in columns 4–6; the amplifier noise is unaffected, so the noisier the technology used, the less the improvement. Note that the no-R_{in} NFs in the fifth column look worse because the reference case also has the noise from R_{in} removed, and is now −130.0 dBu. All noise results in this section are unweighted.

This may appear to be utterly academic, because the cartridge must be loaded with 47 kΩ to get the correct response. This is true, *but it does not have to be loaded with a physical 47 kΩ resistor.* An electronic circuit that has the *V/I* characteristics of a 47 kΩ resistor but lower noise, will do the job very well. Such a circuit may seem like a tall order – it will after all be connected at the very input where noise is critical, but unusually the task is not as difficult as it seems.

Figure 7.20 (a) shows the basic principle. The 47 kΩ R_{in} is replaced with a 1 MΩ resistor whose bottom end is driven with a voltage that is phase-inverted and 20.27 times that at the top

TABLE 7.6 The noise advantages gained by removing or reducing the noise from the 47 kΩ loading resistor R_{in} (from MAGNOISE)

1	2	3	4	5	6	7	8	9
	R_{in} = 47 kΩ EIN (dBu)	R_{in} = 47 kΩ NF (dB)	R_{in} = inf. EIN (dBu)	R_{in} = inf. NF (dB)	Load synth. advantage R_{in} = inf. (dB)	R_{in} = 1M EIN (dBu)	R_{in} = 1M NF (dB)	Load synth. advantage R_{in} = 1M (dB)
2SB737 70 µA	−123.7	2.0	−126.0	4.0	**2.3**	−125.7	3.5	**2.0**
2SB737 100 µA	−123.5	2.2	−125.7	4.3	**2.2**	−125.4	3.8	**1.9**
2SB737 200 µA	−122.7	3.0	−124.4	5.6	**1.7**	−124.2	5.0	**1.5**
5534A	−122.7	3.0	−124.4	5.6	**1.7**	−124.2	5.0	**1.5**
5532A	−120.7	5.0	−121.7	8.3	**1.0**	−121.6	7.6	**0.9**
OPA2134	−119.8	5.9	−120.5	9.5	**0.7**	−120.4	8.8	**0.6**
LM4562	−117.8	7.9	−118.3	11.7	**0.5**	−118.2	11.0	**0.4**
TL072	−113.9	11.8	−114.1	15.9	**0.2**	−114.1	15.1	**0.2**

www.raguvaran.puzl.com

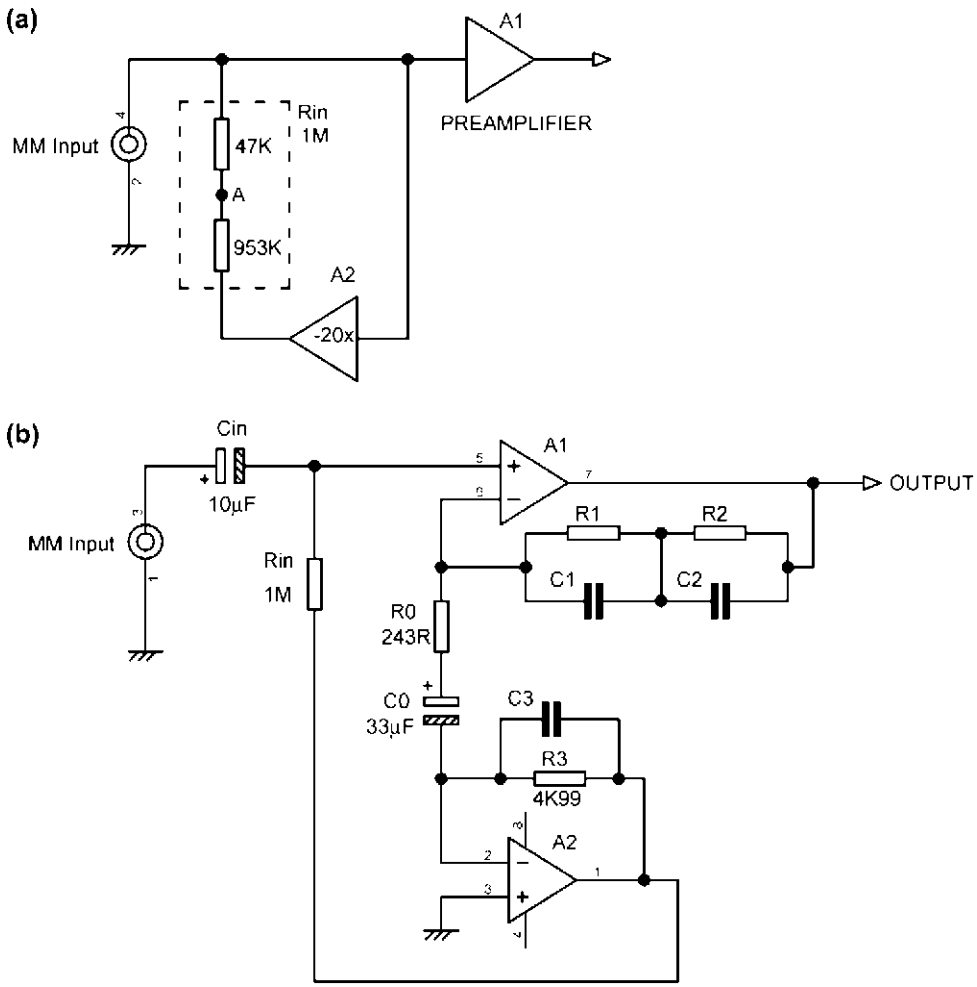


Figure 7.20: Electronic load synthesis. (a) The basic principle. (b) The van de Gevel circuit.

end. If we conceptually split the 1 MΩ resistor into two parts of 47 and 953 kΩ, a little light mathematics shows that with −20.27 times V_{in} at the output of A2, the voltage at the 47 kΩ–953 kΩ junction A is zero, and so far as the cartridge is concerned it is looking at a 47 kΩ resistance to ground. However, the physical component is 1 MΩ, and the Johnson current noise it produces is less than that from a 47 kΩ (Johnson current noise is just the usual Johnson voltage noise applied through the resistance in question). The point here is that the physical resistor value has increased by 21.27 times, but the Johnson noise has only increased by 4.61 times, because of the square root in the Johnson equation; thus, the current noise injected by R_{in} is also reduced by 4.61 times. The improvement gained using a practical 1 MΩ resistor rather than a less practical infinite one is a bit less, as shown in columns 7–9, but with a 5534A amplifier the difference is only 0.2 dB.

www.raguvaran.puzl.com

Other combinations of resistor value and amplifier gain could be used – the higher the resistance, the more the noise advantage, as described earlier in this chapter. The problem is that this also requires higher inverted drive voltage at the bottom of the resistor, and you will soon run into a situation where the inverting amplifier clips before the main path amplifier, restricting headroom.

The implementation made known by van de Gevel [16,19] is shown in Figure 7.20 (b). This ingenious circuit uses the current flowing through the feedback resistor R0 to drive a shunt-feedback stage around A2. With suitable scaling of R3 (note that here it has an E96 value), the output voltage of A2 is at the right level and correctly phase-inverted. When I first saw this circuit I had reservations about connecting R0 to a virtual ground rather than a real one, and thought that extra noise from A2 might find its way back up R0 into the main path. (I hasten to add that these fears may be quite unjustified, and I have not found time so far to put them to a practical test.) The inverting signal given by this circuit is amplified by 20.5 times rather than 20.27, but this has a negligible effect on the amount of noise reduction.

Because of these reservations, I tried out my version of load synthesis as shown in Figure 7.21. This more closely resembles the circuit of Figure 7.20 (a); it is important that the inverting stage A3 does not load the input with its 1 kΩ input resistor R4, so a unity-gain buffer A2 is added. The inverting signal is amplified by 20 times, not 20.27, but once again this has negligible effect on the noise reduction.

In practical measurements with a 5534A as amplifier A1, I found that the noise improvement with the real cartridge load (Shure M75ED 2, cartridge parameters 610 Ω + 470 mH) was

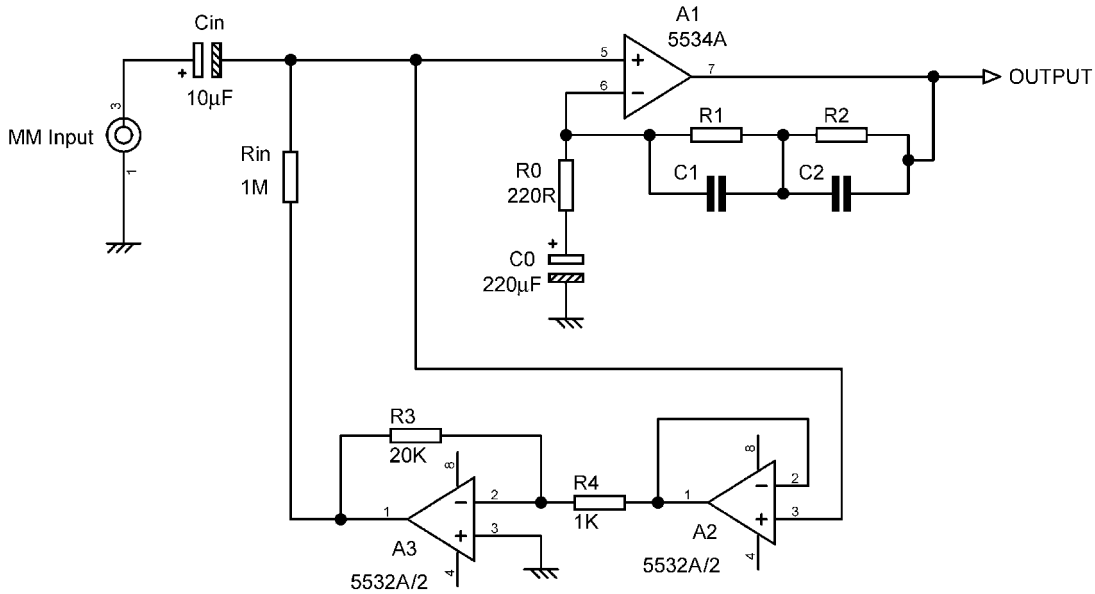


Figure 7.21: Electronic load synthesis: the Self circuit

www.raguvaran.puzl.com

indeed 1.5 dB, just as predicted, which is as nice a matching of theory and reality as you are likely to encounter in this world. There were no HF stability problems. Whether the 1.5 dB is worth the extra electronics is a good question; I say it's worth having.

This technique has been called 'electronic cooling', presumably because it could be regarded as analogous to dipping the loading resistance in liquid nitrogen or whatever to reduce Johnson noise. I must admit I don't like the term as it could be understood to mean that thermoelectric elements have been used to cool down the input stage, a technique I do not think has been used in hi-fi yet. I prefer to call it 'electronic loading', 'active input impedance' or 'load synthesis', the latter being perhaps the most explicit.

Subsonic Filters

In the earlier parts of this chapter we have seen that the worst subsonic disturbances occur in the 2–4 Hz region, due to disc warps, and are about 8 dB less at 10 Hz. We have also seen that the IEC Amendment gives only 14 dB of attenuation at 4 Hz, and in any case is often omitted by the manufacturer or switched out by the user. It is therefore important to provide authoritative subsonic filtering. What needs to be settled is what order filter to use, because some people at least will be concerned about the audibility of LF phase-shifts, and how far into the audio band the filter should intrude. There is nothing approaching a consensus on either point, so it can be a wise move to configure the subsonic filter so it can be switched out.

The third-order filter already described in this chapter also did the job of implementing the IEC Amendment, so it did not have one of the classic filter characteristics. All the filters described here do just the filtering job and it is assumed that the IEC Amendment is implemented elsewhere, if at all.

High-pass filters used for RIAA subsonic are typically of the second-order or third-order Butterworth (maximally flat) configuration, rolling off at rates of 12 and 18 dB/octave respectively, as shown in Figure 7.22. Fourth-order 24 dB/octave filters are much less common, presumably due to worries about the possible audibility of rapid phase changes at the very bottom of the audio spectrum. The Butterworth response is but one of many possible filter alignments; the Bessel response gives a slower roll-off, but aims for linear phase, i.e. a constant delay versus frequency, and so reproduces the shape of transients better. Other filter alignments such as Chebyshev give faster initial roll-offs than the Butterworth, but they do so at the expense of ripples in the passband or stopband gain, which is not helpful if you are aiming for a ruler-flat response after RIAA equalization.

A very handy filter configuration is the well-known Sallen-and-Key type; it has drawbacks when used as a low-pass handling high frequencies (the response comes back up due to the non-zero op-amp output impedance) but works very well for our purposes here. A second-order Sallen-and-Key filter is simple to design; the two series capacitors C1 and C2 are made equal and R2 is made twice the value of R1. Such a filter with a −3 dB point at 20 Hz is shown

www.raguvaran.puzl.com

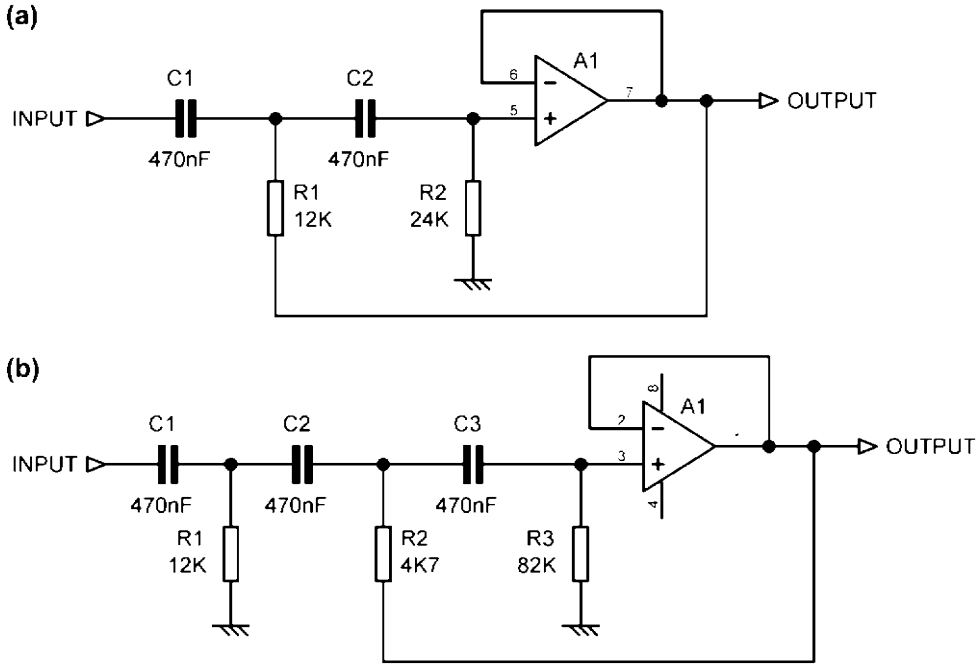


Figure 7.22: Subsonic filters: second-order (a) and third-order (b) Butterworth high-pass filters, both 3 dB down at 20 Hz

in Figure 7.21(a). Other roll-off frequencies can be obtained simply by scaling the component values while keeping C1 equal to C2 and R2 twice R1. The response is 24.0 dB down at 5 Hz, by which time the 12 dB/octave slope is well established, and we are well protected against disc warps. It is, however, only 12.3 dB down at 10 Hz, which gives little protection against arm-resonance problems. Above the −3 dB roll-off point the response is still −0.78 dB down at 30 Hz, which is intruding a little into the sort of frequencies we want to keep. We have to conclude that a second-order filter really does not bifurcate the condiment, and the faster roll-off of a third-order filter is preferable.

Third-order filters are a little more complex. Some versions are made up of a second-order filter cascaded with a first-order roll-off, using two op-amp sections. It can, however, be done with just one, as in Figure 21.1b, which is a third-order Butterworth filter also with a −3 dB point at 20 Hz. The resistor value ratios are now a less friendly 2.53:1.00:17.55, and the circuit shown uses the nearest E24 values to this – which by happy chance come out as E12 values. The frequency response is shown in Figure 7.23, where it can be seen to be 18.6 dB down at 10 Hz, which should keep out any arm-resonance frequencies. It is 36.0 dB down at 5 Hz so disc warp spurii won't have a chance. The 30 Hz response is now only down by an insignificant −0.37 dB, which demonstrates that a third-order filter is much better than a second-order filter for this application. As before, other roll-off frequencies can be had by scaling the component values while keeping the resistor ratios the same.

www.raguvaran.puzl.com

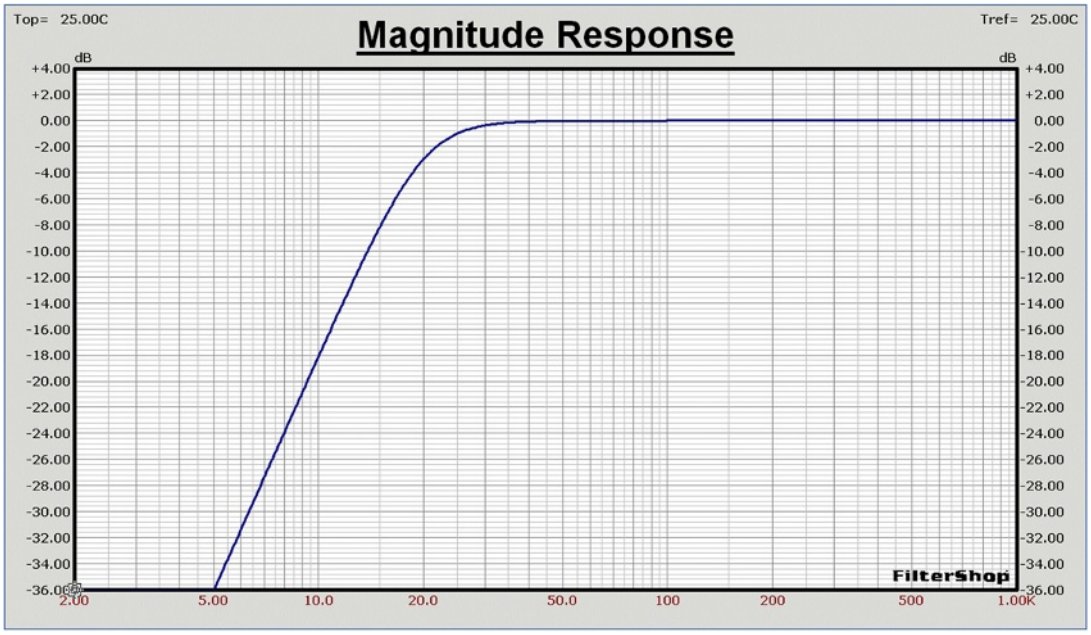


Figure 7.23: Frequency response of a third-order Butterworth subsonic filter, 3 dB down at 20 Hz

When dealing with frequency-dependent networks like filters you need to keep an eye on the input impedance, because it can drop to unexpectedly low values, putting excessive loading on the stage upstream and degrading its linearity. In a high-pass Sallen-and-Key filter, the input impedance is high at low frequencies but falls with increasing frequency. In the third-order version, it tends to the value of R1 in parallel with R3, which here is 10.6 kΩ. This should not worry the previous stage.

Because of the large capacitances, the noise generated by the passive elements in a high-pass filter of this sort is usually well below the op-amp noise. The capacitances do not, of course, generate any noise themselves. With the values used here, SPICE simulation shows that the resistors produce −125.0 dBu of noise at the output (22 kHz bandwidth, 25°C).

Capacitor distortion in electrolytics is (or should be) by now a well-known phenomenon. It is perhaps less well known that non-electrolytics can also generate distortion in filters like these. This has nothing to do with subjectivist musicality, but is all too real and measurable. Details of the problem are given in Chapter 2, where it is concluded that only NP0 ceramic, polystyrene, and polypropylene capacitors can be regarded as free of this effect. The capacitor sizes needed for subsonic filters are large, if impedances and hence noise are to be kept low, which means polypropylene has to be used. Anything larger than 470 nF gets to be big and expensive, so that is the value used here; 220 nF polypropylene is substantially smaller and about half the price. There is more information on this, and on highpass filters in general, in Chapter 5 on filters.

www.raguvaran.puzl.com

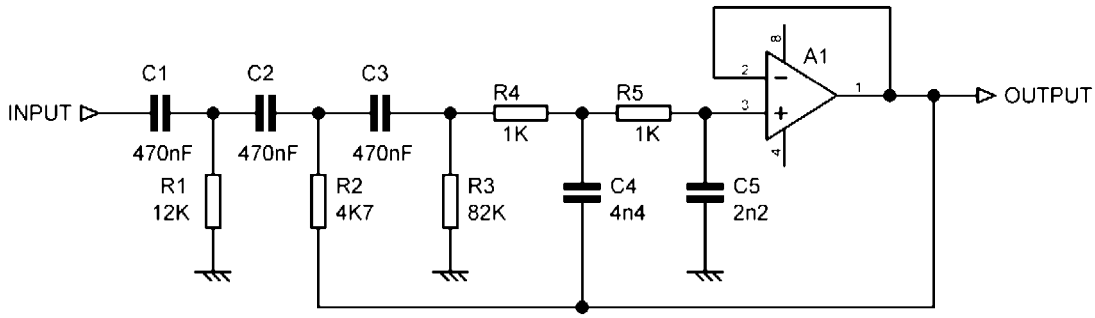


Figure 7.24: A third-order Butterworth subsonic filter combined with a second-order ultrasonic filter

Combining Subsonic and Ultrasonic Filters

Scratches and groove debris create clicks that have a large high-frequency content, some of it ultrasonic and liable to cause slew-rate and intermodulation problems further down the audio chain. It is often considered desirable to filter this out as soon as possible (though of course some people are only satisfied with radio-transmitter frequency responses), but an obstacle to this is the extra cost and power consumption of another filter stage. This difficulty can be resolved by combining an ultrasonic filter with a subsonic filter in the same stage. Combined filters also have the advantage that the signal now passes through one op-amp rather than two, and can be extremely useful if you only have one op-amp section left.

This cunning plan is workable only because the high-pass and low-pass turnover frequencies are widely different. Figure 7.24 shows the third-order Butterworth subsonic filter combined with a second-order 50 kHz Butterworth low-pass filter; the response of the combination is exactly the same as expected for each separately. The low-pass filter is cautiously designed to prevent significant loss in the audio band, and has a −3 dB point at 50 kHz, giving very close to 0.0 dB at 20 kHz. The response is −12.6 dB down at 100 kHz and −24.9 dB at 200 kHz. C4 is made up of two 2n2 capacitors in parallel.

Note that the mid-band gain of the combined filter is −0.15 dB rather than exactly unity. The loss occurs because the series combination of C1, C2, and C3, together with C5, form a capacitative potential divider with this attenuation, and this is one reason why the turnover frequencies need to be widely separated for filter combining to work. If they were closer together then C1, C2, C3 would be smaller, C5 would be bigger, and the capacitative divider loss would be greater.

References

[1] M. Jones, Designing valve preamps: Part 1, Electronics World (March 1996) p. 192.
[2] T. Holman, Dynamic range requirements of phonographic preamplifiers, Audio (July 1977) p. 74.

www.raguvaran.puzl.com

[3] K. Howard, Cut and thrust: RIAA LP equalisation, Stereophile (March 2009).
[4] H.P. Walker, Low-noise audio amplifiers, Electronics World (May 1972) p. 233.
[5] S.P. Lipshitz, On RIAA equalisation networks, J. Audio Eng. Soc. (June 1979) p. 458. onwards.
[6] W.J. Smith, Modern Optical Engineering, McGraw-Hill, 1990, p. 484.
[7] T. Holman, New factors in phonograph preamplifier design, J. Audio Eng. Soc. (May 1975) p. 263.
[8] D. Self, An advanced preamplifier design, Wireless World (November 1976).
[9] D. Self, High performance preamplifier, Wireless World (February 1979).
[10] D. Self, A precision preamplifier, Wireless World (October 1983).
[11] D. Self, Precision preamplifier 96, Electronics World (July/August and September 1996).
[12] B. Vogel, Adventure: noise (calculating RIAA noise), Electronics World (May 2005) p. 28.
[13] J. Dinsdale, Transistor high-quality audio amplifier, Wireless World (January 1965) p. 2.
[14] A.R. Bailey, High performance transistor amplifier, Wireless World (December 1966) p. 598.
[15] http://www.signaltransfer.freeuk.com/
[16] M. van de Gevel, Noise and moving-magnet cartridges, Electronics World (October 2003) p. 38.
[17] B. Hallgren, On the noise performance of a magnetic phonograph pickup, J. Audio Eng. Soc. (September 1975) p. 546.
[18] J. Sherwin, Noise specs confusing? National Semiconductor Application Note AN-104, Linear Applications Handbook (1991).
[19] M. van de Gevel, Private communication (February 1996).

www.raguvaran.puzl.com

This page intentionally left blank

www.raguvaran.puzl.com

CHAPTER 8

Moving-Coil Head Amplifiers

Moving-coil cartridges are generally accepted to have a better tracking performance than moving-magnet cartridges because the moving element is a set of lightweight coils rather than a magnet, which is inevitably made of a relatively dense alloy. Because the coils must be light, they consist of relatively few turns and the output voltage is very low, typically in the range 100–550 μVrms at a velocity of 5 cm/s, compared with 5 mVrms from the average moving-magnet cartridge. Fortunately this low output comes from a very low impedance, which, by various technical means, allows an acceptable signal-to-noise performance to be obtained.

Moving-Coil Cartridge Characteristics

There is much greater variation in impedance and output across the available range of moving-coil (MC) cartridges than for moving-magnet (MM) cartridges. The range quoted above is already much wider, but including the extremes currently on the market (2009) the output range is 40–2500 μV, a remarkably wide span of 62 times, or 36 dB. This is illustrated in Figure 8.1, which shows the results of a survey of 85 different MC cartridges (note that the ranges for the columns are wider on the right side of the diagram). When I first became involved in designing MC amplifiers in 1986, I compiled a similar chart [1] and it is interesting to note that the same features occurred – there are two separate clusters around 100–300 and 500–700 μV, and the lowest output is still 40 μV from an Audio Note cartridge (the loLtd model). The highest output of 2.5 mV comes from the Benz Micro H2, and this is only 6 dB below an MM cartridge.

Assuming that a conventional MM input stage is being used to raise 5 mV to nominal internal level and perform the RIAA equalization, the Audio Note cartridge requires a gain of 125 times or +42 dB. The cartridge cluster around 200 μV output needs 25 times or +28 dB, while the 500 μV cluster requires 10 times or +20 dB. If an amplifier is to cover the whole range of MC cartridges available, some form of gain switching is highly desirable.

Cartridge impedances also vary markedly, over a range from 1 Ω (Audio Note loLtd) to 160 Ω (Denon DL-110 and DL160) with impedance increasing with output level, as you would expect – there are more turns of wire in the coils. The inductance of MC cartridges is very low and their source impedance is normally treated as purely resistive. The recommended load

Small Signal Audio Design; ISBN: 9780240521770
Copyright © 2010 Elsevier Ltd; All rights of reproduction, in any form, reserved.

www.raguvaran.puzl.com

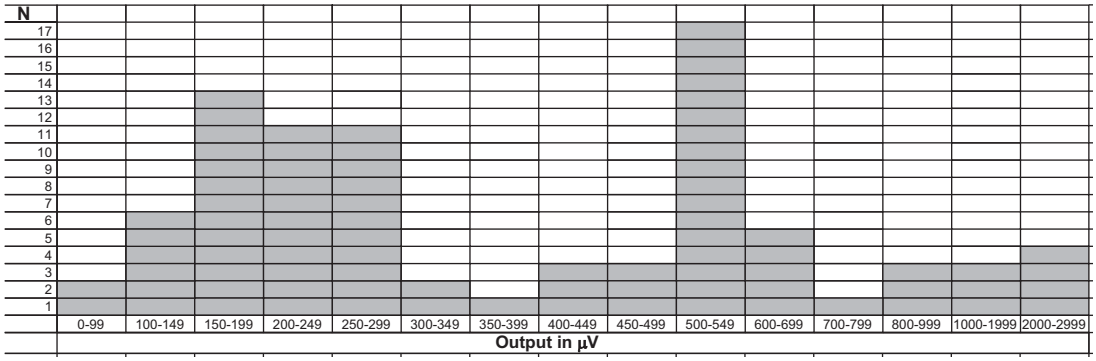


Figure 8.1: The output levels for 85 moving-coil cartridges at 5 cm/s (2009)

impedances are also resistive (unlike the R–C combinations often used with MM cartridges) and are usually quoted as a minimum load resistance. Once more the variation is wide, from 3 Ω (Audio Note loLtd again) to 47 kΩ (Denon DL-110 and DL160 again), but a 100 Ω input load will be high enough for most of the cartridges surveyed and 500 Ω will work for almost all of them. The Audio Note loLtd cartridge is unusual in another way – its magnetic field is produced not by permanent magnets but a DC-powered electromagnet, which presumably requires a very pure supply indeed. The manufacturers whose cartridges were included in the survey are listed in Table 8.1.

Amplification Strategies

There are two ways to achieve the high gains required for these low-output cartridges. In most common methods a standard MM input stage, with RIAA equalization, can be switched to accept either an MM input directly, or the output of a specialized MC input stage that gives the extra gain needed; this may be either a step-up transformer or an amplifier configured to work well with very low source resistances. The large amount of gain required is split between two

TABLE 8.1 Moving-coil cartridge manufacturers whose product data was used to compile Figure 8.1

Audio Note	Immutable Music
Benz Micro	Koetsu
Cardas	Lyra
Clear Audio	Miyabi
Denon	Ortofon
Dynavector	Shelter
Goldring	Sumiko
Grado	van den Hul

www.raguvaran.puzl.com

stages, which makes it easier to achieve. Alternatively, a single stage can be used with switched gain, but this is not too hot an idea:

1. Switchable gain makes accurate RIAA equalization much harder.
2. For good noise performance, the input device operating current needs to be low for MM use (where it sees a high impedance) and high for MC use (where it sees a very low impedance). Making this operating current switchable would be a complicated business.
3. Achieving the very high gain required for MC operation together with low distortion and adequate bandwidth will be a challenge. It is unlikely to be possible with a single op-amp, and so there is little likelihood of any saving on parts.

Moving-Coil Transformers

If you have a very low output voltage and very low impedance, an obvious way to deal with this is by using a step-up transformer to raise the voltage to the level where it can be appropriately applied to a moving-magnet amplifier stage, such as those discussed in Chapter 7. Such a transformer has most of the usual disadvantages, such as frequency response problems and cost, though for hi-fi use the weight is not a difficulty, and non-linearity should not be an issue because of the very low signal levels.

In this application the cost is increased by the need for very high immunity to hum fields. While it is relatively straightforward to make transformers that have high immunity to external magnetic fields, particularly if they are toroidal in construction, it is not cheap, because the mu-metal cans that are required for the sort of immunity necessary are difficult to manufacture. The root of the problem is that the signal being handled is so very small. The transformer is usually working in a modern house, which can have surprisingly large hum fields, and generally presents a hostile environment for very-low-signal transformers, so very good immunity indeed is required; some manufacturers use two nested screening cans, separately grounded, to achieve this, as shown in Figure 8.2. An inter-winding electrostatic screen is usually fitted. A stereo MC input naturally requires *two* of these costly transformers.

MC transformers are designed with low primary winding resistances, typically 2 or 3 Ω, to minimize the Johnson noise contribution from the transformer. Some transformers have windings made of silver rather than copper wire, but the conductivity of silver is only 5% lower than that of copper, and the increase in cost is startling. At the time of writing one brand of silver-wound transformer costs more than £1200 – each, not for a pair.

Because of the great variation in cartridge output levels and impedances, some manufacturers (e.g. Jensen) offer transformers with two or three primary windings, which can be connected in series, parallel, or series-parallel to accommodate a wide variety of cartridges.

www.raguvaran.puzl.com

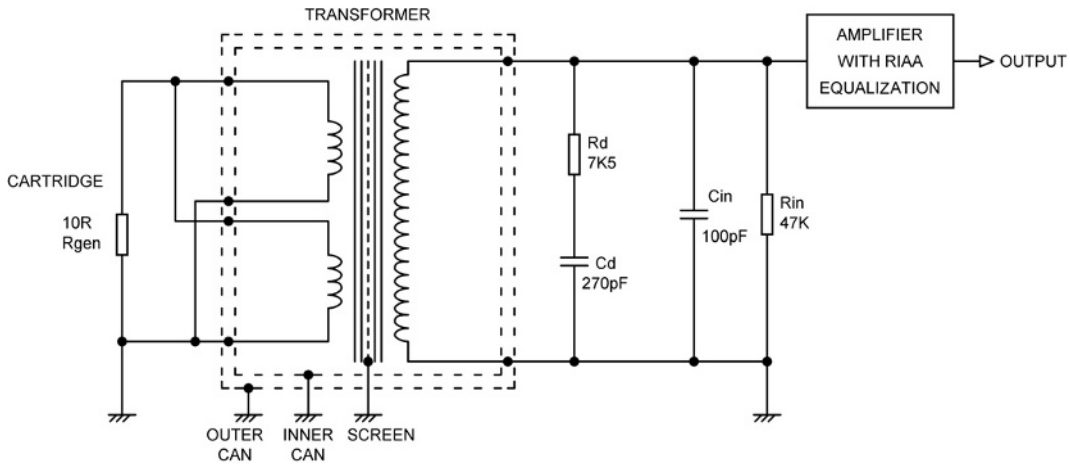


Figure 8.2: A typical MC step-up transformer circuit with twin primaries wired in parallel, dual screening cans, and Zobel network Rd, Cd across the secondary. R_{in}, C_{in} are the usual MM input loading components

A transformer secondary must be correctly loaded to give the flattest possible frequency response, and this usually means that a Zobel R–C network must be connected across it, as explained in Chapter 14 on line inputs. This is Rd and Cd in Figure 8.2, where they have typical values. The values required depend not only on the transformer design but also somewhat on the cartridge impedance, and some manufacturers such as Jensen are praiseworthily thorough in giving secondary loading recommendations for a wide range of cartridge impedances.

The very wide variation in cartridge outputs means that the step-up ratio of the transformer must be matched to the cartridge to get an output around 5 mV that is suitable for an MM input amplifier. For example, Jensen offer basic step-up ratios from 1:8 to 1:37. The maximum ratio is limited not only by transformer design issues but by the fact that the loading on the secondary is, as with all transformers, transferred to the primary divided by the square of the turns ratio. A 1:37 transformer connected to the 47 kΩ input impedance of an MM stage will have an impedance looking into the primary of only 34 Ω; such a transformer would, however, only be used with a very-low-impedance, low-output cartridge, which would be quite happy with such a loading. It is of course possible to arrange the input switching so the 47 kΩ input load is on the MM socket side of the MC/MM switch; the MM amplifier can then have a substantially higher input impedance.

Moving-Coil Input Amplifiers

The high cost of transformers means that there is a strong incentive to come up with an electronic solution to the amplification problem. The only thing that makes it possible to

www.raguvaran.puzl.com

achieve a reasonable signal-to-noise ratio is that the very small signal comes from a very low source impedance.

MC head amplifiers come in many forms, but almost all in use today can be classified into one of the topologies shown in Figure 8.3, all of which use series feedback. The configuration in Figure 8.3(a) is a complementary-feedback pair using a single input transistor chosen to have a low base series resistance R_b. The feedback network must also have a low impedance to prevent its Johnson noise from dominating the overall noise output, and this puts a heavy load on the second transistor. Typically a gain of 47 times will be implemented with an upper feedback resistor of 100 Ω and a lower resistor of 2 Ω, a total load on the amplifier output of 102 Ω. The combination of limited open-loop gain and the heavy load of the feedback network means that both linearity and maximum output level tend to be uninspiring, and the distortion performance is only acceptable because the signals are so small. An amplifier of this type is analyzed in Ref. [2].

Figure 8.3(b) shows a classic configuration where multiple transistors are operated in parallel so that their gains add but their uncorrelated noise partly cancels. Two transistors give a 3 dB improvement, four transistors 6 dB, and so on. The gain block A is traditionally one or two discrete devices, that again have difficulty in driving the low-impedance feedback network. Attention is usually paid to ensuring proper current sharing between the input devices. This can be done by adding low-value emitter resistors to swamp V_{be} variations; they are effectively in series with the input path, and therefore degrade the noise performance unless each resistor is individually decoupled with a large electrolytic. Alternatively, each transistor can be given its own DC feedback loop to set up its collector current. For examples of this kind of circuitry, see Ref. [3].

Figure 8.3(c) shows the series-pair configuration. This simple arrangement uses two complementary input transistors to achieve a 3 dB noise improvement without current-sharing problems as essentially the same collector current goes through each device. The collector signal currents are summed in R_c, which must be reasonably low in value to absorb collector-current imbalances. There is no feedback so linearity is poor. The biasing arrangements are not shown.

Figure 8.3(d) is an enhancement of Figure 8.3(a), with the input transistor inverted in polarity and the spadework of providing open-loop gain and output drive capability entrusted to an op-amp. The much increased feedback gives excellent linearity, and less than 0.002% THD at full output may be confidently expected. However, problems remain. R_{F2} must be very low in value, as it is effectively in series with the input and will degrade the noise performance accordingly. If R_{F2} is 10 Ω (which is on the high side) C_F must be very large, for example 1000 μF, to limit the LF roll-off to −1 dB at 30 Hz. Adopting a quieter 3.3 Ω for the R_{F2} position gives significantly lower noise but demands 4700 μF to give −3 dB at 10 Hz; this is not elegant and leads to doubts as to whether the ESR of

www.raguvaran.puzl.com

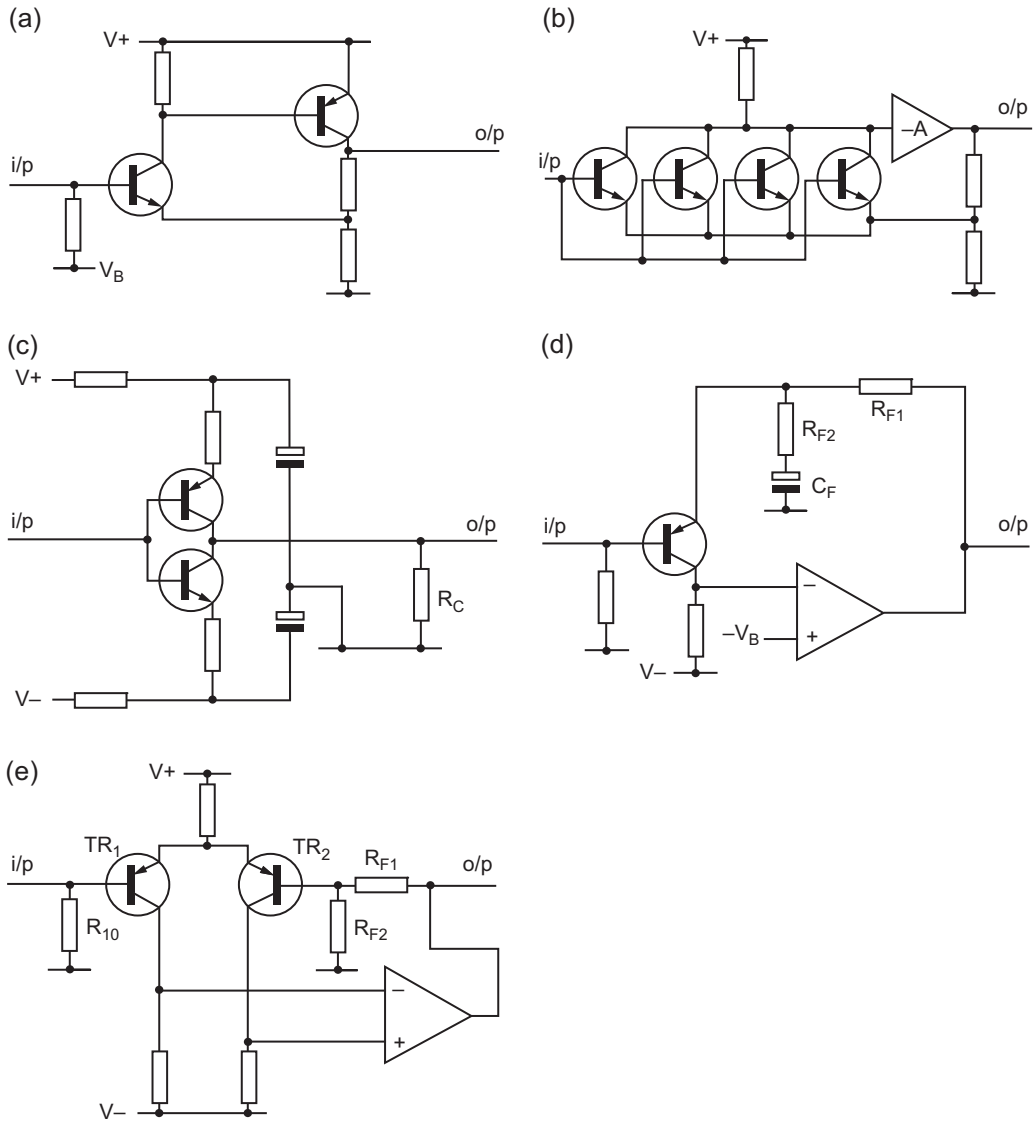


Figure 8.3: Popular MC amplifier configurations

the capacitor will cause trouble. C_F is essential to reduce the gain to unity at DC because there is +0.6 V on the input device emitter, and we don't want to amplify that by 50 times.

The +0.6 V offset can be eliminated by the use of a differential pair, as in Figure 8.3(e). This cancels out the V_{be} of the input transistor TR_1, at the cost of some degradation in noise performance. The pious hope is that the DC offset is so much smaller that if C_F is omitted, and the offset is amplified by the full AC gain, the output voltage swing will not be seriously

www.raguvaran.puzl.com

reduced. The noise degradation incurred by using a differential pair was measured at about 2.8 dB. Another objection to this circuit is that the offset at the output is still non-negligible, about 1 V, mostly due to the base bias current flowing through R10. A DC-blocking capacitor on the output is essential.

A New Moving-Coil Amplifier Configuration

Finding none of these configurations satisfactory, I evolved the configuration shown as a block diagram in Figure 8.4. There is no C_F in the feedback loop, and indeed no overall DC feedback at all. The input transistor and the op-amp each have their own DC feedback systems. The transistor relies on simple shunt negative feedback via DC loop 1; the op-amp has its output held precisely to a DC level of 0 V by the integrator A2, which acts as DC loop 2. This senses the mean output level, and sets up a voltage on the non-inverting input of A1 that is very close to that at Q1 collector, such that the output stays firmly at zero. The time-constant is made large enough to ensure that an ample amount of open-loop gain exists at the lowest audio frequencies. Too short a time-constant will give a rapid rise in distortion as frequency falls. Any changes in the direct voltage on the Q1 collector are completely uncoupled from the output, but AC feedback passes through R_{F1} as usual and ensures that the overall linearity is near perfect, as is often the case with transistor op-amp hybrid circuits. Due to the high open-loop gain of A the AC signal on the Q1 collector is very small and so shunt AC feedback through DC loop 1 does not significantly reduce the input impedance of the overall amplifier, which is about 8 kΩ.

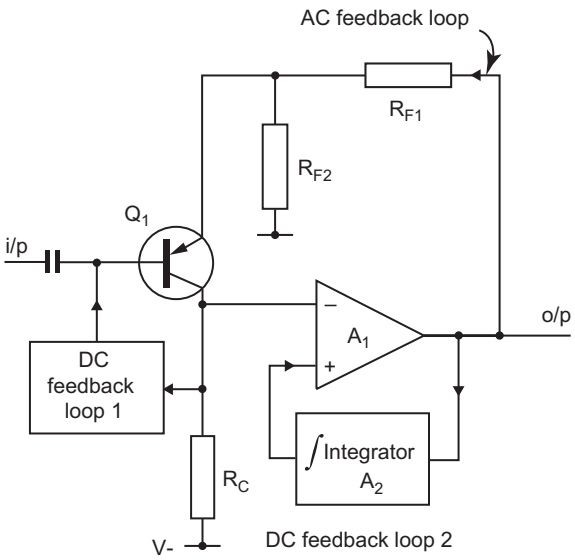


Figure 8.4: Block diagram of the MC preamplifier, with two DC feedback loops

www.raguvaran.puzl.com

As we have seen, MC cartridges vary greatly in their output, and different amplifier gain settings are highly desirable. Usually it would be simple enough to alter R_{F1} or R_{F2}, but here it is not quite so simple. The resistance R_{F2} is not amenable to alteration, as it is kept to the low value of 3.3 Ω by noise considerations, while R_{F1} must be kept up to a reasonable value so that it can be driven to a full voltage swing by an op-amp output. This means a minimum of 500 Ω for the 5534/2. It is intriguing that amplifiers whose output is measured in millivolts are required to handle so much current.

These two values fix a minimum closed-loop gain of about 44 dB, which is much too high for all but the most insensitive cartridges. My solution was to use a ladder output attenuator to reduce the overall gain; this would be anathema in a conventional signal path, because of the loss of headroom involved, but since even an output of 300 mVrms would be enough to overload virtually all MM amplifiers, we can afford to be prodigal with it. If the gain of the head amplifier is set to be a convenient 200 times (+46 dB) then adding output attenuation to reduce the overall gain to a more useful +20 dB still allows a maximum output of 480 mVrms. Lesser degrees of attenuation to give intermediate gains allow greater outputs, and these are summarized in Table 8.2. For testing an Ortofon MC10 was used with +26 dB of gain, giving similar output levels to those of MM cartridges. This highly successful cartridge was in production for 30 years and has only recently been superseded; its impedance is 3.3 Ω.

A final constraint on the attenuator is the need for low output impedances so the succeeding MM input stage can give a good noise performance. The MM input should have been optimized to give its best noise figure with relatively high source impedances, but a low source impedance will still reduce its actual noise output. This means that an output attenuator will need low resistor values, imposing yet more loading on the unfortunate op-amp. This problem was solved by making the attenuator ladder an integral part of the AC feedback loop, as shown in Figure 8.5. This is practicable because it is known that the input impedance of the following MM stage will be too high at 47 kΩ to cause significant gain variations.

TABLE 8.2 Gain options and maximum outputs

Gain	Gain (dB)	Max. output (rms)
10×	+20	480 mV
20×	+26	960 mV
50×	+34	2.4 V
100×	+40	4.6 V
200×	+46	10 V

www.raguvaran.puzl.com

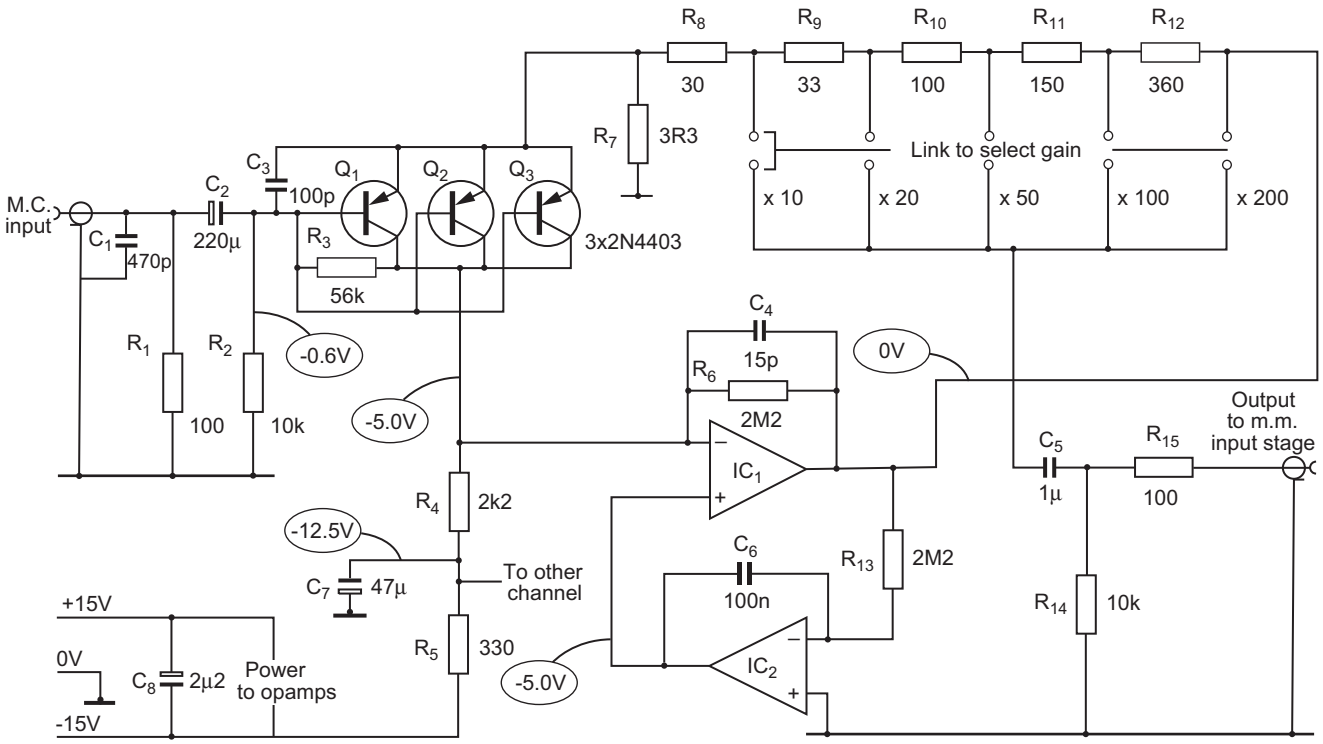


Figure 8.5: Circuit diagram of the MC preamplifier

www.raguvaran.puzl.com

The Complete Circuit

This is shown in Figure 8.5, and closely follows Figure 8.4, though you will note that the input devices have suddenly multiplied themselves by three. Capacitor C1 is soldered on the back of the MC input phono sockets and is intended for EMC immunity rather than cartridge response modification. If the need for more capacitative or resistive loading is felt, then extra components may be freely connected in parallel with R1. If R1 is raised in value, then load resistances of 5 kΩ or more are possible, as the impedance looking into C2 is about 8 kΩ. Capacitor C2 is large to give the input devices the full benefit of the low source impedance, and its value should not be altered. Resistors R2, R3 make up DC loop 1, setting the operating conditions of Q1, Q2, Q3 while R4 is the collector load, decoupled from the supply rail by C9 and R5, which are shared between two stereo channels. Op-amp IC1 is a half of a 5532, providing most of the AC open-loop gain, and is stabilized at HF by C4. R6 has no real effect on normal operation, but is included to give IC1 a modicum of DC negative feedback and hence tidy behavior at power-up, which would otherwise be slow due to the charging time of C2. IC2, half of a TL 072, is the integrator that forms DC loop 2, its time-constant carefully chosen to give ample open-loop gain from IC1 at low frequencies, while avoiding peaking in the LF response that could occur due to the second time-constant of C2.

The ladder resistors R8–R12 make up the combined feedback network and output attenuator, the gain being selected by a push-on link in the prototype. A rotary switch could be used instead, but this should *not* be operated with the system volume up as this will cause loud clicks, due to the emitter current (about 4 mA) of Q1–Q3 flowing through R7, which causes voltage drops down the divider chain. Note that the current through R7 flows down the ground connection back to the PSU. Output resistor R15 ensures stability when driving screened cables, and C5 is included to eliminate any trace of DC offset from the output.

The power supply rails do not need to be specially quiet and a normal op-amp supply is quite adequate.

Performance

The input transistor originally chosen was the 2N4403, a type that was acknowledged as superior for this kind of application for some years, due to its low R_b of about 40 Ω. A single device used in the circuit of Figure 8.5 gives an EIN of −138 dB with a 4 mA collector current and a 3.3 Ω source resistance. The Johnson noise from 3.3 Ω is −147.4 dBu, so we have an NF of 9.4 dB. It was then consistently found that putting devices in parallel without any current-sharing precautions whatever always resulted in a significant improvement in noise performance. On average, adding a second transistor reduced noise by 1.2 dB, and adding a third reduced it by another 0.5 dB, giving an EIN of −139.7 dBu and an NF of 7.7 dB. Beyond this further multiplication was judged unprofitable, so a triple-device input was settled

www.raguvaran.puzl.com

on. The current sharing under these conditions was checked by measuring the voltage across 100 Ω resistors temporarily inserted in the collector paths. With 3.4 mA as the total current for the array it was found after much device swapping that the worst case of imbalance was 0.97 mA in one transistor and 1.26 mA in another. The transistors were not all from the same batch. It appears that, for this device at least, matching is good enough to make simple paralleling practical.

A superior device for low source impedances was the purpose-designed 2SB737, with a stunningly low R_b of 2 Ω. Three of them improved the EIN to −141.0 dBu and the NF to 6.4 dB, albeit at significant cost. Sadly it is now obsolete (why, for heaven's sake?) but can still be obtained from specialized suppliers such as the Signal Transfer Company [4].

You will have spotted that R7, at 3.3 Ω, generates as much noise as the source impedance; this only degrades the NF by 1.4 dB, rather than 3 dB, as most of the noise comes from the transistors.

It would be instructive to compare this design with those of other MC preamplifiers, but this is not at all easy since, as their noise performance is specified in so many different ways, it is virtually impossible to reduce them all to a similar form, particularly without knowing the spectral distribution of the noise. (This article has dealt throughout with unweighted noise referred to the input, over a 400 Hz–20 kHz bandwidth, and with RIAA equalization *not* taken into account.) Nonetheless, I suggest that this design is quieter than most, being within almost 6 dB of the theoretical minimum, with clearly limited scope for improvement. Burkhard Vogel has written an excellent article on the calculation and comparison of MC signal-to-noise ratios [5].

The performance is summarized in Table 8.3. Careful grounding is needed if the noise and crosstalk performance quoted is to be obtained.

When connected to an RIAA-equalized MM stage as described in Chapter 7, the noise output from the MM stage is −93.9 dBu at 10 times gain and −85.8 dBu at 50 times. In the 10 times case the MC noise is actually 1.7 dB lower than for MM mode.

TABLE 8.3 Moving-coil head amp performance figures

Input overload level	48 mVrms
Equivalent input noise	−141.0 dBu, unweighted, without RIAA equalization (3.3 Ω source res.)
Noise figure	6.4 dB (3.3 Ω source res.)
THD	Less than 0.002% at 7 Vrms out (maximum gain) at 1 kHz
	Less than 0.004% at 40 Hz – 20 kHz
Frequency response	+0, −2 dB, 20 Hz – 20 kHz
Crosstalk	Less than −90 dB, 1 – 20 kHz (layout dependent)
Power consumption	20 mA at ±15 V, for two channels

www.raguvaran.puzl.com

References

[1] D. Self, Design of moving-coil head amplifiers, Electronics & Wireless World (December 1987) p. 1206.
[2] Nordholt, Van Vierzen, Ultra low noise preamp for moving-coil phono cartridges, JAES (April 1980) pp. 219–223.
[3] J. Barleycorn (a.k.a. S. Curtis), HiFi For Pleasure (August 1978) pp. 105–106.
[4] http://www.signaltransfer.freeuk.com/
[5] B. Vogel, The sound of silence (calculating MC preamp noise), Electronics World (October 2006) p. 28.

www.raguvaran.puzl.com

CHAPTER 9

Volume and Balance Controls

Volume Controls

A volume control is the most essential knob on a preamplifier – in fact the unhappily named 'passive preamplifiers' usually consist of nothing else but a volume control and an input selector switch. Volume controls in one guise or another are also freely distributed on the control surfaces of mixing consoles, examples being the auxiliary sends and the faders.

A volume control for a hi-fi preamplifier needs to cover at least a 50 dB range, with a reasonable approach to a logarithmic (i.e. linear-in-dB) law, and have a channel balance that is better than ± 1 dB over this range if noticeable stereo image shift is to be avoided.

The simplest volume control is a potentiometer. These components, which are invariably called 'pots' in practice, come with various control laws, such as linear, logarithmic, anti-logarithmic, and so on. The control law is still sometimes called the 'taper', which is a historical reference to when the resistance element was actually physically tapered, so the rate of change of resistance from track-end to wiper could be different at different angular settings. Pots are no longer made this way, but the term has stuck around. An 'audio-taper' pot usually refers to a logarithmic type intended as a volume control.

All simple volume controls have the highest output impedance at the wiper at the -6 dB setting. For a linear pot this is when the control is rotated halfway towards the maximum, at the 12 o'clock position. For a log pot it will be at a higher setting, around three o'clock. The maximum impedance is significant because it usually sets the worst-case noise performance of the following amplification stage. The resistance value of a volume control should be as low as possible, given the loading/distortion characteristics of the stage driving it. This is sometimes called 'low-impedance design'. Lower resistances mean:

1. less Johnson noise from the track resistance;
2. less noise from the current-noise component of the following stage;
3. less likelihood of capacitive crosstalk from neighboring circuitry;
4. less likelihood of hum and noise pickup.

Small Signal Audio Design; ISBN: 9780240521770
Copyright © 2010 Elsevier Ltd; All rights of reproduction, in any form, reserved.

www.raguvaran.puzl.com

TABLE 9.1 Pot law identification letters

ALPS code letter	Pot characteristic
A	Logarithmic
B	Linear
C	Anti-logarithmic
RD	Reverse log

Volume Control Laws

A linear pot is a simple thing – the output is proportional to the angular control setting, and this is usually pretty accurate, depending only on the integrity of the mechanical construction. Linear pots are given the code letter 'B' (see Table 9.1 for more code letters).

Log controls are rather less satisfactory. A typical log pot is not a precision attenuator with a fixed number of dB attenuation for each 10 degrees of shaft rotation. It is instead made up of two linear slopes that roughly approximate a logarithmic law, produced by superimposing two sections of track made with different resistivity material, the overlap usually being at the bottom end of the control setting, such as 20% of full rotation (see Figure 9.1). These pots are usually given the code letter 'A'.

Anti-logarithmic pots are the same only constructed backwards, so that the slope change is at the top end of the control setting; these are typically used as gain controls for amplifying stages rather than as volume controls. These pots are usually given the code letter 'C'.

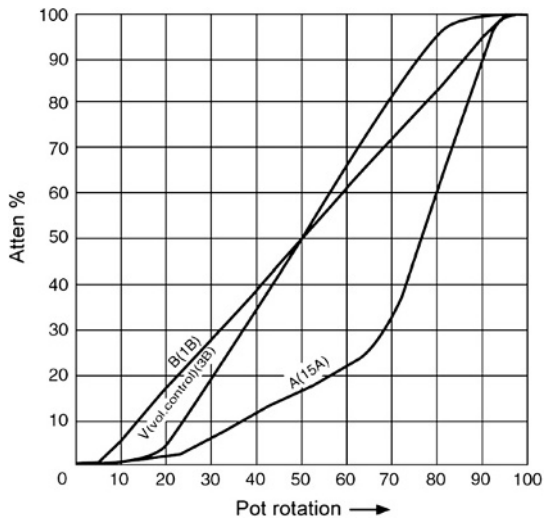


Figure 9.1: The control laws of typical linear and log pots

www.raguvaran.puzl.com

There is a more extreme version of the anti-logarithmic law where the slope change occurs at 10% of rotation instead of 20%. These pots are useful where you want to control the gain of an amplifying stage over a wide range, and still have something like a linear-in-dB control law. Typically they are used to set the gain of microphone input amplifiers, which can have a gain range of 50 dB or more. These pots are given the code 'RD', which stands for reverse D law; I don't think I have ever come across a non-reverse D law. Some typical laws are shown in Figure 9.1.

Please note that the code letters are not adhered to quite as consistently across the world as one might hope. The codes given in Table 9.1 are those used by ALPS, one of the major pot makers, but other people use quite different allocations; for example Radiohm, another major manufacturer, calls linear pots A and log pots B, but they agree that anti-log pots should be called C. Radiohm have several other laws called F, T, S, and X; for example, S is a symmetrical law apparently intended for use in balance controls. It clearly pays to check very carefully what system the manufacturer uses when you're ordering parts.

The closeness of approach to an ideal logarithmic law is not really the most important characteristic of a volume control. Of greater importance is the matching between the two halves of a stereo volume control. It is common for the channel balance of log pots to deteriorate quite markedly at low volume settings, causing the stereo image to shift as the volume is altered. You may take it from me that customers really do complain about this, and a good deal of ingenuity has been applied in attempts to extract good performance from indifferent components.

An important point in the design of volume controls is that their offness – the amount of signal that gets through when the control is at its minimum – is not very critical. This is in glaring contrast to a level control such as an auxiliary send on a mixer channel (see Chapter 17), where the maximum possible offness is very important indeed. A standard log pot will usually have an offness of the order of −90 dB with respect to fully up, and this is quite enough to render even a powerful hi-fi system effectively silent.

Loaded Linear Pots

Since ordinary log pots are not very accurate, many other ways of getting a log law have been tried. An approximation to a logarithmic law can be made by loading the wiper of a linear pot with a fixed resistor R1 to ground, as shown in Figure 9.2. Trace 1 in Figure 9.3 (for a linear pot with no loading) makes it clear that the use of an unmodified linear law for volume control really is not viable; the attenuation is all cramped up at the bottom end.

Adding a loading resistor much improves the law, but the drawback is that this technique really only works for a limited range of attenuation – in fact it only works well if you are looking for a control that varies from around 0 to −20 dB. It is therefore suitable for power amp level controls and aux master gain controls (see Chapter 17 for details of the latter) but is unlikely to

www.raguvaran.puzl.com

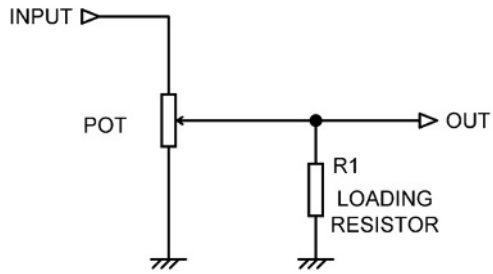


Figure 9.2: Resistive loading of a linear pot to approximate a logarithmic law

be useful for a preamplifier gain control, which needs a much wider logarithmic range. Figure 9.3 shows how the law varies as the value of the loading resistor is changed, and it is pretty clear that whatever its value, the slope of the control law around the middle range is only suitable for emulating the ideal log law labeled '20'. The value of the loading resistor for each trace number is given in Table 9.2.

Figure 9.3 shows that with the optimal loading value (trace 2) the error in emulating a 0 to −20 dB log range is very small, lying within ±0.5 dB over the range 0 to −16 dB; below this the error grows rapidly. This error for trace 2 only is shown in Figure 9.4. Obtaining this accuracy naturally relies on having the right ratio between the pot track resistance and the loading resistor. The resistance of pot tracks is not controlled as closely as for fixed resistors, their tolerance usually being specified as ±20%, so this does present a significant problem. The only solution is to make the loading resistor trimmable, and this approach has been used by at least one console manufacturer.

Another snag to this approach is that when the control is fully up, the loading resistor is placed directly across the input to the volume control, reducing its impedance drastically and possibly

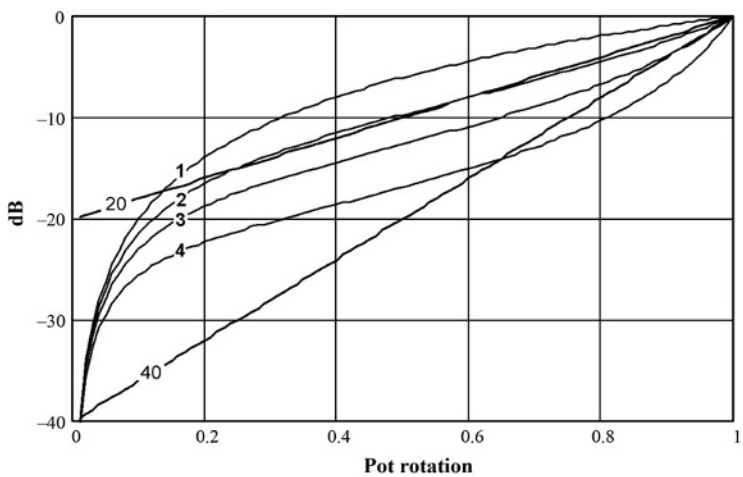


Figure 9.3: Resistive loading of a linear pot: the control laws plotted

www.raguvaran.puzl.com

TABLE 9.2 The loading resistor values used in Figure 9.3

Trace number	Loading resistor R1 value
1	None
2	4k7
3	2k2
4	1 kΩ

causing unhappy loading effects on the stage upstream. However, the main problem is that you are stuck with a 0 to −20 dB range, inadequate for a volume control on a preamplifier.

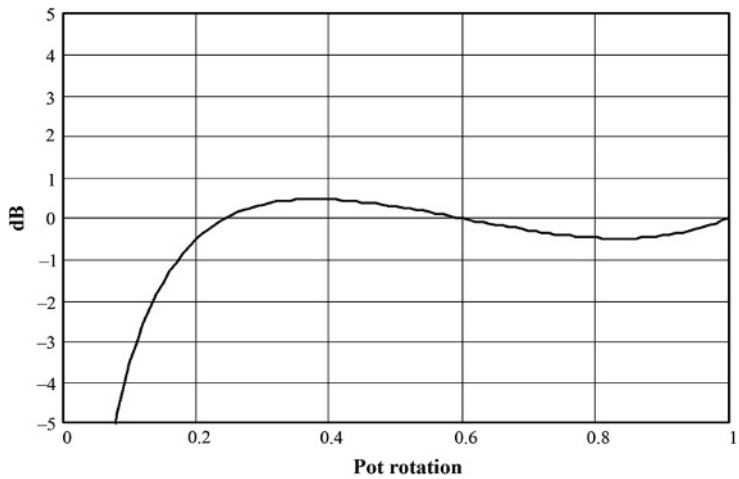


Figure 9.4: Resistive loading of a linear pot: the deviation from an ideal 20 dB log law

The following stage must have a high enough impedance to not affect the volume control law; this obviously also applies to plain logarithmic pots, and to all the passive controls described here.

Dual-Action Volume Controls

In the previous section on loaded linear pots, we have seen that the control law is only acceptably logarithmic over a limited range – too limited for effective use as a volume control in a preamplifier, or as a fader or send control in a mixer. Another technique that can be used to approximate a log law is the cascading of pots, so that their attenuation laws sum in dB. This approach does of course require the use of a four-gang pot for stereo, which is usually objectionable because of increased cost, problems in sourcing, and a worsened volume-control feel. Nonetheless the technique can be useful, so we will give it a quick look.

www.raguvaran.puzl.com

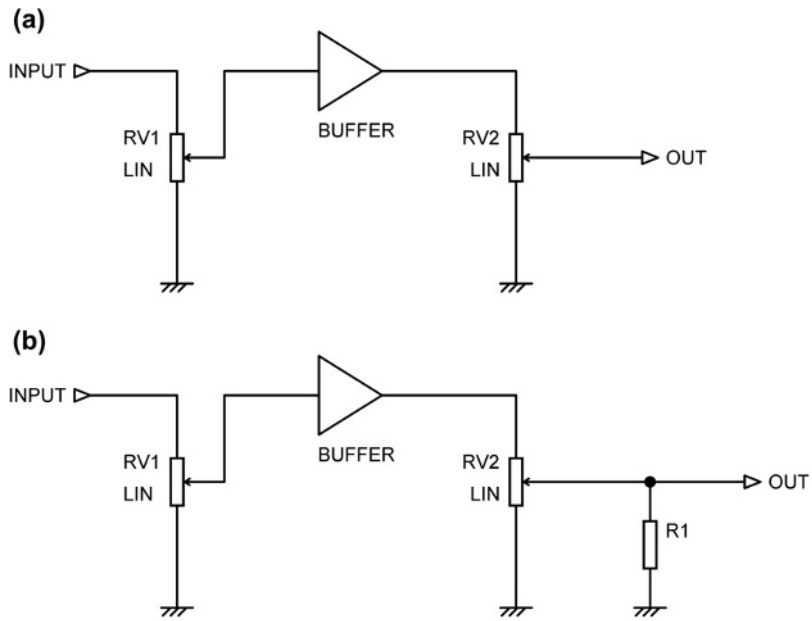


Figure 9.5: Dual-section volume controls: (a) shows two linear pots cascaded and (b) is a linear pot cascaded with a loaded linear pot

It is assumed there is no interaction between the two pots, so the second pot does not directly load the wiper of the first. This implies a buffer stage between them, as shown in Figure 9.5. There is no need for this to be a unity-gain stage, and in fact several stages can be interposed between the two pots. This gives what is usually called a *distributed gain control*, which can be configured to give a better noise/headroom compromise than a single volume control.

Figure 9.6 shows a linear law (trace 1) and the square law (trace 2) made by cascading two linear pots; 20 and 30 dB ideal log lines are shown for comparison, and it is pretty clear that while the square law is much more usable than the linear law, it is still a long way from perfect. It does, however, have the useful property that, assuming the wiper is very lightly loaded by the next stage, the gain is dependent only upon control rotation and not the ratio between fixed 1% resistors and a ±20% pot track resistance. Figure 9.7 shows the deviation of the square law from the 30 dB line; the error peaks at just over +3 dB before it plunges negative.

A *much* better attempt at a log law can be made by cascading a linear pot with a loaded linear pot; the resulting law is shown in Figure 9.8. Trace 1 is the law of the linear pot alone, and trace 2 is the law of a loaded linear 10 kΩ pot with a 2.0 kΩ loading resistor R1 from wiper to ground, alone. The combination of the two is trace 3, and it can be seen that this gives a very good fit to the 40 dB ideal log line, and good control over a range of at least 35 dB.

www.raguvaran.puzl.com

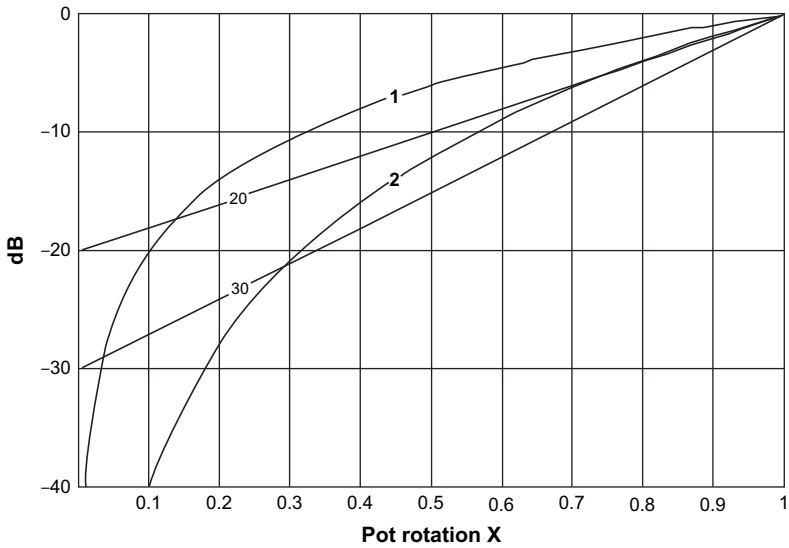


Figure 9.6: The dual-section volume control. Trace 1 is a linear pot, while trace 2 is a square law obtained by cascading two ganged linear pots

Figure 9.9 shows the deviation of the combined law from the 40 dB line; the error now peaks at just over ±1 dB. Unfortunately, adding the loading resistor means that once more the gain is dependent on the ratio between a fixed resistor and the pot track resistance.

Passive volume controls of various types can of course also be cascaded with an active volume control stage. Active volume stages are dealt with later in this chapter.

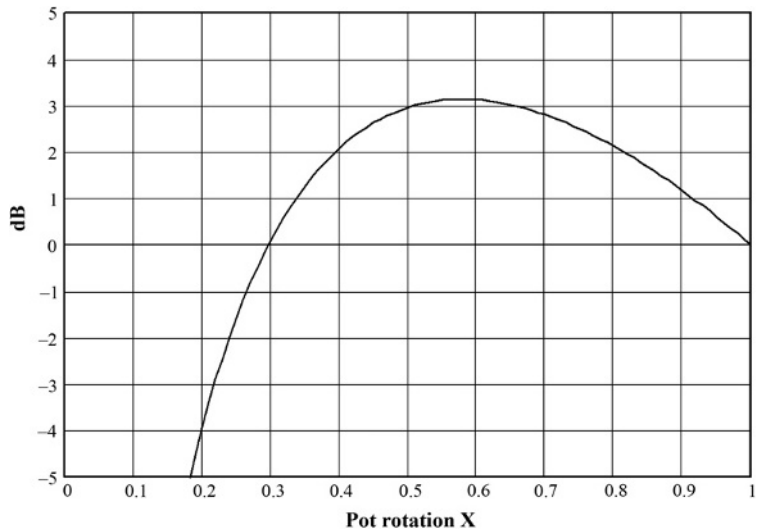


Figure 9.7: Dual-section volume control: the deviation of the square law from the ideal 30 dB log line

www.raguvaran.puzl.com

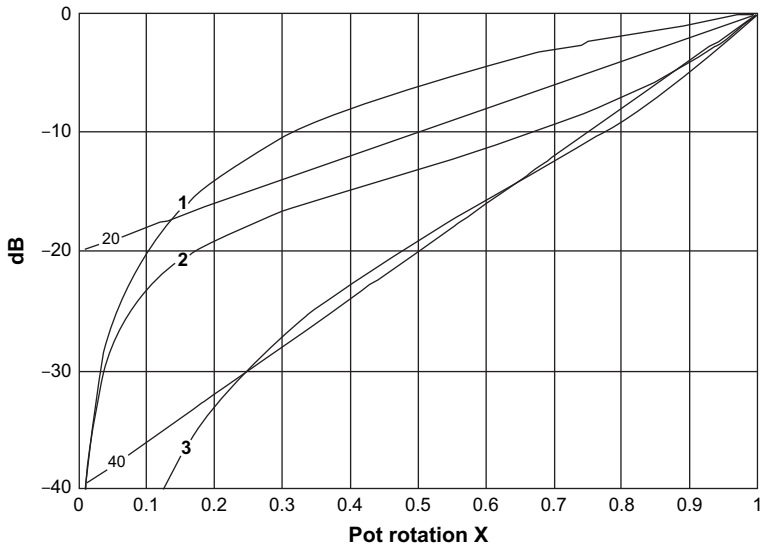


Figure 9.8: Dual-section volume control with improved law: linear pot cascaded with loaded linear pot. Trace 3 is the combination of traces 1 and 2 and closely fits the 40 dB log line

Tapped Volume Controls

The control law of a linear pot can be radically altered if it has a center tap (see Chapter 17 for the use of tapped pots for LCR panning). This can be connected to a potential divider that has a low impedance compared with the pot track resistance, and the attenuation at the tap point

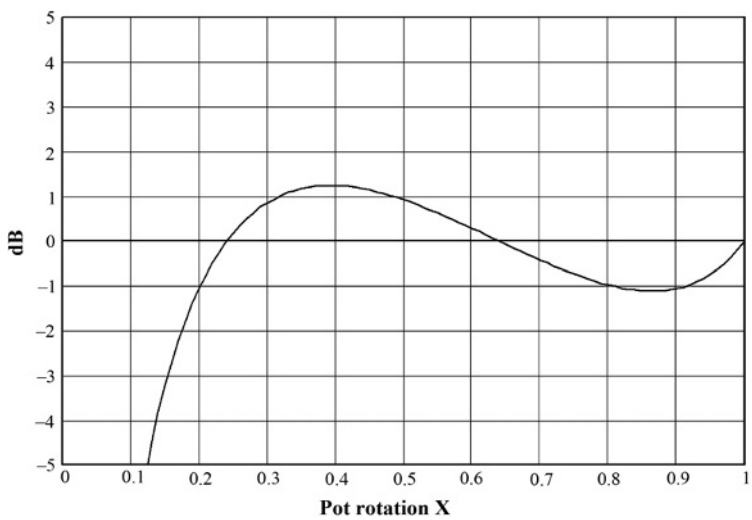


Figure 9.9: Dual-section volume control with improved law: deviation of the control law from the ideal 40 dB log line. Much better than in Figure 9.7

www.raguvaran.puzl.com

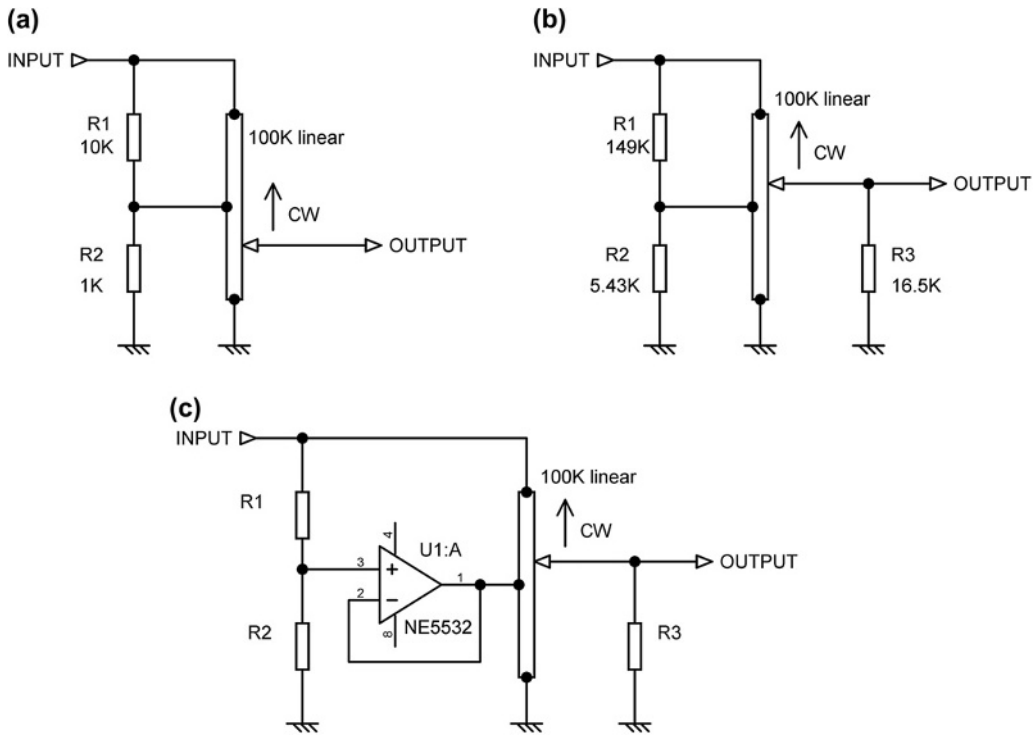


Figure 9.10: Tapped volume control: (a) unloaded; (b) loaded; (c) with active control of tap voltage

altered independently of other parameters. Figure 9.10 shows both unloaded and loaded versions of the arrangement.

The unloaded version shown in Figure 9.10(a) is arranged so that the attenuation at the tap is close to −20 dB. This gives the law shown in Figure 9.11; it approximates to a 33 dB log line, but there is an abrupt change of slope as the pot wiper crosses the tapping point.

Note that the track resistance has to be a good deal higher than that of the fixed resistors, so that they control the level at the tap. This means that with the values shown the source impedance at the wiper can be as high as 12.5 kΩ when it is halfway between the tap and one end, and this may degrade the noise performance, particularly if the following stage has significant current noise. A normal 10 kΩ pot, of whatever law, has a maximum output impedance of only 2.5 kΩ. To match this figure the values shown would have to be scaled down by a factor of 5, so that R1 = 2k, R2 = 200 Ω, and the track resistance is a more normal 10 kΩ. This scaling is quite practical, the load on the previous stage now being 1.62 kΩ.

www.raguvaran.puzl.com

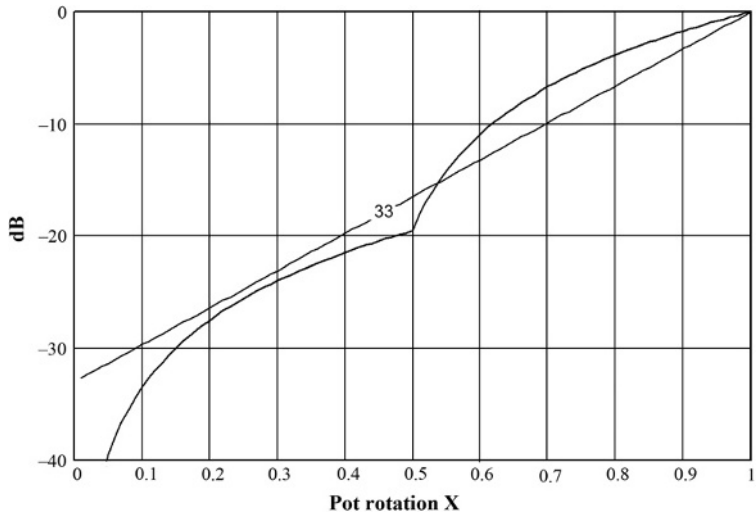


Figure 9.11: The law of an unloaded tapped volume control with −20 dB at the tap

The law error is shown in Figure 9.12. This version of a tapped linear volume control covers a range of about 30 dB, and almost keeps the errors within ±3 dB, but that abrupt change of slope at the tap point is somewhat less than ideal.

A much better approach to a log law is possible if a loading resistor is added to the wiper, as with the loaded linear pot already examined (see Figure 9.10(b) for the circuit arrangement and Figure 9.13 for the control law). If the resistors are correctly chosen, using the ratios given in Figure 9.10(b), the law can be arranged to have no change of slope at the tapping, and the

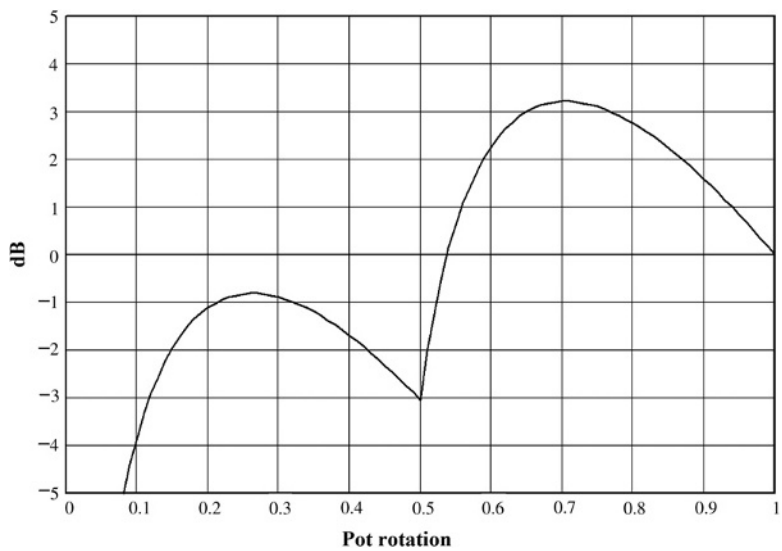


Figure 9.12: The deviation of the unloaded tapped volume control law from an ideal 33 dB log line

www.raguvaran.puzl.com

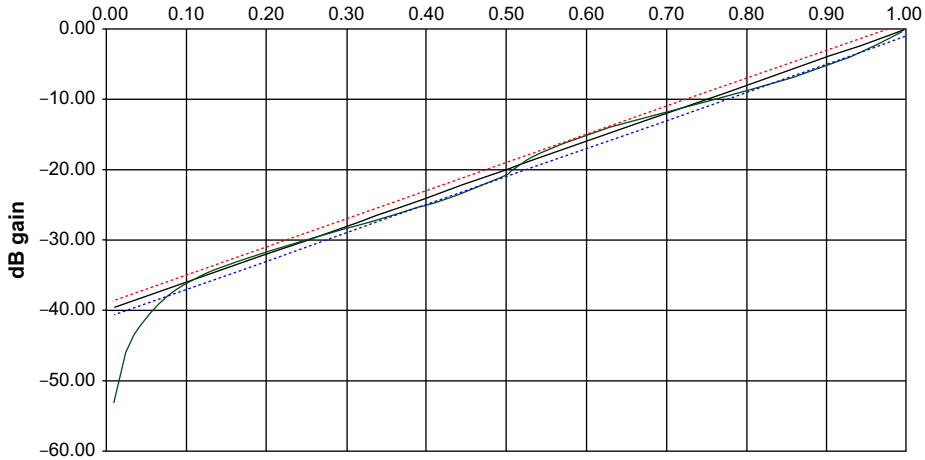


Figure 9.13: The law of a loaded and tapped volume control with −20 dB at the tap. The dotted lines shows the width of a ±1 dB error band around the 40 dB ideal line

deviation from a 40 dB ideal line will be less than 1 dB over the range from 0 to −38 dB. The exact values of the resistors are given rather than the nearest preferred values. It is clear that this is a very effective way of giving an accurate law over a wide range, and it is widely used in making high-quality slide faders. Note that the accuracy still depends on good control of the track resistance value compared with the fixed resistors.

If the value of the resistors connected to the tap is such that the loading effect on the previous stage becomes excessive, one possible solution is shown in Figure 9.10(c), where R1 and R2 are kept reasonably high in value, and a unity-gain op-amp buffer now holds the tap point in a vice-like grip due to its low output impedance. The disadvantage is that the noise of the op-amp, and of R1 in parallel with R2, is fed directly through at full level when the wiper is near the tap. Often this will be at a lower level than the noise from the rest of the circuitry, but it is a point to watch. Note that because of the low-impedance drive from the op-amp, the values of the fixed resistors will need to be altered from those of Figure 9.10(b) to get the best control law.

If the best possible noise performance is required, then it is better to increase the drive capability of the previous stage and keep the resistors connected to the tap low in value; if this is done by paralleling op-amps then the noise performance will be improved rather than degraded.

Slide Faders

So far as the design of the adjacent circuitry is concerned, a fader can normally be regarded as simply a slide-operated logarithmic potentiometer. Inexpensive faders are usually made using the same two-slope carbon-film construction as for rotary log volume controls, but the more expensive and sophisticated types use a conductive-plastic track with multiple taps connected

www.raguvaran.puzl.com

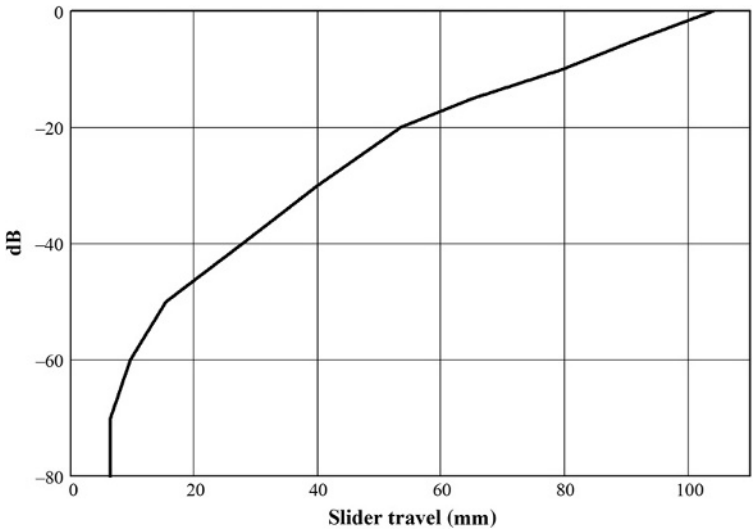


Figure 9.14: A typical control law for a 104 mm fader, showing how the attenuation is spread out over the upper part of the travel

to a resistor ladder, as described in the previous section. This allows much better control over the fader law.

High-quality faders typically have a conductive-plastic track, contacted by multiple gold-plated metal fingers to reduce noise during movement. A typical law for a 104 mm travel fader is shown in Figure 9.14. Note that a fader does not attempt to implement a linear-in-dB log scale; the attenuation is spread out over the top part of the travel and much compressed at the bottom. This puts the greatest ease of control in the range of most frequent use; there is very little point in giving a fader precise control over signals at −60 dB.

A so-called infinity-off feature is incorporated into the more sophisticated faders. Figure 9.15(a) shows the straightforward construction used in smaller and less expensive mixers. There is some end resistance R_{ef} at the bottom of the resistive track that compromises the offness, and it is further compromised by the voltage drop down the resistance R_g of the ground wire that connects the fader to the channel PCB.

Figure 9.15(b) shows an infinity-off fader. When the slider is pulled down to the bottom of its travel, it leaves the resistive track and lands on an end section with a separate connection back to the channel module ground. No signal passes down this ground wire and so there is no voltage drop along its length. This arrangement gives extremely good maximum attenuation, orders of magnitude better than the simple fader, though of course 'infinity' is always a tricky thing to claim.

Faders are sometimes fitted with fader-start switches; these are microswitches that are actuated when the slider moves from the 'off' position.

www.raguvaran.puzl.com

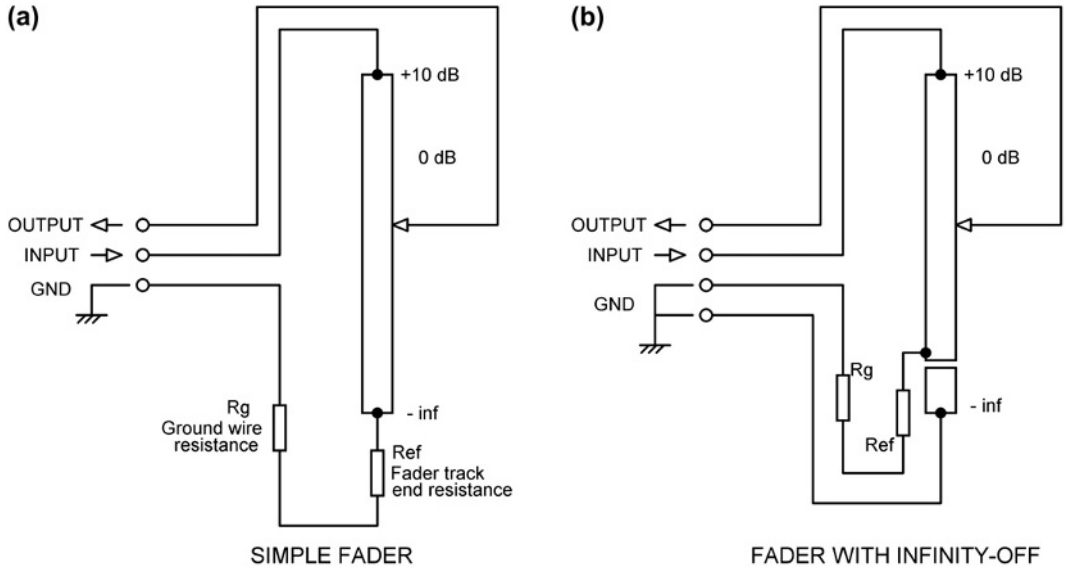


Figure 9.15: A simple fader (a) and a more sophisticated version (b) with an 'infinity-off' section to maximize offness

Active Volume Controls

Active volume controls have many advantages. As explained in the section on preamplifier architectures, the use of an active volume control removes the dilemma concerning how much gain to put in front of the volume control and how much to put after it. An active gain control must fulfill the following requirements:

1. The gain must be smoothly variable from a maximum, usually well above unity, down to effectively zero, which in the case of a volume control means at least −70 dB. This at once rules out series-feedback amplifier configurations as these cannot give a gain of less than 1. Since the use of shunt feedback implies a phase inversion, this can cause problems with the preservation of absolute polarity.

2. The control law relating shaft rotation and gain should be a reasonable approximation to a logarithmic law. It does not need to be strictly linear-in-decibels over its whole range; this would give too much space to the high-attenuation end, say around −60 dB, and it is better to spread out the middle range of −20 to −50 dB, where the control will normally be used. These figures are naturally approximate as they depend on the gain of the power amplifier, speaker sensitivity, and so on. A major benefit of active gain controls is that they give much more flexible opportunities for modifying the law of a linear pot than does the simple addition of a loading resistor, which was examined and found somewhat wanting earlier in this chapter.

www.raguvaran.puzl.com

3. The opportunity to improve channel balance over the mediocre performance given by the average log pot should be firmly grasped. Most active gain controls use linear pots and arrange the circuitry so that these give a quasi-logarithmic law. This approach can be configured to remove imbalances due to track resistance tolerances.

4. The noise gain of each amplifier involved should be as low as possible.

5. As for passive volume controls, the circuit resistance values should be as low as practicable to minimize Johnson noise and capacitative crosstalk.

Figure 9.16 shows a collection of possible active volume configurations, together with their gain equations. Each amplifier block represents an inverting stage with a large gain $-A$, i.e. enough to give plenty of negative feedback at all gain settings. It can be regarded as an op-amp with its non-inverting input grounded. Figure 9.16(a) simply uses the series resistance of a log pot to set the gain. While you get the noise/headroom benefits of an active volume control, the retention of a log pot with its two slopes and resulting extra tolerances means that the channel balance is no better than that of an ordinary passive volume control using a log pot. It may in fact be worse, for the passive volume control is truly a potentiometer, and if it is lightly loaded differences in track resistance due to process variations should at least partially cancel, and one can at least rely on the gain being exactly 0 dB at full volume. Here, however, the pot is actually acting as a variable resistance, so variations in its track resistance compared with the fixed R1 will cause imbalance; the left and right gain will not even be the same with the control fully up. Given that pot track resistances are usually subject to a ±20% tolerance, it would be possible for the left and right channel gains to be 4 dB different at full volume. This configuration is not recommended.

Figure 9.16(b) improves on Figure 9.16(a) by using a linear pot and attempting to make it quasi-logarithmic by putting the pot into both the input and feedback arms around the amplifier. It is assumed that a maximum gain of 20 dB is required; it is unlikely that a preamplifier design will require more than that. The result is the law shown in Figure 9.17, which can be seen to approximate fairly closely to a linear-in-decibels line with a range of 0 to −44 dB. This is a result of the essentially square-law operation of the circuit, in which the numerator of the gain equation increases as the denominator increases. This is in contrast to the loaded linear pot case described earlier, which approximates to a 20 dB line.

The deviation of the control law from the 44 dB line is plotted in Figure 9.18, where it can be seen that between control rotations of 0.1 and 1, and a gain range of almost 40 dB, the maximum error is ±2.5 dB. This sort of deviation from an ideal law is not very noticeable in practice. A more serious issue is the way the gain heads rapidly south at rotations less than 0.1, with the result that volume drops rapidly towards the bottom of the travel, making it more difficult to set low volumes to be where you want them. Variations in track resistance tend to

www.raguvaran.puzl.com

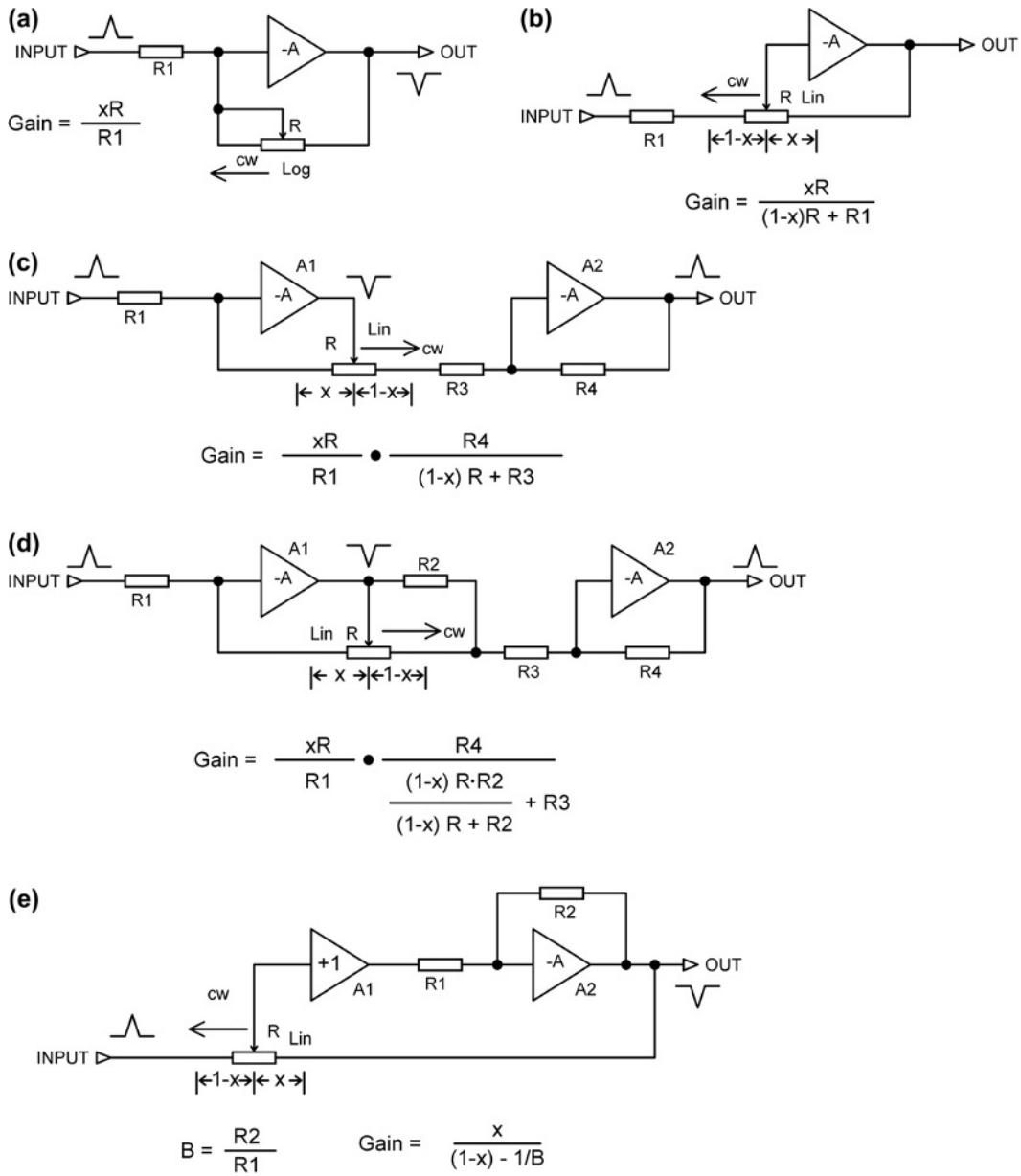


Figure 9.16: Active volume control configurations

cancel out for middle volume settings, but at full volume the gain is once more proportional to the track resistance and therefore subject to large tolerances.

The configuration in Figure 9.16(c) also puts the pot into an input arm and a feedback arm, but in this case in separate amplifiers: the feedback arm of A1 and the input arm of A2. It requires

www.raguvaran.puzl.com

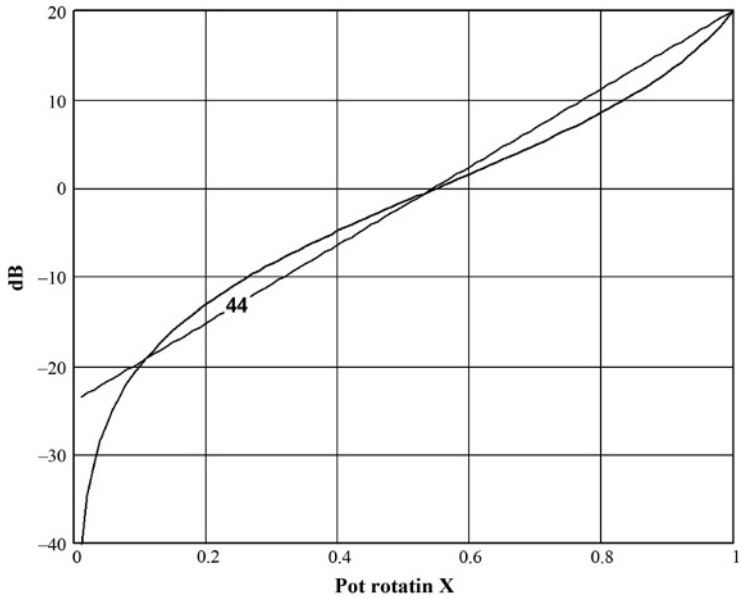


Figure 9.17: The control law of the active volume control in Figure 9.16(b)

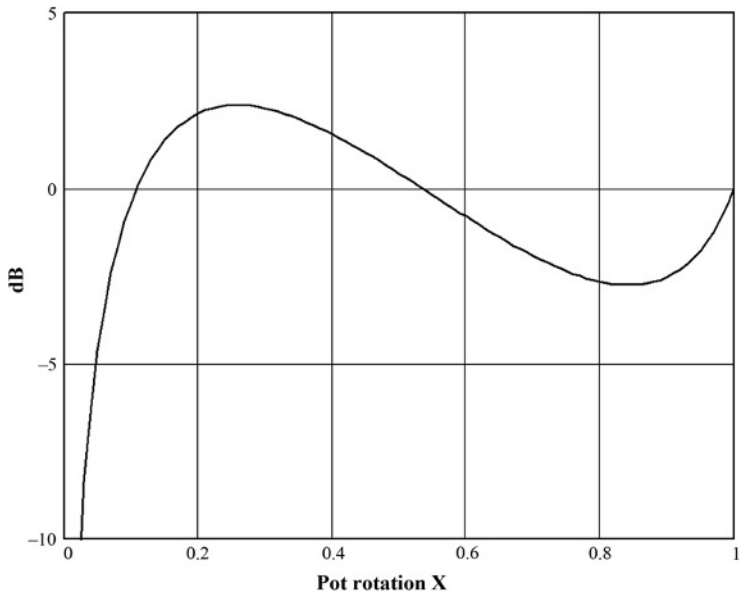


Figure 9.18: The deviation of the control law in Figure 9.17 from an ideal 44 dB logarithmic line

two amplifier stages, but as a result the output signal is in the correct phase. When configured with R = 10 kΩ, R1 = 10 kΩ, R3 = 1 kΩ, and R4 =10 kΩ, it gives exactly the same law as Figure 9.16(b), with the same maximum error of ±2.5 dB. It therefore may be seen as a pointless extra complication, but in fact the extra resistors involved give a greater degree of design freedom.

www.raguvaran.puzl.com

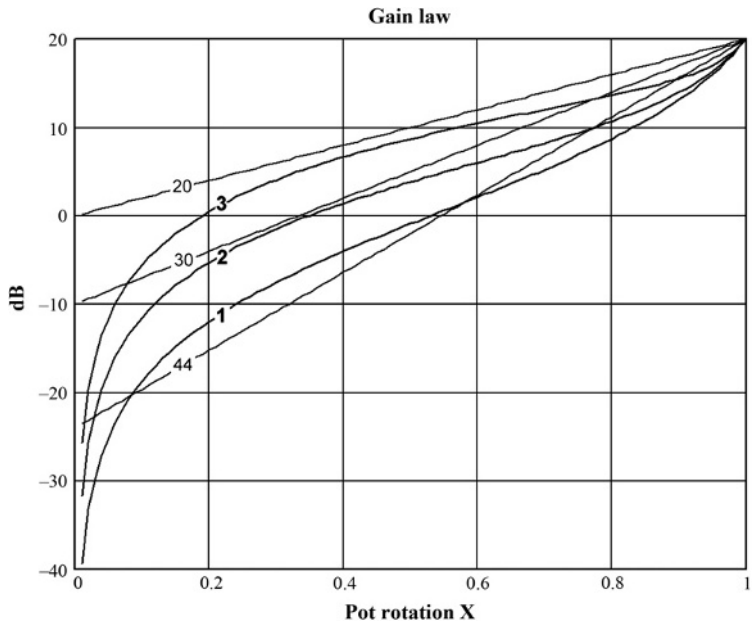


Figure 9.19: The control law of the active volume control in Figure 9.16(d)

In some cases a linear-in-decibels line with a range of 44 dB, which is given by the active gain stages already looked at, is considered too rapid; less steep laws can be obtained from a modified version of Figure 9.16(c), by adding another resistor R2 to give the arrangement in Figure 9.16(d). This configuration was used in the famous Cambridge Audio P50 integrated amplifier, introduced in 1970. When R2 is very high, the law approximates to that of Figure 9.16(c). With R2 reduced to 4 kΩ, the law is modified to trace 2 in Figure 9.19; the law is shifted up, but in fact the slope is not much altered, and is not a good approximation to the ideal 30 dB log line, labeled '30'. When R2 is reduced to 1 kΩ, the law is as trace 3 in Figure 9.19, and is a reasonable fit to the ideal 20 dB log line, labeled '20'. Unfortunately, varying R2 can do nothing to help the way that all the laws fall off a cliff below a control rotation of 0.1, and in addition the problem remains that the gain is determined by the ratio between fixed resistors, for which a tolerance of 1% is normal, and the pot track resistance, with its ±20% tolerance. For this reason none of the active gain controls considered so far is going to help with channel balance problems.

The Baxandall Active Volume Control

The active volume control configuration in Figure 9.16(e) is due to Peter Baxandall. Like so many of the innovations conceived by that great man, it authoritatively solves the problem it addresses [1]. Figure 9.20 shows the law obtained with a maximum gain of +20 dB; the best-fit ideal log line is now 43 dB. There is still a rapid fall-off at low control settings.

www.raguvaran.puzl.com

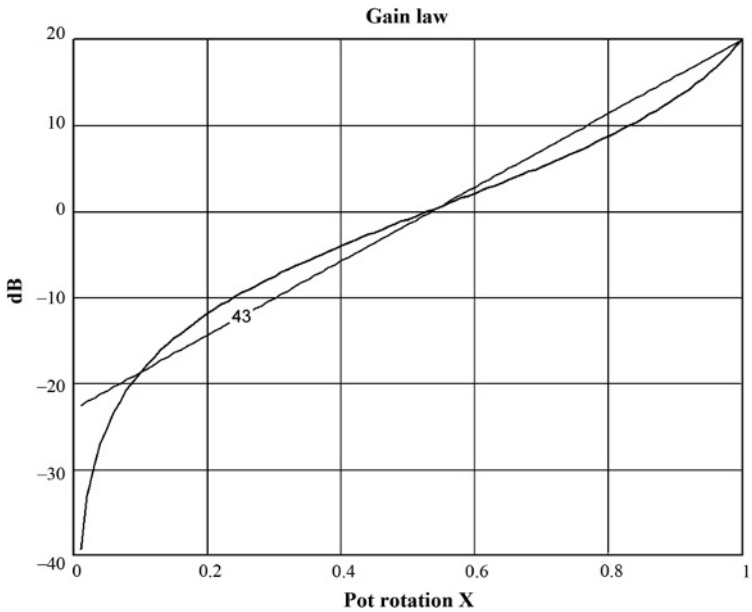


Figure 9.20: The control law of the Baxandall active volume control in Figure 9.16(e)

You will note that there are no resistor or track resistance values in the gain equation; the gain is only a function of the pot rotation and the maximum gain set up by R1, R2. As a result quite ordinary dual linear pots can give very good channel matching. When I tried a number of RadioOhm 20-mm-diameter linear pots, the balance was almost always within 0.3 dB over a 46 dB gain range, with occasional excursions to an error of 0.6 dB.

However, the one problem that the Baxandall configuration cannot solve is channel imbalance due to mechanical deviation between the wiper positions. I have only once found that the Baxandall configuration did not greatly improve channel balance; in that case the linear pots I tried, which came from the same Chinese source as the log pots that were provoking customer irritation, had such poor mechanical alignment that the balance improvement obtained was small, and not worth the extra circuitry.

Note that all the active gain configurations require a low-impedance drive if they are to give the designed gain range – don't try feeding them from, say, the wiper of a balance control pot. The Baxandall configuration inherently gives a phase inversion that can be highly inconvenient if you are concerned with preserving absolute phase, but this can be undone by an inverting tone-control stage.

An important point is that while at first glance the Baxandall configuration looks like a conventional shunt-feedback control, its action is modified by the limited gain set by R1 and R2. This means that the input impedance of the stage falls as the volume setting is increased,

www.raguvaran.puzl.com

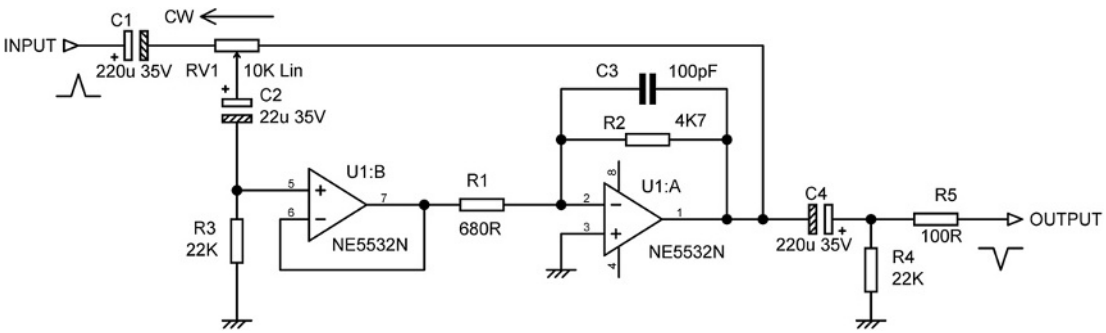


Figure 9.21: A practical Baxandall active volume control, as used in several preamplifier designs

but does *not* drop to zero. With the values shown in Figure 9.21, input impedance falls steadily from a maximum of 10 kΩ at zero gain, to a minimum of 1.27 kΩ at maximum gain. If the preceding stage is based on a 5532 it will have no trouble driving this.

Another consequence of the gain of the A2 stage is that the signals handled by the buffer A1 are never very large. This means that R1, and consequently R2, can be kept low in value to reduce noise without placing an excessive load on the buffer.

All the active volume controls examined here, including the otherwise superior Baxandall configuration, give a gain law that falls rapidly in the bottom tenth of control rotation. It is not easy to see that there is any cure for this.

A Practical Baxandall Active Gain Stage

I have designed several preamplifiers using a Baxandall active volume control [2,3]. The practical circuitry I employed is shown in Figure 9.21; this includes arrangements to deal with the significant bias currents of the 5532 op-amp [1]. The maximum gain is set to +17 dB by the ratio of R1, R2, to amplify a 150 mV line input to 1 V with a small safety margin.

This active volume-control stage gives the usual advantages of lower noise at gain settings below maximum, and excellent channel balance that depends solely on the mechanical alignment of the dual linear pot – all mismatches of its electrical characteristics are canceled out. Note that in both the preamplifier designs referenced here all the pots were identical at 10 kΩ linear, apart from the question of center detents, which are desirable only on the balance, and treble and bass boost/cut controls.

The values given here are as used in the Precision Preamplifier 96 [3]. Compared with Ref. [2], noise has been reduced slightly by an impedance reduction on the gain-definition network R1, R2. The limit on this is the ability of buffer U1:B to drive R1, which has a virtual earth at its

www.raguvaran.puzl.com

other end. C3 ensures HF stability if there are excess phase-shifts due to stray capacitance. C1 prevents any DC voltages from circuitry upstream from reaching the volume-control stage. The input bias current of U1:B will produce a small voltage drop across R3, and C2 prevents this from reaching the control pot. Since two terminals of the pot are DC blocked, it is now permissible to connect the third terminal to the output of U1:A as no current can flow through the control. The offset voltage at the output of U1:B will be amplified by U1:A, but should still be much too small to have any significant effect on available voltage swing, and it is prevented from leaving this stage by DC-blocking capacitor C4. R4 is a drain resistor to prevent voltage building up on the output due to leakage through C4, and R5 ensures stability by isolating the stage from load capacitance downstream, such as might be caused by the use of screened cable. Note that R4 is connected before R5, to prevent any loss of gain. The amount of loss is of course very small in one stage, but it can build up irritatingly in a large system if this precaution is not observed.

This gain control is used in the second section of the complete preamplifier described in Chapter 11. Table 9.3 gives the noise performance.

TABLE 9.3 Noise performance of active gain control at various volume settings

Setting	Noise out (dBu)
Zero gain	−114.5
Unity gain	−107.4
Full volume	−90.2

The figure at full volume may look a bit mediocre, but results from the use of +17 dB of gain; at normal volume settings the noise output is well below −100 dBu.

Potentiometers and DC

As noted in the previous section, it is never a good idea to have DC flowing through any part of a potentiometer. If there is a DC voltage between the ends of the pot track there will be rustling noises generated as the wiper moves up and down over minor irregularities in track resistance.

Feeding a bias current through a wiper to the next stage tends to create more serious noise because the variations in wiper contact resistance are greater. This tends to get worse as the track surface becomes worn. This practice is often acceptable for FET input op-amps like the TL072, but it is definitely not a good idea for bipolar op-amps such as the 5532 because the bias current is much greater and so therefore is the noise on wiper movement. AC coupling is

www.raguvaran.puzl.com

essential when using bipolar op-amps. If you are using electrolytic capacitors then make sure that the coupling time-constant is long enough for capacitor distortion to be avoided (see Chapter 2).

Motorized Potentiometers

Motorized pots are simply ordinary pots driven by an attached electric motor, heavily geared down and connected to the control shaft through a slipping clutch. This clutch allows manual adjustment of the volume when the motor is off, and prevents the motor stalling when the pot hits the end of its travel; limit switches are not normally used. Motorized pots are now considerably cheaper than they used to be, due to the manufacture of components in China that represent a very sincere homage to designs by ALPS and others, and they appear in lower to middle-range integrated amplifiers. Motorization can be added to any control that uses a rotary pot.

In many ways motorized pots are the ideal way to implement a remote-controlled volume function. There is no variable gain electronics to add noise and distortion, manual control is always available if you can't find the remote, and their setting is inherently non-volatile, as the knob stays where it was left when you switch off.

A disadvantage is that the 'feel' of a motorized pot is pretty certain to be worse than a normal control's, because of the need for the slipping clutch between the control shaft and the motor; a large-diameter knob helps with this. The channel balance is of course no better than if the same pot was used as a manual control.

Since the motor has to be able to run in either direction, and it is simplest to run it from a single supply, an H-bridge configuration, as shown in Figure 9.22, is used to drive it. Normally all

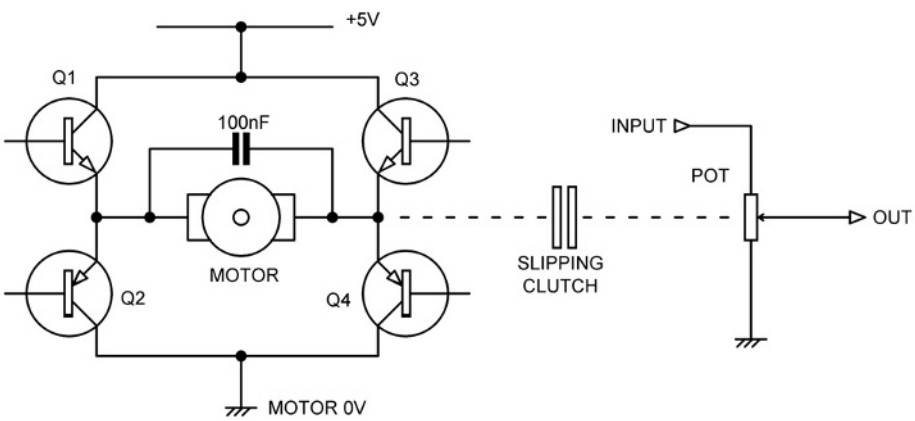


Figure 9.22: Control circuitry for a motorized volume control

www.raguvaran.puzl.com

four transistors are off. To run the motor in one direction Q1 and Q4 are turned on; to run in the other direction Q2 and Q3 are turned on. The H-bridge and associated logic to interface with a microcontroller can be obtained in convenient ICs such as the BA6218 by Rohm. This IC can supply an output current of up to 700 mA. Two logic inputs allow four output modes: forward, reverse, idle (all H-bridge transistors off), and dynamic braking (motor shorted via ground). The logic section prevents input combinations that would turn on all four devices in the H-bridge and create (briefly) electronic mayhem.

It would appear that there might be problems with electrical noise from the motor getting into the audio circuitry, but I have not myself found this to be a problem. In the usual version the motor is screened with a layer of what appears to be grain-oriented silicon steel (GOSS) to keep magnetic effects under control, and the motor terminals are a long way from the audio terminals. A 100 nF capacitor across the motor terminals, and as close to them as practicable, is always a good idea. The motor should be driven from a separate non-audio supply unless you're really looking for trouble; motorized controls with 5 V motors are popular as they can run off the same +5 V rail as a housekeeping microcontroller. Once again I have never had problems with motor noise interfering with a microcontroller.

Linear faders, as used on mixing consoles, are also sometimes motorized, not for remote control, but to allow previously stored fader movements to be played back in an automatic mixdown system. A linear servo track next to the audio track allows accurate positioning of the fader. Motorization has usually been done by adding a small electric motor to one end of the fader, and moving the control knob through a mechanism of string and pulleys that is strongly reminiscent of an old radio dial. Such arrangements have not always been as reliable as one might have hoped.

Stepped Volume Controls

The great feature of potentiometer-based volume controls is that they have effectively infinite resolution, and you can set exactly the level you require. If, however, you are prepared to forgo this and accept a volume control that changes in steps, a good number of new possibilities open up, and in return promise much greater law accuracy and channel balance. The technologies available include rotary-switched resistive attenuators, relay-switched resistive attenuators, switched multi-tap transformers, and specialized volume-control ICs. These options are examined below.

The obvious question is: how small a step is needed to give satisfactory control? If the steps are made small enough they are imperceptible, provided there are no switching transients, but there are powerful economic reasons for not using more steps than necessary. Two-decibel steps are widely considered acceptable for in-car entertainment (implemented by an IC) but my view is that serious hi-fi requires 1-dB steps.

www.raguvaran.puzl.com

Switched Attenuator Volume Controls

For high-end products where the imperfections of a ganged-potentiometer volume control are not acceptable, much superior accuracy can be achieved by using switched attenuators to control level. It is well known that the BBC for many years used rotary faders that were stud-switched attenuators working in 2-dB steps; some of these were still in use in 1961.

The normal practice is to have a large rotary switch that selects a tap on a resistor ladder; since the ladder can be made of 1% tolerance components the channel matching is much better than that of the common dual-gang pot. A stereo control naturally requires two resistor ladders and a two-pole switch. The snag is of course the much greater cost; this depends to a large extent on how many control steps are used. The resistor ladders are not too costly, unless exotic super-precision parts are used, but two-pole switches with many ways are neither cheap nor easy to obtain.

At the time of writing, one commercial preamp offers 12 5-dB steps; the component cost has been kept down by using separate switches for left and right channels, which is not exactly a triumph of ergonomics. Another commercial product has an 11-position ganged switched attenuator. In my opinion neither offers anything like enough steps.

The largest switches readily available are made up from one-pole 12-way wafers. The most common version has a break-before-make action, but this causes clicky transients due to the interruption of the audio waveform. Make-before-break versions can usually be obtained and these are much more satisfactory as the level changes as the switch is rotated are much smaller, and the transients correspondingly less obtrusive.

Moving beyond 12-way, a relatively popular component for this sort of thing is a 24-position switch from the ELMA 04 range; this seems to be the largest number of ways they produce. A bit of care is needed in selecting the part required as one version has 10 μm of silver on the contracts with a protective layer of only 0.2-μm-thick gold. This very thin layer is for protection during storage and transport only and in use will wear off quite quickly exposing the silver, which is then subject to tarnishing, with the production of non-conductive silver sulfide; this version should be avoided unless used in a sealed environment. Other versions of these switches have 3-μm-thick gold on the contacts, which is much more satisfactory. They can be obtained with one, two, or four 24-way wafers, but they are not cheap. At the time of writing (January 2009) two-wafer versions are being advertised by third parties at about $130 each, which means that in almost all cases the volume switch is going to cost a lot more than the rest of the preamplifier parts put together.

A switched attenuator can be made very low impedance to minimize its own Johnson noise and the effect of the current noise of the following stage. The limiting factors are that the attenuator input must present a load that can be driven with low distortion from the preceding stage, and that the resistor values at the bottom of the ladder do not become too

www.raguvaran.puzl.com

small for convenience. An impedance of around 1000 Ω from top to bottom is a reasonable choice. This means that the highest output impedance (which is at the −6 dB setting, a very high setting for a volume control) will be only 250 Ω. This has a Johnson noise of only −128.2 dBu (22 kHz bandwidth, 25°C) and is unlikely to contribute much to the noise output of a system. The choice of around 1000 Ω does assume that your chosen range of resistors goes down to 1 Ω rather than stopping at 10 Ω, though if necessary the lower values can of course be obtained by paralleling resistors.

Assuming you want to use a 12-way switch, a possible approach is to use 12 steps of 4 dB each, covering the range 0 to −44 dB. My view is that such steps are too large, and 2 dB is much more usable, but you may not want to spend a fortune on a rotary switch with more ways. In fact, 1-dB steps are really required so you can always get exactly the volume you want, but implementing this with a single rotary switch is going to be very difficult, as it implies something like a 60- or 70-way switch. My own preamplifier just happens to be a relay-switched Cambridge Audio 840E with some 90-odd 1-dB steps.

There is no reason why the steps have to be equal in size, and it could be argued that the steps should be smaller around the center of the range and larger at the top and bottom, to give more resolution where the volume control is most likely to be used. It does not really matter if the steps are not exactly equal – the vital thing is that they should be identical for the left and right channels to avoid image shift.

Figure 9.23(a) shows a typical switched attenuator of this sort, with theoretically exact resistor values, so the size of each step is 4.0000 dB; in fact each step is accurate to five decimal places. In view of resistor tolerances, such accuracy may appear pretty pointless, apart from the warm feeling it produces, but it costs nothing to start off with more accuracy than you need. The total resistance of the ladder is 1057 Ω, which is a very reasonable load on a preceding stage, so long as it is implemented with 5532s or similar.

A lot depends on what range of resistor values you have access to. When I began to design preamplifiers, E12 resistors (12 values per decade) were the norm, and E24 resistors (24 values per decade) were rather rare and expensive. This is no longer the case, and E24 is freely available. Figure 9.23(b) shows the same attenuator with the nearest E24 value used for each resistor. The attenuation at each tap is still very accurate, the error never exceeding 0.12 dB except for the last tap, which is 0.25 dB high; this could be cured by making R12 a parallel combination of 6.8 and 330 Ω and making R11 3.9 Ω, which reduces the last-tap error to 0.08 dB.

The next most prolific resistor range is E96, with no less than 96 values in a decade – nobody seems interested in making an E48 range as such, though it is of course just a subset of E96. Using the nearest E96 value in the attenuator, we get Figure 9.23(c), where the taps are accurate to within 0.04 dB; remember that this assumes that each resistor is exactly the value it

www.raguvaran.puzl.com

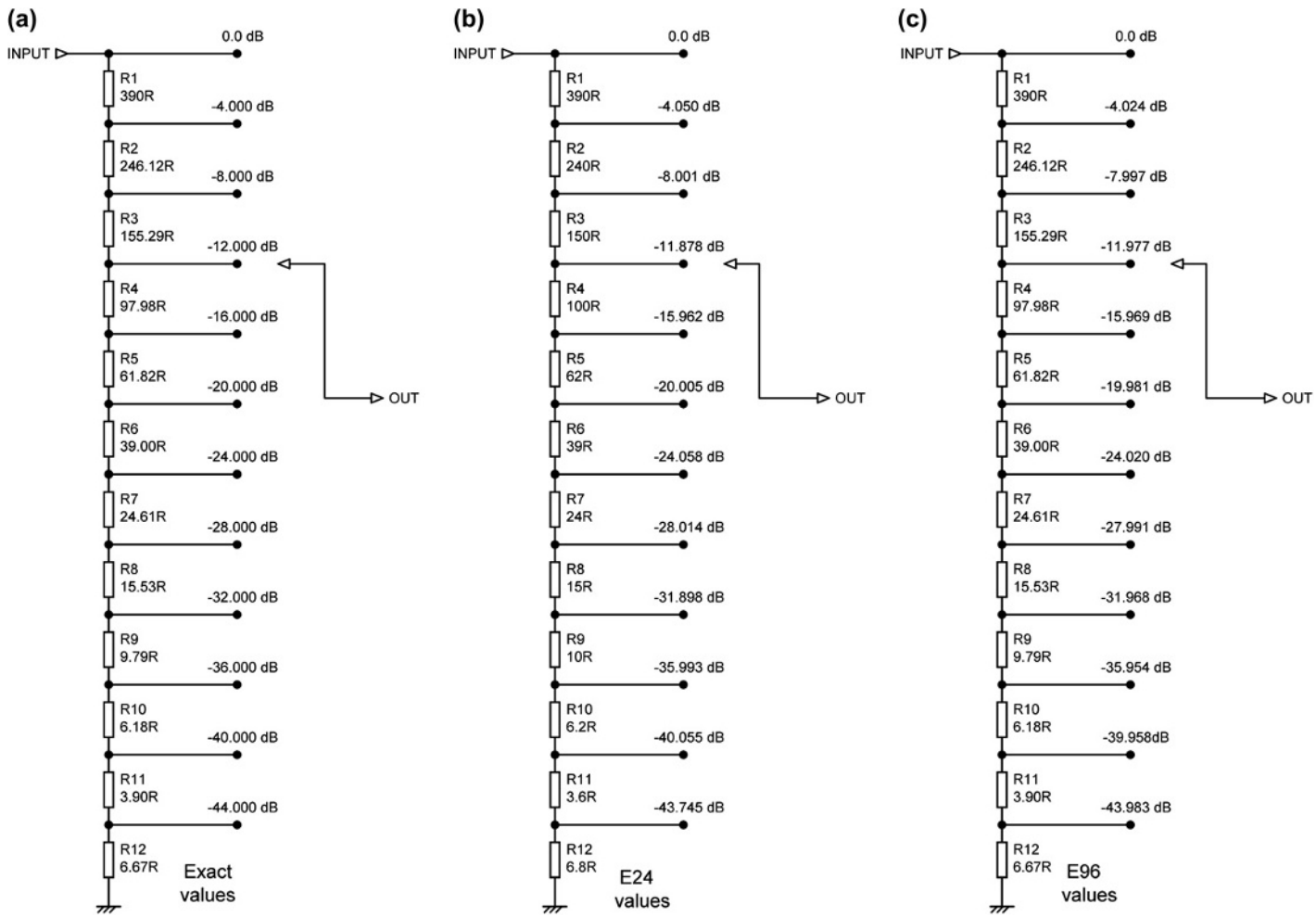


Figure 9.23: A 12-way switched attenuator volume control. (a) gives the theoretical resistor values, (b) is the best accuracy with E24 resistor values, and (c) shows how the step errors are much reduced by using E96 values

www.raguvaran.puzl.com

should be and does not incorporate any tolerances. The improvement in accuracy is not enormous, and if you can get a tighter tolerance in the E24 range than E96, E24 is the preferred option.

There is also an E192 resistor range, but it is rather rare, and there seems to be no pressing need to use so many different values in volume-control attenuators.

Almost all the work in the design of a switched attenuator is the calculation of the resistor values in the ladder. A spreadsheet model is the best way to do this. Putting in likely component values and attempting to tweak them by hand is a most unpromising approach, because all the values interact and you will boil off your sanity. A systematic approach using a spreadsheet is the best way. This is probably the simplest:

1. Decide the approximate total resistance of the resistor ladder.
2. Decide the step size, say 4 dB.
3. Take a two-resistor potential divider, where the sum of the resistors equals the desired total resistance, and choose exact values for a 4 dB attenuation; work back by trial and error to get an exact E24 value for the top resistor. (An exact E24 value at the top of the ladder is not essential, but it is convenient.)
4. Now split the bottom resistor into two, so that another 4 dB attenuator results. Check that the attenuation really is correct, because an error will propagate through the rest of the process, and you will have to go back and do it again from that point. Repeat this step until you have enough resistors in the ladder for the number of taps required.
5. When the table of resistor values is complete, for each resistor pick the nearest value from the E-series you are using.

You will then have constructed something like the spreadsheet shown in Table 9.4, which gives the E24 resistor values shown in Figure 9.23(b). The chosen value at the top of the ladder is 390 Ω. The spreadsheet has been set up to give more information than just the resistor value and the tap attenuation; it gives the voltage at each tap for a given input (10 V in this case) in column 5, the output impedance of each tap in column 6, the step size in dB in column 8, and the absolute error for each tap in column 9. It also gives the total resistance of the ladder, the current through it for the specified input voltage, and the resulting power dissipation in each resistor in column 7. The last parameter is unlikely to be a major concern in an audio attenuator, but if you're working with very low impedances to minimize noise, it is worth keeping an eye on.

Before the design work begins you must consider the stages before and after the attenuator. It is highly recommended that the attenuator input is driven from a very low impedance such as the output of an op-amp with plenty of negative feedback, so the source impedance can be effectively considered as zero and does not enter into the calculations. The loading on the output of the

www.raguvaran.puzl.com

TABLE 9.4 A spreadsheet that gives all the relevant information about a divider ladder for a switched attenuator

1	2	3	4	5	6	7	8	9	10
		Input = 10 V			Divider current = 0.0096		Amps		
		dB ref. = 0 V		V	Total res. = 1046.6				
Nearest E24 value	Ohms	ratio	(dB)	Output (V)	Tap Z Ohms	Power (mW)	Step (dB)	Tap error (dB)	Tap no.
R1	390					35.6			1
		0.62736	−4.050	6.2736	244.672		4.0496	−0.0496	2
R2	240					21.9			
		0.39805	−8.001	3.9805	250.77		3.952	−0.0012	3
R3	150					13.7			
		0.25473	−11.878	2.5473	198.69		3.877	0.1216	4
R4	100					9.1			
		0.15918	−15.962	1.5918	140.08		4.084	0.0379	5
R5	62					5.7			
		0.09994	−20.005	0.9994	94.15		4.043	−0.0050	6
R6	39					3.6			
		0.06268	−24.058	0.6268	61.49		4.053	−0.0575	7
R7	24					2.2			
		0.03975	−28.014	0.3975	39.95		3.956	−0.0137	8
R8	15					1.4			
		0.02542	−31.898	0.2542	25.92		3.884	0.1020	9
R9	10					0.9			
		0.01586	−35.993	0.1586	16.34		4.095	0.0065	10
R10	6.2					0.6			
		0.00994	−40.055	0.0994	10.30		4.061	−0.0549	11
R11	3.6					0.3			
		0.00650	−43.745	0.0650	6.76		3.690	0.2546	12
R12	6.8					0.6			

attenuator is more of a problem. You can either take loading on the output into account in the calculations, in which case the shunting effect of the load must be incorporated into step 3 above, or else make the input impedance of the next stage so high that it has a negligible effect.

As an example, the E24 network in Figure 9.23(b) was calculated with no allowance for loading on the output. Its highest output impedance is 253 Ω at tap 3, so this is the worst case for both loading sensitivity and noise. If a load of 100 kΩ is added, the level at this tap is only pulled down by 0.022 dB. A 100 kΩ input impedance for a following stage is easy to arrange, so the extra computation required in allowing for loading is probably not worthwhile unless for some very good reason the loading is much heavier. The Johnson noise of 253 Ω is still only −128.2 dBu.

www.raguvaran.puzl.com

TABLE 9.5 Resistor values and accuracy for a 2-dB step switched attenuator

		Switched attenuator: 2-dB steps							
		Exact		Using E24 values			Using E96 values		
Tap		Ohms	Step (dB)	Ohms	Step (dB)	Error (dB)	Ohms	Step (dB)	Error (dB)
1	R1	220.0000	0.0000	220	0.0000	0.0000	220	0.0000	0.0000
2	R2	174.7490	−2.0000	180	−2.0435	−0.0435	174	−2.0208	−0.0208
3	R3	138.8080	−4.0000	130	−4.1427	−0.1427	137	−4.0114	−0.0114
4	R4	110.2591	−6.0000	110	−6.0660	−0.0660	110	−5.9942	0.0058
5	R5	87.5819	−8.0000	82	−8.1104	−0.1104	86.6	−7.9967	0.0033
6	R6	69.5688	−10.0000	68	−10.0251	−0.0251	69.8	−9.9799	0.0201
7	R7	55.2605	−12.0000	56	−12.0124	−0.0124	54.9	−11.9914	0.0086
8	R8	43.8950	−14.0000	43	−14.0788	−0.0788	44.2	−13.9838	0.0162
9	R9	34.8670	−16.0000	33	−16.0850	−0.0850	34.8	−16.0046	−0.0046
10	R10	27.6958	−18.0000	27	−18.0166	−0.0166	27.4	−18.0107	−0.0107
11	R11	22.0000	−20.0000	22	−19.9959	0.0041	22.0	−19.9987	0.0013
12	R12	17.4749	−22.0000	18	−22.0272	−0.0272	17.4	−22.0077	−0.0077
13	R13	13.8808	−24.0000	13	−24.1361	−0.1361	13.7	−24.0093	−0.0093
14	R14	11.0259	−26.0000	11	−26.0577	−0.0577	11.0	−25.9916	0.0084
15	R15	8.7582	−28.0000	8.2	−28.1000	−0.1000	8.66	−27.9934	0.0066
16	R16	6.9569	−30.0001	6.8	−30.0120	−0.0120	6.98	−29.9758	0.0242
17	R17	5.5260	−32.0001	5.6	−31.9960	0.0040	5.49	−31.9863	0.0137
18	R18	4.3895	−34.0001	4.3	−34.0580	−0.0580	4.42	−33.9773	0.0227
19	R19	3.4867	−36.0001	3.3	−36.0588	−0.0588	3.48	−35.9964	0.0036
20	R20	2.7696	−38.0001	2.7	−37.9839	0.0161	2.74	−38.0004	−0.0004
21	R21	2.1999	−40.0001	2.2	−39.9548	0.0452	2.20	−39.9857	0.0143
22	R22	1.7475	−42.0000	1.8	−41.9753	0.0247	1.74	−41.9913	0.0087
23	R23	1.3881	−44.0000	1.5	−44.0701	−0.0701	1.37	−43.9886	0.0114
24	R24	5.3609	−46.0001	5.1	−46.3095	−0.3095	5.36	−45.9657	0.0343
				Total res. = 1055 Ω			Total res. = 1065 Ω		
				Avg abs. error = 0.0654 dB			Avg abs. error = 0.0117 dB		

If you feel you can afford 24-way switches, then there is rather more flexibility in design. You could cover from 0 to −46 dB in 2-dB steps, 0 to −57.5 dB in 2.5-dB steps, or 0 to −69 dB in 3-dB steps. There are almost infinite possibilities for adopting varying step sizes.

Table 9.5 shows the resistor values for 0 to −46 dB in 2-dB steps, Table 9.6 gives those for 0 to −57.5 dB in 2.5-dB steps, and Table 9.7 gives the values for 0 to −69 dB in 3-dB steps. All

www.raguvaran.puzl.com

TABLE 9.6 Resistor values and accuracy for a 2.5-dB step switched attenuator

		Switched attenuator: 2.5-dB steps							
		Exact		Using E24 values			Using E96 values		
Tap		Ohms	Step (dB)	Ohms	Step (dB)	Error (dB)	Ohms	Step (dB)	Error (dB)
1	R1	220.0000	0.0000	220	0.0000	0.0000	220	0.0000	0.0000
2	R2	164.9809	−2.4999	160	−2.5471	−0.0471	165	−2.4997	0.0003
3	R3	123.7182	−4.9999	120	−5.0208	−0.0208	124	−4.9907	0.0093
4	R4	92.7756	−7.4999	91	−7.4864	0.0136	93.1	−7.4901	0.0099
5	R5	69.5719	−10.0000	68	−9.9726	0.0274	69.8	−9.9928	0.0072
6	R6	52.1715	−12.5000	51	−12.4441	0.0559	52.3	−12.4958	0.0042
7	R7	39.1231	−15.0000	39	−14.9066	0.0934	39.2	−14.9974	0.0026
8	R8	29.3382	−17.5000	30	−17.4130	0.0870	29.4	−17.4982	0.0018
9	R9	22.0005	−20.0000	22	−19.9969	0.0031	22	−19.9995	0.0005
10	R10	16.4981	−22.5000	16	−22.5429	−0.0429	16.5	−22.4952	0.0048
11	R11	12.3718	−25.0000	12	−25.0153	−0.0153	12.4	−24.9899	0.0101
12	R12	9.2775	−27.5000	9.1	−27.4790	0.0210	9.31	−27.4890	0.0110
13	R13	6.9572	−30.0000	6.8	−29.9627	0.0373	6.98	−29.9914	0.0086
14	R14	5.2172	−32.5000	5.1	−32.4310	0.0690	5.23	−32.4939	0.0061
15	R15	3.9123	−35.0000	3.9	−34.8893	0.1107	3.92	−34.9949	0.0051
16	R16	2.9338	−37.5000	3.0	−37.3900	0.1100	2.94	−37.4948	0.0052
17	R17	2.2000	−40.0000	2.2	−39.9658	0.0342	2.2	−39.9950	0.0050
18	R18	1.6498	−42.5000	1.6	−42.5014	−0.0014	1.65	−42.4892	0.0108
19	R19	1.2372	−45.0000	1.2	−44.9601	0.0399	1.24	−44.9818	0.0182
20	R20	0.9277	−47.5001	0.91	−47.4057	0.0943	0.931	−47.4783	0.0217
21	R21	0.6957	−49.9999	0.68	−49.8653	0.1347	0.698	−49.9771	0.0229
22	R22	0.5217	−52.4999	0.51	−52.3019	0.1981	0.523	−52.4748	0.0252
23	R23	0.3912	−54.9997	0.39	−54.7183	0.2817	0.392	−54.9694	0.0306
24	R24	1.1731	−57.4997	1.20	−57.1626	0.3374	1.18	−57.4608	0.0392
				Total res. = 866 Ω			Total res. = 881 Ω		
				Avg abs. error = 0.0816 dB			Avg abs. error = 0.0113 dB		

three tables give the nearest E24 and E96 values, and the resulting errors. Note the first two versions start off with a 220 Ω resistor at the top, but on moving from 2- to 2.5-dB steps, the resistors toward the bottom naturally get smaller to give the greater attenuation required, and the total ladder resistance falls from 1060 to 866 Ω. To prevent the resistor values becoming inconveniently small, the 3-dB step version starts off with a higher value resistor of 430 Ω at the top; this increases the total resistance of the ladder to 1470 Ω and raises the maximum

www.raguvaran.puzl.com

TABLE 9.7 Resistor values and accuracy for a 3-dB step switched attenuator

		Switched attenuator: 3-dB steps							
		Exact		Using E24 values			Using E96 values		
Tap		Ohms	Step (dB)	Ohms	Step (dB)	Error (dB)	Ohms	Step (dB)	Error (dB)
1	R1	430.0000	0.0000	430	0.0000	0.0000	430	0.0000	0.0000
2	R2	304.4210	−3.0000	300	−2.9981	0.0019	301	−3.0066	−0.0066
3	R3	215.5135	−6.0000	220	−5.9437	0.0563	215	−5.9727	0.0273
4	R4	152.5719	−9.0000	150	−8.9930	0.0070	154	−8.9583	0.0417
5	R5	108.0126	−12.0000	110	−11.9281	0.0719	107	−11.9800	0.0200
6	R6	76.4671	−15.0000	75	−14.9622	0.0378	76.8	−14.9437	0.0563
7	R7	54.1346	−18.0000	56	−17.8769	0.1231	54.9	−17.9414	0.0586
8	R8	38.3243	−21.0000	39	−20.9466	0.0534	38.3	−20.9732	0.0268
9	R9	27.1315	−24.0000	27	−23.9857	0.0143	27.4	−23.9652	0.0348
10	R10	19.2077	−27.0000	20	−26.9606	0.0394	19.1	−26.9917	0.0083
11	R11	13.5980	−30.0000	13	−30.0906	−0.0906	13.7	−29.9736	0.0264
12	R12	9.6266	−33.0000	10	−32.9681	0.0319	9.53	−32.9950	0.0050
13	R13	6.8151	−36.0000	6.8	−36.0924	−0.0924	6.81	−35.9624	0.0376
14	R14	4.8248	−38.9999	4.7	−39.1209	−0.1209	4.87	−38.9497	0.0503
15	R15	3.4157	−42.0000	3.3	−42.0755	−0.0755	3.40	−41.9678	0.0322
16	R16	2.4181	−45.0000	2.4	−44.9831	0.0169	2.43	−44.9436	0.0564
17	R17	1.7119	−48.0000	1.8	−47.9476	0.0524	1.74	−47.9433	0.0567
18	R18	1.2119	−51.0001	1.2	−51.1090	−0.1090	1.21	−50.9841	0.0159
19	R19	0.8580	−54.0001	0.82	−54.1167	−0.1167	0.866	−53.9772	0.0228
20	R20	0.6074	−57.0002	0.62	−57.0034	−0.0034	0.604	−57.0068	−0.0068
21	R21	0.4300	−60.0003	0.43	−60.0776	−0.0776	0.432	−59.9949	0.0051
22	R22	0.3044	−63.0003	0.30	−63.1079	−0.1079	0.301	−63.0148	−0.0148
23	R23	0.2155	−66.0003	0.22	−66.0982	−0.0982	0.215	−65.9860	0.0140
24	R24	0.5224	−69.0003	0.51	−69.2132	−0.2132	0.523	−68.9771	0.0229
				Total res. = 1473 Ω			Total res. = 1470 Ω		
				Avg abs. error = 0.0701 dB			Avg abs. error = 0.0281 dB		

output impedance to 368 Ω at tap 3. The Johnson noise of 368 Ω is naturally higher at −126.5 dBu, but this is still very low compared with the likely noise from amplifier stages downstream.

The exact values on the left-hand side of each table can be scaled to give the total divider resistance required, but it will then be necessary to select the nearest E24 or E96 value manually. If you are increasing all the resistances by a factor of 10, then the same E24 or E96

www.raguvaran.puzl.com

values with a zero added can be used. The average of the absolute error for all 24 steps is shown at the bottom of each section; note that simply taking the average error would give a misleadingly optimistic result because the errors are of random sign and would partially cancel.

If 24 volume steps are not enough for you, then a different approach is required. Using relays, a microcontroller and an inexpensive shaft encoder, it is relatively simple to come up with a system that emulates a rotary switch with as many ways as desired, but you will need an awful lot of relays. Another approach is to use relays in a ladder attenuator, which greatly reduces the relay count. This is dealt with in the next section. As soon as a microcontroller is introduced, then there is of course the possibility of infrared remote control.

Another solution is to use two rotary switches: one for coarse volume control and the other for vernier volume control, the latter perhaps in 1-dB steps. While this is a workable approach, it is not exactly user-friendly. A buffer stage is usually desirable to prevent the second attenuator from loading the first one; the second attenuator should present a constant load, but it is a heavy one. The total loading of the two attenuators could also present an excessive load to the stage before the first attenuator.

A rotary switch does not of course have the smooth feel of a good potentiometer. To mitigate this, a switched volume control may need a large-diameter weighted knob, possibly with some sort of silicone damping. Sharp detents and a small knob do not a good volume control make.

Relay-Switched Volume Controls

The ultimate development in stepped volume controls is the relay ladder attenuator. This allows any number of steps to be employed.

The design of a good relay volume control is not as simple as it may appear. If you try to use a binary system, perhaps based on an R–2R network, you will quickly find that horrible transients erupt on moving from, say, 011111 to 100000. This is because relays have an operating time that is both long by perceptual standards, and somewhat unpredictable. It is necessary to use logarithmic resistor networks with some quite subtle relay timing.

Currently the highest expression of this technology is the Cambridge Audio 840A Integrated amplifier and 840E preamplifier. The latter uses second-generation relay volume control technology. A rather unexpected feature of relay volume controls is that having purchased a very sophisticated preamplifier, where the precision relay control is the major unique feature, some customers then object to the sound of the relays clicking inside the box when they adjust the volume setting. You just can't please some people.

www.raguvaran.puzl.com

Transformer-Tap Volume Controls

There are other ways of controlling volume than with resistors. At the time of writing (2009) there is at least one passive preamplifier on the market that controls volume by changing the taps on the secondary of a transformer. This should at least give a low output impedance, but there are several potential problems with this idiosyncratic approach – transformers are well known to fall much further short of being an ideal component than most electronic parts do. They can introduce frequency response irregularities, LF distortion, and hum. They are relatively heavy and expensive, and the need for a large number of taps on the secondary puts the price up further. A multi-way switch to select the desired tap must also be paid for.

The unit I have in mind has only 12 steps, apparently each of 4 dB. These are very coarse steps for a volume control – unacceptably so, I would have thought. Twenty-four-way switches are available (see the section on switched attenuator volume controls above) but they are expensive, and there might be some interesting challenges in fitting 24 taps into a transformer design.

Integrated Circuit Volume Controls

Specialized volume-control ICs based on switched networks have been around for a long time, but have only recently reached a level of development where they can be used in high-quality audio. Early versions had problems with excessively large volume steps, poor distortion performance, and nasty glitching on step changing. The best contemporary ICs are greatly improved; a modern example is the PGA2310 from Burr-Brown (Texas Instruments), which offers two independent channels with a gain range from +31.5 to −95.5 dB in 0.5-dB steps, and is spec'd at 0.0004% THD at 1 kHz for a 3 Vrms input level. The IC includes a zero crossing detection function intended to give noise-free level transitions. The concept is to change gain settings on a zero crossing of the input signal, and so minimize audible glitches. Gain control is by means of a three-wire serial interface, which is normally driven by a rotary encoder and a microcontroller such as one of the PIC series.

The Yamaha YAC523 has a similar specification but incorporates seven gain-controlled channels for AV applications.

Balance Controls: Passive

The balance control on a preamplifier adjusts the relative gain of the left and right channels, in order to move the stereo stage to the left or right. As you would expect, increasing the left gain moves things to the left. A balance control does not have to make radical changes to signal level to do its job. Introducing a channel gain imbalance of 10 dB is quite enough to shift the sound image completely to one side, so it appears to be coming from one loudspeaker only,

www.raguvaran.puzl.com

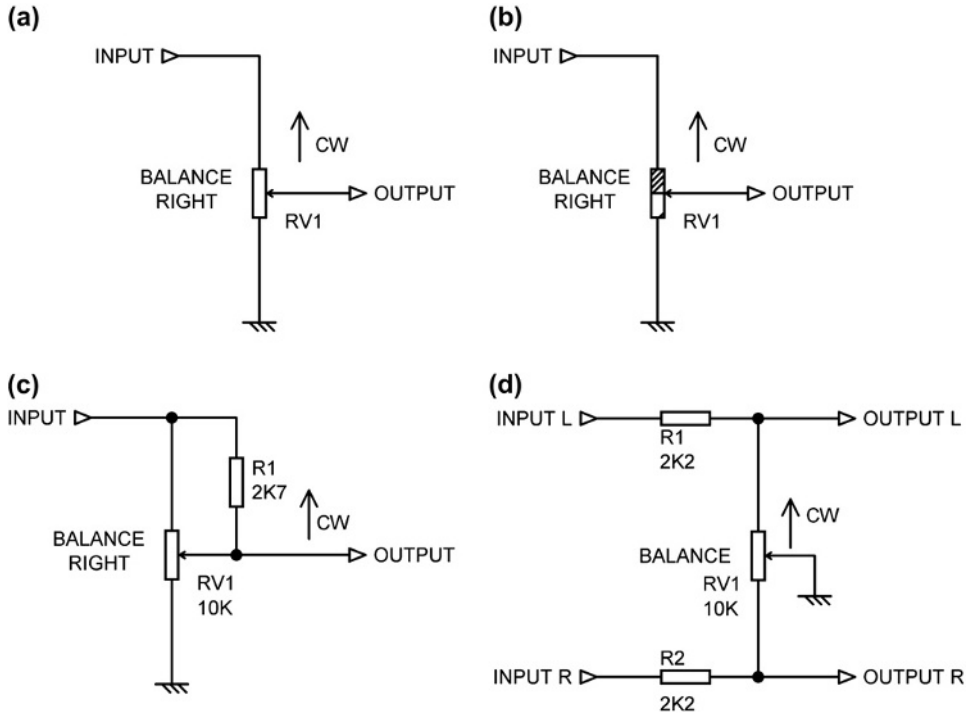


Figure 9.24: Passive balance controls

and there is nothing to be gained by having the ability to fade out one channel completely. A central detent on the control is highly desirable.

The ideal hi-fi balance control is *not* a mixer panpot (see Chapter 17 on mixer subsystems). Since the balance control is usually a set-and-forget function that does not require readjustment unless the listening room is rearranged, there is no requirement that the overall volume remains exactly constant when the balance is adjusted. The ideal hi-fi balance control will have no attenuation when set centrally, and when moved to left or right will attenuate only one channel without affecting the gain of the other; to implement this passively in a pot requires a special design, but this is not a problem as they are available from several manufacturers. If there *is* attenuation at the central position then it needs to be made up by extra amplification either before or after the balance control, and this means that either the overload margin or the noise performance will be compromised to some degree.

Figure 9.24 shows the various forms of passive balance control; only the right channel is shown in the first three cases. Anticlockwise movement of the control is required to introduce attenuation into the right channel and shift the sound image to the left. In Figure 9.24(a) a simple pot is used. If this has a linear track it will give a 6 dB loss when set centrally – this

www.raguvaran.puzl.com

means that an extra 6 dB of gain has to be built into the system, and the noise/headroom compromise is too great. If a log law is used for one channel, and an anti-log law for the other, the central loss can be made smaller, say 3 dB or less.

However, the uncertainties of log-law tolerancing means it will not be possible to guarantee that the channel gains are identical when the control is centralized. The ultimate expression of this is the special balance pot shown in Figure 9.24(b), where the top half of the track is made of a low-resistance material, so that neither channel is attenuated at the central position. On moving the control clockwise the right channel wiper stays on the low-resistance section and gain stays at 0 dB, while the left wiper moves on to the normal log/anti-log section and the signal is attenuated. Another balance control option is shown in Figure 9.24(c), where a linear pot is used with a pull-up resistor so the attenuation at the central position is less than 6 dB. With the values shown the loss for each channel with the control central is −2.6 dB. More information on what is essentially a panpot configuration (though not in this case a panpot law) can be found in Chapter 17.

The economical method of Figure 9.24(d) was once popular as only a single pot section is required. Unfortunately it has the snag that the relatively high resistance between wiper and track causes serious degradation of the interchannel crosstalk performance. If the resistance values are reduced to lower Johnson noise, the track-wiper resistance is unlikely to decrease proportionally and the crosstalk will be worsened. With the values shown the loss for each channel with the control central is −3.2 dB. This loss can be reduced by decreasing the value of R1, R2 with respect to the pot, but this puts a correspondingly heavier load on the preceding stages when the control is well away from central.

Some preamplifier designs have attempted to evade the whole balance control problem by having separate but concentric volume knobs for left and right channels. The difficulty here is that almost all the time the volume only will require adjustment, and the balance function will be rarely used; it is therefore highly desirable that the left and right knobs are linked together in some sort of high-friction way so that the two normally move together. This introduces some awkward mechanical complications.

Active Balance Controls

An active balance control is configured so that it makes a small adjustment to the gain of each channel rather than introducing attenuation, so the noise/headroom compromise can be avoided altogether. Since all active preamplifiers have at least one gain stage, the extra complication is likely to be minimal. It is not desirable to add an extra active stage just to implement the balance function.

Figure 9.25(a) shows an active balance control that requires only a single-gang pot. However, it suffers from the same serious disadvantage as the passive version in

www.raguvaran.puzl.com

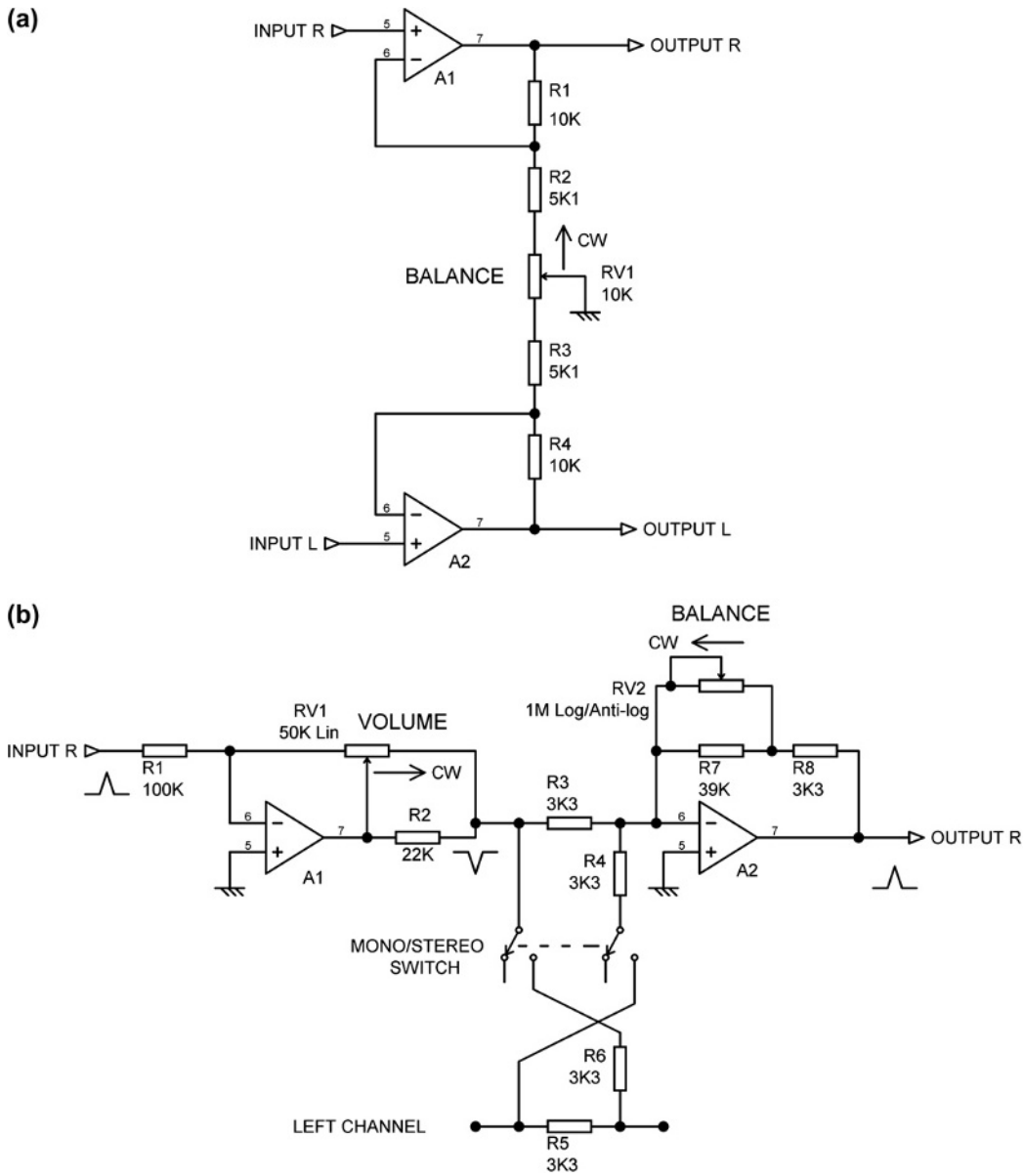


Figure 9.25: Active balance controls

Figure 9.24(d): the wiper connection acts as a common impedance in the two channels and causes crosstalk. This kind of balance control cannot completely fade out one channel as it is not possible to reduce the stage gain below unity, and in fact even unity cannot be achieved with this configuration. With the values shown the gain for each channel with the control central is +6.0 dB. With the control fully clockwise the gain increases to +9.4 dB,

www.raguvaran.puzl.com

and decreases to +4.4 dB with it fully anticlockwise; the range is deliberately quite restricted. It is a characteristic of this arrangement that the gain increase on one channel is greater than the decrease on the other.

Figure 9.25(b) shows an active balance control combined with an active gain control and mono/stereo switching. This is the configuration shown above in Figure 9.16(d), and was used in the original Cambridge Audio P-series amplifiers; there A1 was a simple two-transistor inverting stage and A2 an even simpler single transistor. The left-hand section of volume control RV1 is the feedback resistance for A1, while the right-hand section forms part of the input resistance to shunt stage A2, both changing to give a quasi-logarithmic law when the control is altered. The balance control RV2 is a variable resistance in the shunt-feedback network of A2. The mono/stereo switch feeds the virtual-earth node of A2 with both channels via R3, R4 when in mono mode. The circuit handily avoids phase inversion of the output.

Mono/Stereo Switches

It was once commonplace for preamplifiers to have mono/stereo switches, which allowed a mono source to be played over both channels of a stereo amplifier system. Some of these were configured so that the two channels were simply joined together somewhere in the middle of the preamp stages, which was not very satisfactory unless the unused input was terminated in a low impedance to minimize noise. More sophisticated versions allowed either the left or the right input to be routed to both outputs.

Since all modern sources are at least stereo, mono/stereo switches are now rarely if ever fitted.

Width Controls

Another facility that has become rare is the width control. Summing a small proportion of each channel into the other reduces the width of the sound image, and this was sometimes advocated as a small width reduction would make the image less associated with the loudspeakers, and so give a stronger illusion of acoustic reality. This of course runs directly counter to more contemporary views that very high levels of interchannel isolation are required to give a good stereo image. This is flat-out untrue; it was established long ago by the BBC, in extensive testing before the introduction of stereo broadcasting, that a stereo separation of 20–25 dB is enough to give the impression of full image width.

By cross-feeding anti-phase signals, the width of a stereo image can be increased. A famous circuit published by Mullard back in 1972 [4] gave continuous variation between mono, normal, and enhanced-width stereo. It was stated that anti-phase cross-feed of greater than 24% should not be used as it would cause the sound image to come apart into two halves.

www.raguvaran.puzl.com

References

[1] P. Baxandall, Audio gain controls, Wireless World (November 1980) pp. 79–81.
[2] D. Self, A precision preamplifier, Electronics World (October 1983).
[3] D. Self, Precision Preamplifier 96, Electronics World (July/August and September 1996).
[4] M.J. Rose (Ed.), Transistor Audio and Radio Circuits, second ed., Mullard, 1972, p. 180.

www.raguvaran.puzl.com

This page intentionally left blank

www.raguvaran.puzl.com

CHAPTER 10

Tone Controls and Equalizers

Facilities that alter the shape of the frequency response are called tone controls when they are incorporated in hi-fi systems, and equalization (or EQ) in mixing consoles.

Tone controls have suffered at the hands of fashion for some years now. It has been claimed that tone controls cause an audible deterioration even when set to the flat position. This is usually blamed on 'phase-shift'. For a long time tone controls on a preamp damaged its chances of street (or rather sitting room) credibility, for no good reason. A tone control set to 'flat' – assuming it really is flat – cannot possibly contribute any extra phase-shift unless you have accidentally built in an all-pass filter, which would require truly surreal levels of incompetence, and so the control really must be inaudible. My view is that hi-fi tone controls are absolutely indispensable for correcting room acoustics, loudspeaker shortcomings, or the tonal balance of the source material, and that a lot of people are suffering suboptimal sound as a result of this fashion. It is commonplace for audio critics to suggest that frequency-response inadequacies should be corrected by changing loudspeakers; this is an extraordinarily expensive way of avoiding tone controls.

The equalization sections of mixers have a rather different function, being creative rather than corrective (from now on I am going to just call it EQ). The aim is to produce a particular sound, and to this end mixer EQ is much more sophisticated than that found on most hi-fi preamplifiers. There will be middle controls as well as bass and treble (which in the mixing world are more often called LF and HF) and these introduce a peak or dip into the middle range of the audio band. On more complex consoles the middle frequencies are infinitely variable, and the most advanced examples have variable Q as well, to control the width of the peak or dip introduced. No one has so far suggested that mixing consoles should be built without EQ.

It is not necessary to litter these pages with equations to determine center frequencies and so on. In each case, altering the range of frequency controlled can be done very simply by scaling the capacitor values given. If a stage gives a peaking cut/boost at 1 kHz, but you want 2.5 kHz, then simply reduce the values of all the capacitors by a factor of 2.5 times.

Scaling the associated resistors instead would give the same frequency response, but may also affect noise performance, because if the resistor values are raised, Johnson noise will increase. Distortion will increase if reducing the resistor values places excessive loading on the op-amps

Small Signal Audio Design; ISBN: 9780240521770
Copyright © 2010 Elsevier Ltd; All rights of reproduction, in any form, reserved.

www.raguvaran.puzl.com

used. Another consideration is that potentiometers come in a very limited number of values, usually multiples of 1, 2, and 5. Changing the capacitors is simpler and more likely to be trouble free.

Passive Tone Controls

For many years all tone controls were passive, simply providing frequency-selective attenuation. These came in a bewildering variety of forms; take a look at the chapter on 'Tone compensation and tone control' in the famous *Radio Designer's Handbook* [1]. This is 42 pages long with 90 references. The final edition was published in 1953, and interestingly does *not* include the Baxandall configuration. Some of the circuits are incredibly complex, requiring multi-section switches and tapped inductors to give quite limited tone-control possibilities. Figure 10.1 shows one of the simpler arrangements [2], which is probably the best-known passive tone control configuration. It was described by Sterling [3], though I have no idea if he originally invented it.

The arrangement gives about ±18 dB of treble and bass boost and cut, the curves looking something like those of a Baxandall control but with less symmetry.

Such circuitry has several disadvantages. When set to flat it gives a loss in each network of 20.8 dB, which means a serious compromise in either noise (if the make-up gain is after the tone-control network) or headroom (if the make-up gain is before the tone-control network). In the days of valves, when these networks were popular, headroom may have been less of a problem, but given the generally poor linearity of valve circuitry, the increased level probably gave rise to significantly more distortion.

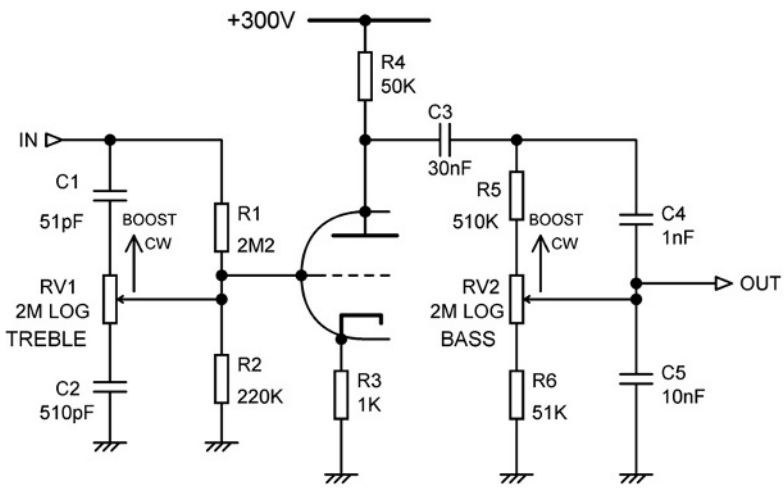


Figure 10.1: A pre-Baxandall passive tone control, with severe limitations

www.raguvaran.puzl.com

Another problem is that if linear pots are used the flat position corresponds to one-tenth of the rotation. It is therefore necessary to use log pots to get the flat setting to somewhere near the center of control travel, and their large tolerances in law and value mean that the flat position is actually rather variable, and is unlikely to be the same for the two channels of stereo.

This circuit counts as a passive tone control because the valve in the middle of it is simply providing make-up gain for the treble network and is not in a frequency-dependent feedback loop. In the published circuit there was another identical valve stage immediately after the bass network to make up those losses. There are some other interesting points about this valve-based circuit: it runs at a much higher impedance than solid-state versions, using 2 MΩ pots rather than 10 kΩ, and it uses a single supply rail, and that is at an intimidating 300 V. Circuitry running at such high impedances is susceptible to capacitative hum pickup.

Baxandall Tone Controls

The Baxandall bass-and-treble tone control swept all other versions before it. The original Baxandall tone control, as famously published in *Wireless World* in 1952 [4], was in fact rather more complex than the simplified version which has become universal. The original circuit required a center-tapped treble pot to give minimum interaction between the two controls, but such specialized components are almost as unwelcome to manufacturers as they are to home constructors, and the form of the circuit that became popular is shown in Figure 10.2, some control interaction being regarded as acceptable. The advantages of this circuit are its simplicity and its feedback operation, the latter meaning that there are no awkward compromises between noise and headroom, as there are in the passive circuit above. It also gives symmetrical cut and boost curves and is easily controlled.

It is not commonly realized that the Baxandall tone control comes in two versions. Either one or two capacitors can be used to define the bass time-constants, and the two arrangements give

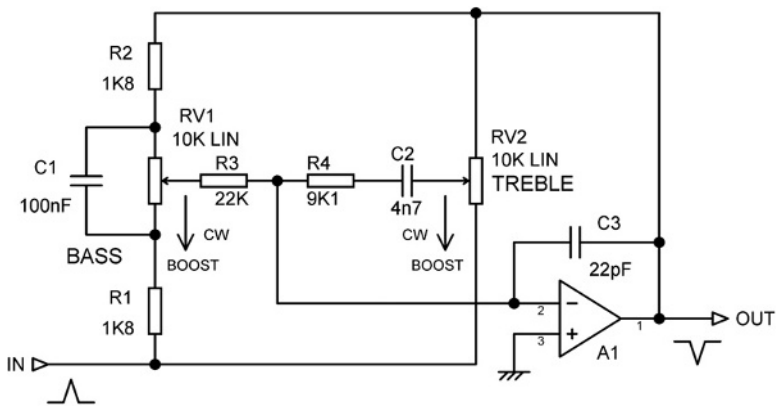


Figure 10.2: The one-capacitor Baxandall tone control

www.raguvaran.puzl.com

rather different results at the bass end. The original Baxandall design was the two-capacitor version.

In the descriptions that follow, I have used the term 'break frequency' to indicate where the tone control begins to take action. This is defined as the frequency where the response is ±1 dB away from flat with maximum cut or boost applied.

The Baxandall Two-Capacitor LF Control

This is probably the most common form of the Baxandall tone control, simply because it saves a capacitor. The circuit is shown in Figure 10.2. At high frequencies the impedance of C1 is small and the bass control RV1 is effectively shorted out and R1, R2 give unity gain. At low frequencies RV1 is active and controls the gain, ultimately over a ±15 dB range at very low frequencies. R1, R2 are then end-stop resistors that set the maximum boost or cut.

At high frequencies again, C2 has a low impedance and treble control RV2 is active, with maximum boost or cut set by end-stop resistor R4; at low frequencies C2 has a high impedance and so RV2 has no effect. Resistor R3 is chosen to minimize interaction between the controls.

The one-capacitor version is distinguished by its fixed LF break frequency, as shown by the bass control response in Figure 10.3. The treble control response in Figure 10.4 is similar.

In these figures the control travel is in 11 equal steps of a linear pot (central plus five steps on each side). The bass curves are ±1.0 dB at 845 Hz, while the treble curves are ±1.0 dB at 294 kHz. This sort of overlap is normal with the Baxandall configuration. Phase spikes are shown at input and output to underline that this stage phase-inverts, which can be inconvenient; phase spikes will be seen in most of the diagrams that follow in this chapter.

An HF-stabilizing capacitor C3 is shown connected around A1. This is often required to ensure HF stability at all control settings, depending on how much stray capacitance to introduce extra phase-shifts there is in the physical layout. The value required is best determined by experiment. This capacitor is not shown in most of the diagrams that follow, to keep them as uncluttered as possible, but the likely need for it should not be forgotten.

It is important to remember that the input impedance of this circuit varies both with frequency and control settings, and it can fall to rather low values.

Taking the circuit values shown in Figure 10.2, the input impedance varies with frequency as shown in Figure 10.5, for treble (HF) control settings, and as in Figure 10.6 for bass (LF) control settings. With both controls central the input impedance below 100 Hz is 2.9 kΩ. At low frequencies C1, C2 have no effect and the input impedance is therefore half the LF pot resistance plus the 1k8 end-stop resistor, adding up to 6.8 kΩ. In parallel is 5 kΩ, half the impedance of the HF pot, as although this has no direct connection to the summing point

www.raguvaran.puzl.com

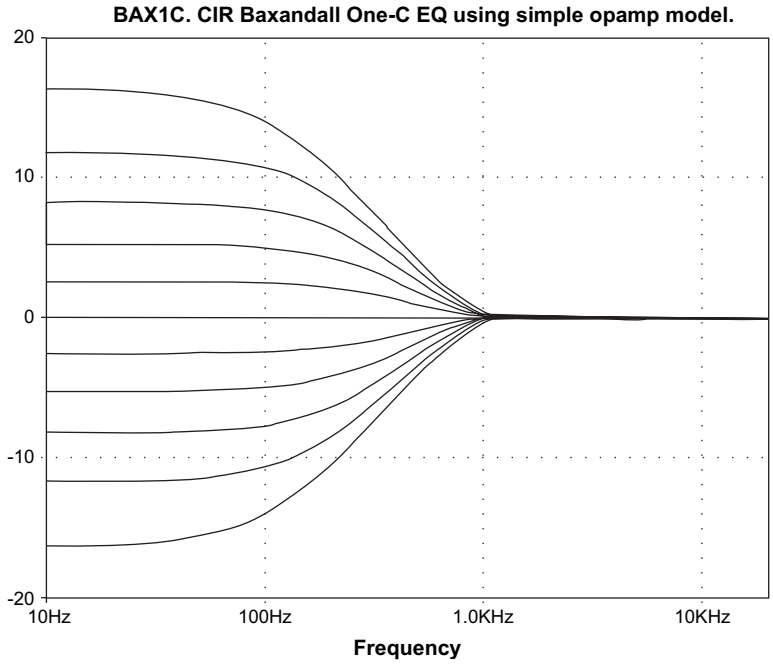


Figure 10.3: Bass control frequency responses. The effect of the LF control above the 'hinge point' at 1 kHz is very small

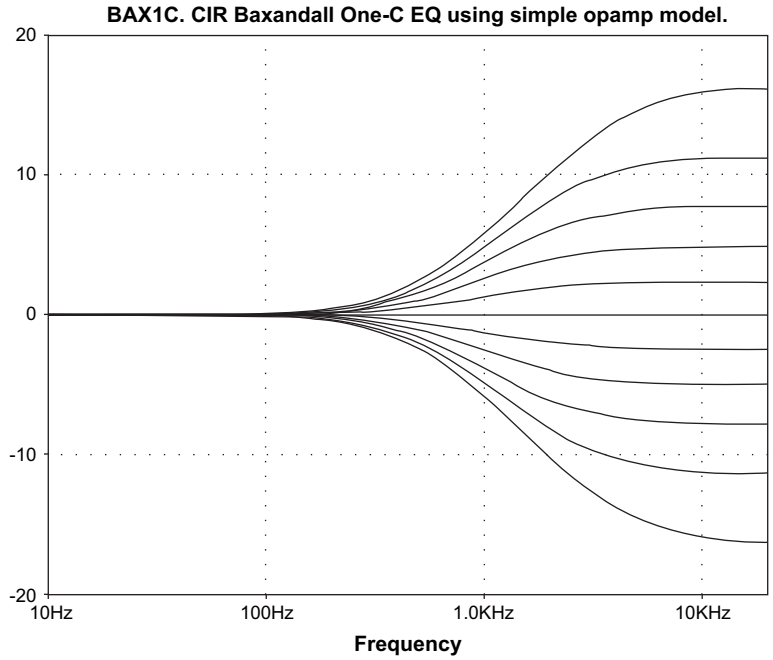


Figure 10.4: Treble control frequency responses

www.raguvaran.puzl.com

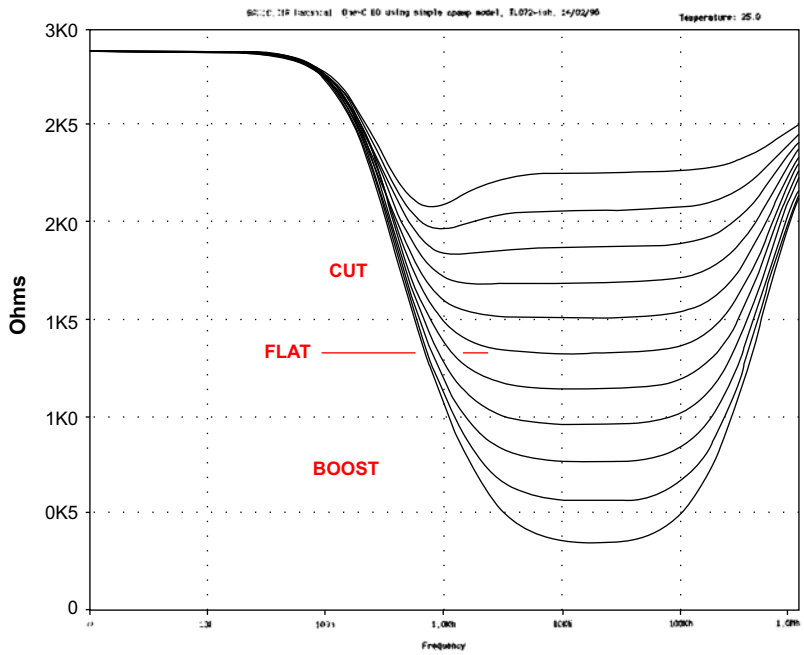


Figure 10.5: Input impedance variation with frequency, for 11 treble control settings

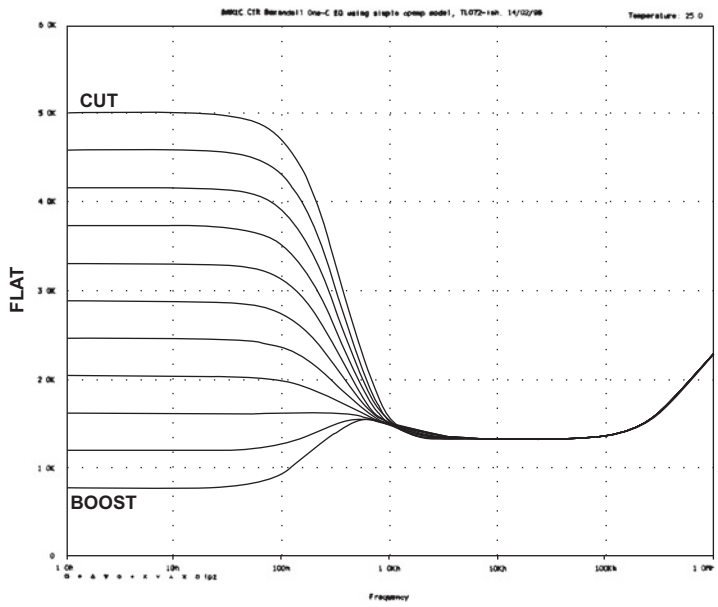


Figure 10.6: Input impedance variation with frequency, for 11 bass control settings

www.raguvaran.puzl.com

(C2 being effectively open circuit), its other end is connected to the output, which is the input inverted. Hence the center of the pot is approximately at virtual earth. The parallel combination of 6.8 and 5 kΩ is 2.9 kΩ.

At frequencies above 100 Hz (still with both controls central) the input impedance falls because C1 is now low impedance, and the LF pot is shorted out. The input impedance is now 1.8 kΩ in parallel with 5 kΩ, which is 1.3 kΩ. This is already about as much loading as you would want to put on a TL072, and we haven't applied boost or cut yet.

When the HF control is moved from its central position, the HF impedance is higher at full cut at 2.2 kΩ. It is, however, much lower at 350 Ω at full boost. There are very few op-amps that can give full output into such a low impedance, but this is not quite as serious a problem as it first appears. The input impedances are only low when the circuit is boosting; therefore driving the input at the full rail capability is not relevant, for if you do the output will clip long before the stage driving it. Nonetheless, op-amps such as the 5532 will show increased distortion driving too heavy a load, even if the level is a long way below clipping, so it is a point to watch.

The input impedances can of course be raised by scaling the impedance of the whole circuit. For example, multiply all resistor values by 4, and quarter the capacitor values to keep the frequency response the same. The downside to this is that you have doubled the Johnson noise from the resistors, and so made the stage noisier.

When the LF control is moved from its central position, the input impedance variations are similar, as shown in Figure 10.6. At full cut the input impedance is increased to 5.0 kΩ; at full boost it falls to 770 Ω. This variation begins below 1 kHz, but is only fully established below 100 Hz.

The Baxandall Two-Capacitor Tone Control

The two-capacitor version of the Baxandall tone control is shown in Figure 10.7, and while it looks very similar there is a big difference in the LF end response curves, as seen in Figure 10.8. The LF break frequency rises as the amount of cut or boost is increased. It is my view that this works much better in a hi-fi system, as it allows small amounts of bass boost to be used to correct loudspeaker deficiencies without affecting the whole of the bass region. In contrast, the one-capacitor version seems to be more popular in mixing consoles.

Note that the treble control here has been configured slightly differently, and there are now two end-stop resistors, at each end of the pot, rather than one attached to the wiper; the frequency response is identical, and this variation is only shown to make that point. An extra resistor is required without any corresponding benefit.

www.raguvaran.puzl.com

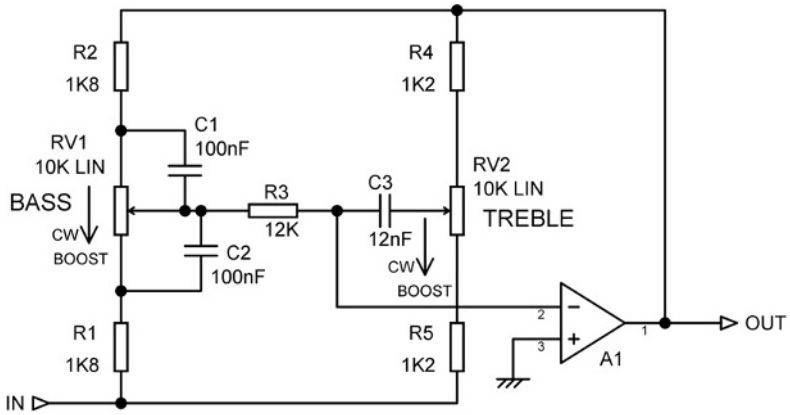


Figure 10.7: The two-capacitor Baxandall tone control

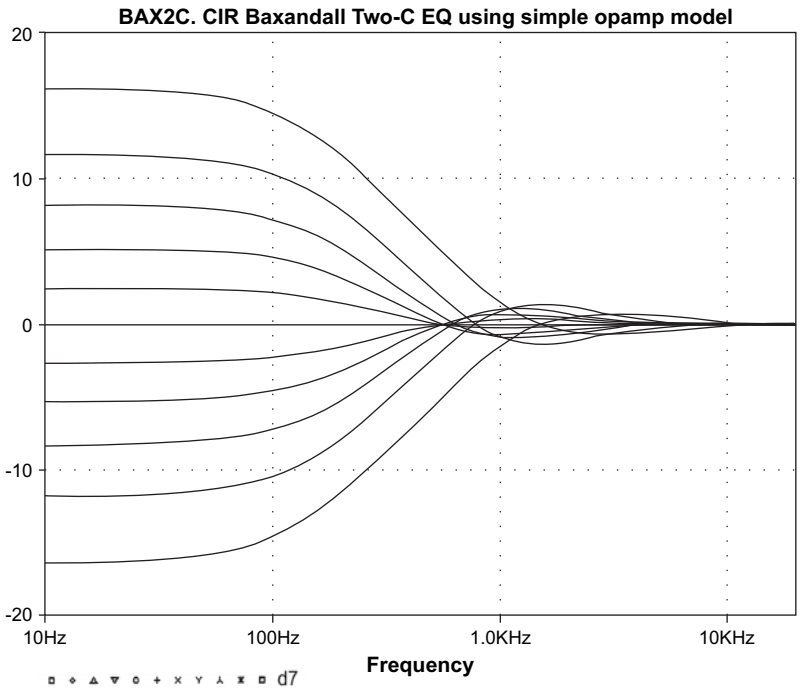


Figure 10.8: Bass control frequency responses for the two-capacitor circuit. Compare Figure 10.3

The input impedance of this version shows variations similar to those of the one-capacitor version. With controls central, at LF the input impedance is 3.2 kΩ; from 100 Hz to 1 kHz it slowly falls to 1.4 kΩ. The changes with HF and LF control settings are similar to those of the one-capacitor version, and at HF the impedance falls to 370 Ω.

www.raguvaran.puzl.com

The Baxandall Two-Capacitor HF Control

The treble control can also be implemented with two capacitors, as in Figure 10.9.

In this case the frequency response results are identical to those of the one-capacitor version, but there are interesting differences in the loading presented to the preceding stage. With controls central, at LF the input impedance is 6.8 kΩ, which is usefully higher; from 100 Hz to 1 kHz it slowly falls to the previous value of 1.4 kΩ.

When the LF control is varied, at full cut the input impedance is increased to 11.7 kΩ, and at full boost it falls to 1.9 kΩ (see Figure 10.10). On varying the HF control, at full cut the input impedance is increased to 2.3 kΩ; at full boost it falls to 420 Ω. These values are higher because with this configuration C3, C4 effectively disconnect the HF pot from the circuit at low frequencies. In some cases the higher input impedance may justify the cost of an extra capacitor.

One disadvantage of the Baxandall tone control is that it inherently phase-inverts. This is decidedly awkward, because relatively recently the hi-fi world has decided that absolute phase is important; in the recording world keeping the phase correct has always been a rigid requirement.

An important point about all of the circuits shown so far is that they assume an FET-input op-amp (such as the TL072, or a more sophisticated FET part) will be used to minimize the bias currents flowing. Therefore all the pots and switches are directly connected to the op-amp without any explicit provision for preventing DC flowing through the pots. Excessive DC would make the pots scratchy and crackly when they are moved; this does not sound nice. It is, however, long established that typical FET bias currents are low enough to prevent such effects

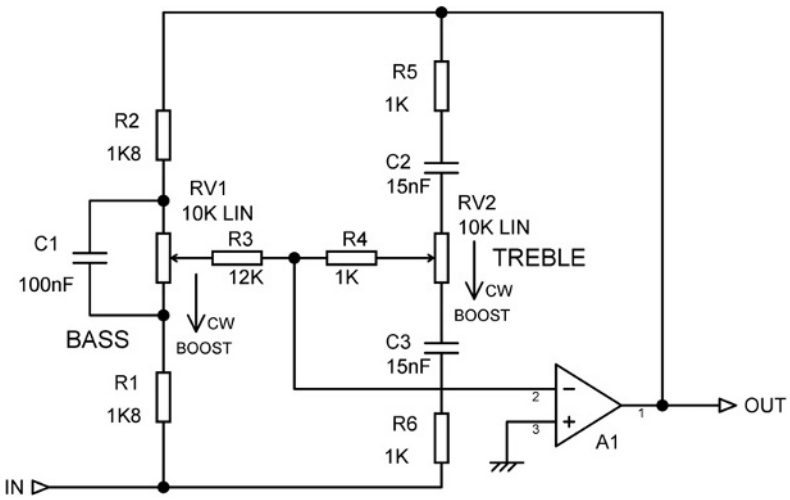


Figure 10.9: Circuit of the two-HF-capacitor version

www.raguvaran.puzl.com

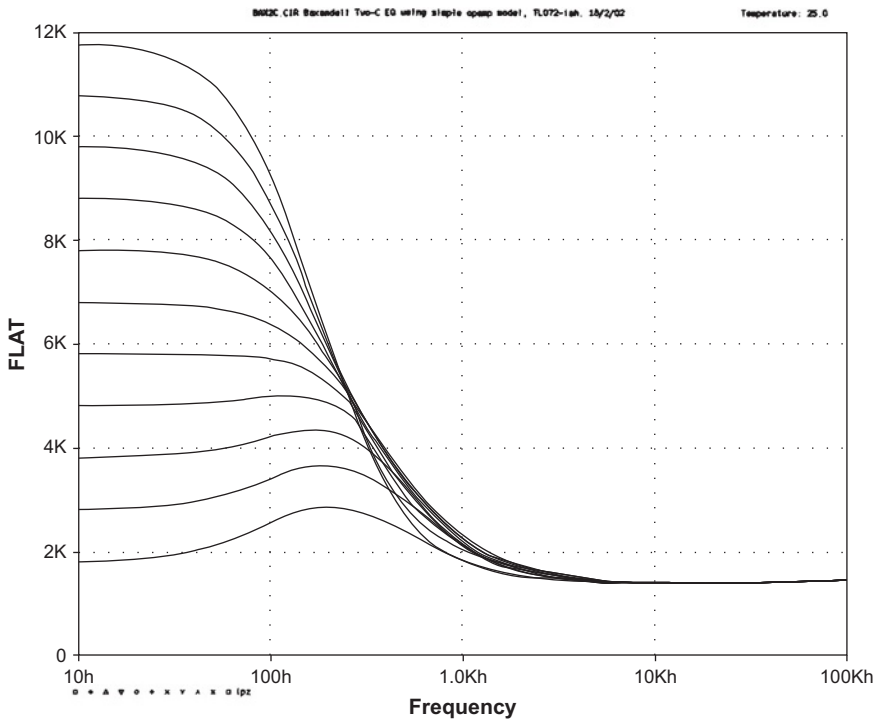


Figure 10.10: Input impedance variation with frequency, for 11 bass control settings; two-HF-capacitor version

in circuits like these. Substituting a bipolar op-amp such as the 5532 will improve the noise and distortion performance markedly, at the expense of the need to make provision for the much greater bias currents by adding DC-blocking capacitors.

Switched-Frequency Baxandall Controls

While the Baxandall approach gives about as much flexibility as one could hope for from two controls, there is often a need for more. The most obvious elaboration is to make the break frequencies variable.

This is straightforward if a small number of switched break frequencies meets the case. The frequency response of the Baxandall configuration is set by its RC time-constants, and one obvious way to make a variable-frequency version is to make the capacitors switchable. Changing the R part of the RC is far less practical as it would require changing the potentiometer values as well. Figure 10.11 shows a typical tone control with three switched treble frequencies, as used in a preamplifier design I did in 1976 [5]. The break frequencies are 1, 3, and 5 kHz. I found break frequencies higher than that to be a bit too subtle. Note that the circuit impedances are much higher than I would use nowadays.

www.raguvaran.puzl.com

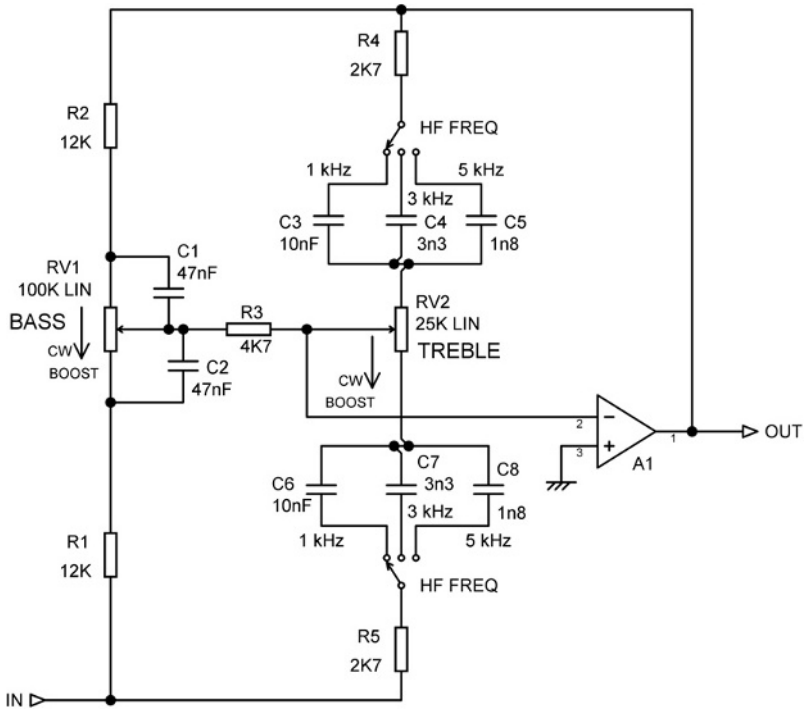


Figure 10.11: A Baxandall tone control with three switched treble turnover frequencies

A parameter switchable in steps is inconvenient, compared with having it infinitely variable. Also, multi-way switches are more expensive than variable resistors, and so there is a strong incentive to make the operating frequencies variable without switches and banks of capacitors. To do this requires a different configuration, as described in the next section.

Variable-Frequency LF and HF EQs

Since we are now moving more into the world of mixing consoles rather than preamplifiers, I shall stop using 'bass' and 'treble' and switch to 'LF' and 'HF', which of course means the same thing. The circuit shown in Figure 10.12 gives HF equalization only, but with a continuously variable break frequency, and is used in many mixer designs. It is similar to the Baxandall concept in that it uses op-amp A1 in a shunt-feedback mode so that it can provide either cut or boost, but the resemblance ends there.

R1 and R2 set the basic gain of the circuit to −1, and ensure that there is feedback at DC to establish the operating conditions; there is no DC feedback path to the non-inverting input of A1. When the wiper of RV1 is at the input end, extra input signal is fed forward and gain increases. When the wiper of RV1 is at the output end, more output is fed back, the total amount of NFB increases, and gain is reduced.

www.raguvaran.puzl.com

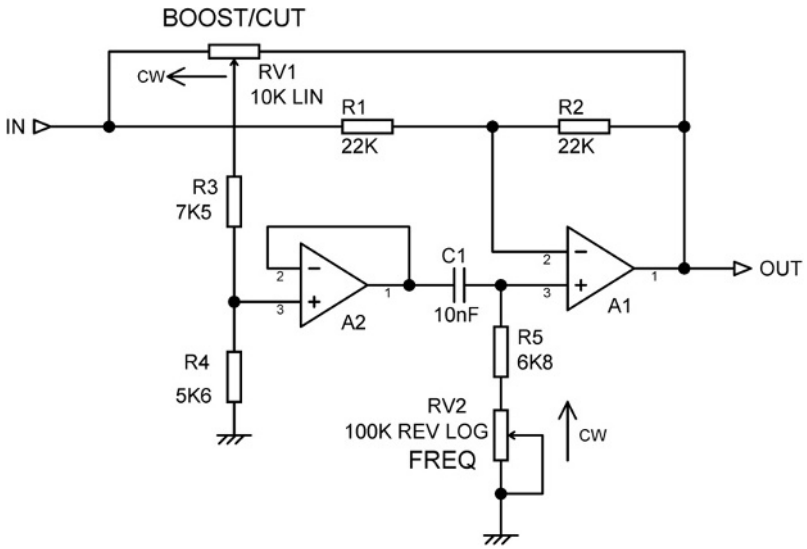


Figure 10.12: A variable-frequency HF shelving circuit. The break frequency range is 1–15 kHz

The signal tapped off is scaled by divider R3, R4, which set the maximum cut/boost. The signal is then buffered by voltage-follower A2, and fed to the frequency-sensitive part of the circuit, a high-pass RC network made up of C1 and (R5 + RV2). The frequency-setting control RV2 has a relatively high value at 100 kΩ. This is because the range of variation is usually at least 10:1, and a lower value would give excessive loading on A2 at the high-frequency end.

A2 prevents interaction between the amount of cut/boost and the break frequency. Without it cut/boost would be less at high frequencies because R5 would load the divider R1, R2. In cheaper products this interaction may be acceptable, and A2 could be omitted. In EQ circuitry it is the general rule that the price of complete freedom from control interaction is more op-amps. The frequency range can be scaled simply by altering the value of C1, as shown in Figure 10.13. The range of frequency variation is controlled by the value of end-stop resistor R5, subject to restrictions on loading A2 if distortion is to be kept low. The corresponding LF EQ is made by swapping the positions of C1 and R5, RV2 in Figure 10.12.

The component values given are the E24 values that give the closest approach to ±15 dB cut/boost at the control extremes. When TL072 op-amps are used this circuit is stable as shown, assuming the usual supply-rail decoupling. If other types are used a small capacitor (say 33 pF) across R2 may be required.

In Chapter 11 a tone control with variable HF and LF frequencies is examined as part of a complete preamplifier design. 5532 op-amps are used and the resulting need for DC-blocking capacitors is looked at in detail.

www.raguvaran.puzl.com

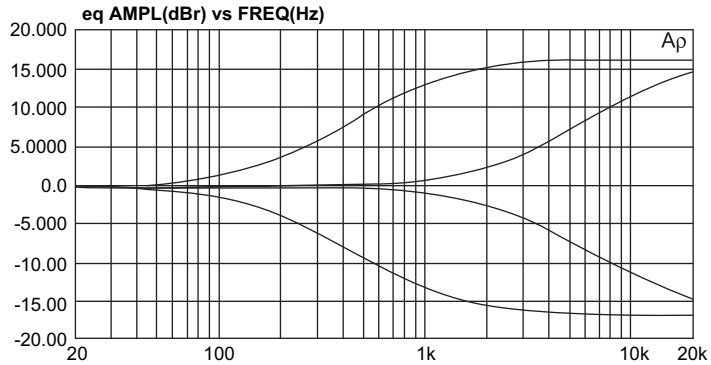


Figure 10.13: The response of the shelving circuit with C1 increased to 100 nF. Maximum cut/boost is slightly greater than ±15 dB

Tone-Balance Controls

A tone-balance or tilt circuit is operated by a single control, and affects not just part of the audio spectrum, but most or all of it. Typically the high frequencies are boosted as the low frequencies are cut, and vice versa. It has to be said that the name is unfortunate as it implies it has something to do with interchannel amplitude balance, which it has not. A 'stereo tone-balance control' alters the frequency response of both channels equally and does not introduce amplitude differences between them, whereas a 'stereo balance control' is something quite different. It is clearer to call a tone-balance control a tilt control.

Tone-balance controls are (or were) supposedly useful in correcting the overall tonal balance of recordings in a smoother way than a Baxandall configuration, which concentrates on the ends of the audio spectrum. An excellent (and very clever) approach to this was published by Ambler in 1970 [6] (see Figure 10.14). The configuration is very

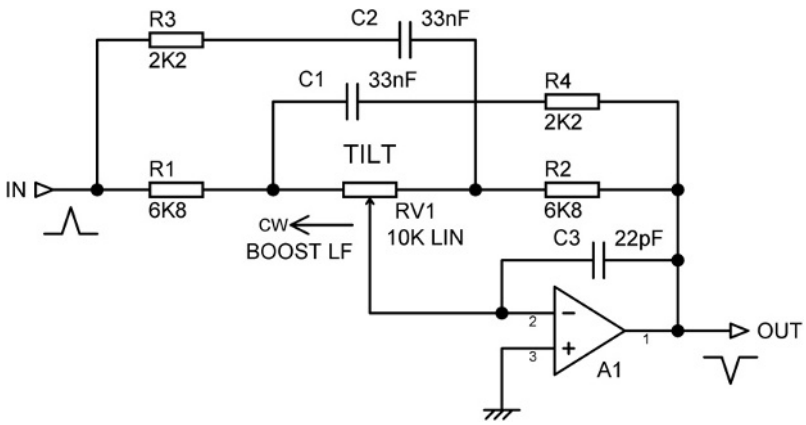


Figure 10.14: Tone-balance control of the Ambler type

www.raguvaran.puzl.com

similar to the Baxandall's – the ingenious difference here is that the boost/cut pot effectively swaps its ends over as the frequency goes up. At low frequencies C1, C2 do nothing, and the gain is set by the pot, with maximum cut and boost set by R1, R2. At high frequencies, where the capacitors are effectively short-circuit, R3, R4 overpower R1, R2 and the control works in reverse. The range available with the circuit shown is ±8 dB at LF and ±6.5 dB at HF. This may seem ungenerous, but because of the way the control works, 8 dB of boost in the bass is accompanied by 6.5 dB of cut in the treble, and a total change of 14.5 dB in the relative level of the two parts of the spectrum should be enough for anyone. The measured frequency response at the control limits is shown in Figure 10.15; the response is not quite flat with the control central due to component tolerances.

The need for one set of end-stop resistors to take over from the other puts limits on the cut/boost that can be obtained without the input impedance becoming too low; there is of course also the equivalent need to consider the impedance the op-amp sees when driving the feedback side of the network.

The input impedance at LF, with the control set to flat, is approximately 12 kΩ, which is the sum of R3 and half of the pot resistance. At HF, however, the impedance falls to 2.0 kΩ. Please note that this is not a reflection of the values of the HF end-stop resistors R3, R4, but just a coincidence. When the control is set to full treble boost, the input impedance at HF falls as low as 620 Ω. The impedance at LF holds up rather better at full bass boost, as it cannot fall below the value of R1, i.e. 6.8 kΩ.

The impedances of the circuit shown here have been reduced by a factor of 10 from the original values published by Ambler, to make them more suitable for use with op-amps. The original (1970) gain element was a two-transistor inverting amplifier with limited linearity and load-driving capability. Here a stabilizing capacitor C3 is shown explicitly, just to remind you that you might need one.

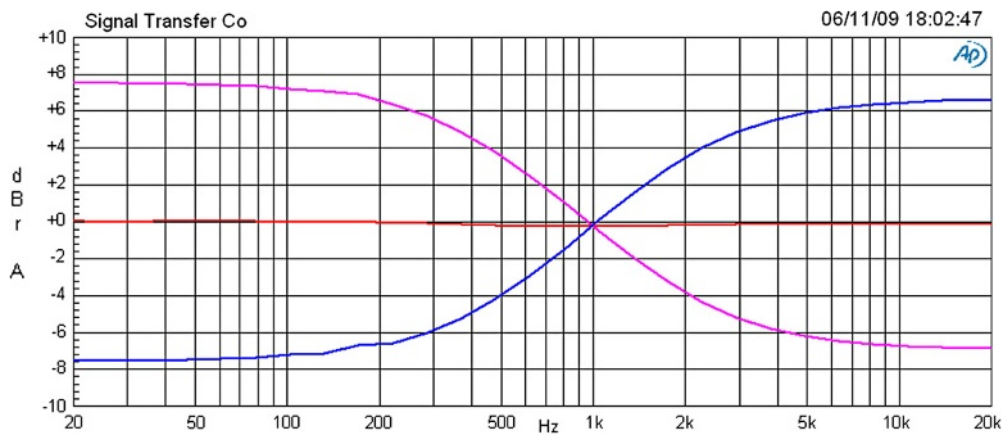


Figure 10.15: Frequency response of tone-balance control

www.raguvaran.puzl.com

A famous example of the use of a tilt control is the Quad 44 preamplifier. The tilt facility is combined with a bass cut/boost control in one quite complicated stage, and it is not at all obvious if the design is based on the Ambler concept. Tilt controls have never really caught on and remain rare.

Middle Controls

A middle control affects the center of the audio band rather than the bass and treble extremes. It must be said at once that middle controls, while useful in mixers, are of very little value in a preamplifier. If the middle frequency is fixed, then the chances that this frequency and its associated Q correspond with room shortcomings or loudspeaker problems are remote in the extreme. Occasionally middle controls appeared on preamps in the 1970s, but only rarely and without much evidence of success in the marketplace. One example is the Metrosound ST60 (1972), which combined a three-band Baxandall tone control (more on this below) with slider controls in one package. The middle control had a very wide bandwidth centered on 1 kHz, and it was suggested that it could be used to depress the whole middle of the audio band to give the effect of a loudness control.

Middle controls come into their own in mixers and other sound-control equipment, where they are found in widely varying degrees of sophistication. In recording applications middle controls play a vital part in 'voicing' or adjusting the timbres of particular instruments, and the flexibility of the equalizer, and its number of controls defines the possibilities open to the operator.

The obvious first step is to add a fixed middle control to the standard HF and LF controls. Unfortunately, this is not much more useful in a mixer than in a preamplifier. In the past this was addressed by adding more fixed middles, so a line-up with a high middle and a low middle would be HF–HMF–LMF–LF, but this takes up more front panel space (which is a very precious resource in advanced and complex mixers, and ultimately defined by the length of the human arm) without greatly improving the EQ versatility.

The minimum facilities in a mixer input channel for proper control are the usual HF and LF controls plus a sweep middle with a useful range of center frequency. This also requires four knobs but is much more useful.

Fixed-Frequency Baxandall Middle Controls

Figure 10.16 shows a middle version of a Baxandall configuration. The single control RV1 now has around it both the time-constants that were before assigned to the separate bass and treble controls. R4 and R5 maintain unity gain at DC, and keep the stage biased correctly.

www.raguvaran.puzl.com

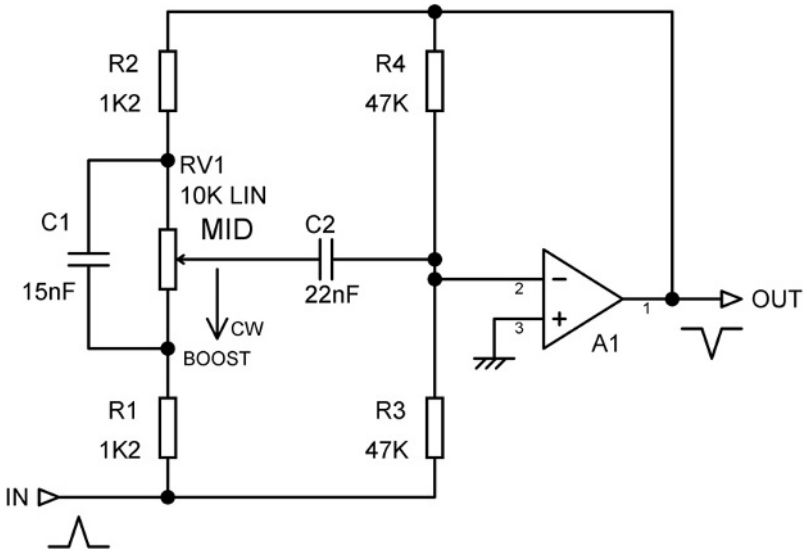


Figure 10.16: Fixed middle control of the Baxandall type. Center frequency 1.26 kHz

As the input frequency increases from the bottom of the audio band, the impedance of C2 falls and the position of the pot wiper begins to take effect. At a higher frequency, the impedance of C1 becomes low enough to effectively tie the two ends of the pot together, so that the wiper position no longer has effect, and the circuit reverts to having a fixed gain of unity. The component values shown give a mid frequency of 1.26 kHz, at a Q of 0.8, with a maximum boost/cut of ± 15 dB. The Q value is only valid at maximum boost/cut; with less the curve is flatter and the effective Q lower. It is not possible to obtain high values of Q with this approach.

This circuit gives the pleasingly symmetrical curves shown in Figure 10.17, though it has to be said that the benefits of exact symmetry are visual rather than audible.

As mentioned above, in the simpler mixer input channels it is not uncommon to have two fixed mid controls; this is not the ideal arrangement, but it can be implemented very neatly and cheaply, as in Figure 10.18. There are two stages, each of which has two fixed bands of EQ. It has the great advantage that there are two inverting stages so the output signal ends up back in phase. The first stage needs the extra resistors R5, R6 to maintain DC feedback.

There will inevitably be some control interaction with this scheme. It could be avoided by using four separate stages, but this is most unlikely to be economical for mixers with this relatively simple sort of EQ. To minimize interaction, the control bands are allocated between the stages to keep the frequencies controlled in a stage as far apart as possible, combining HF with LO MID, and LF with HI MID, as shown here.

www.raguvaran.puzl.com

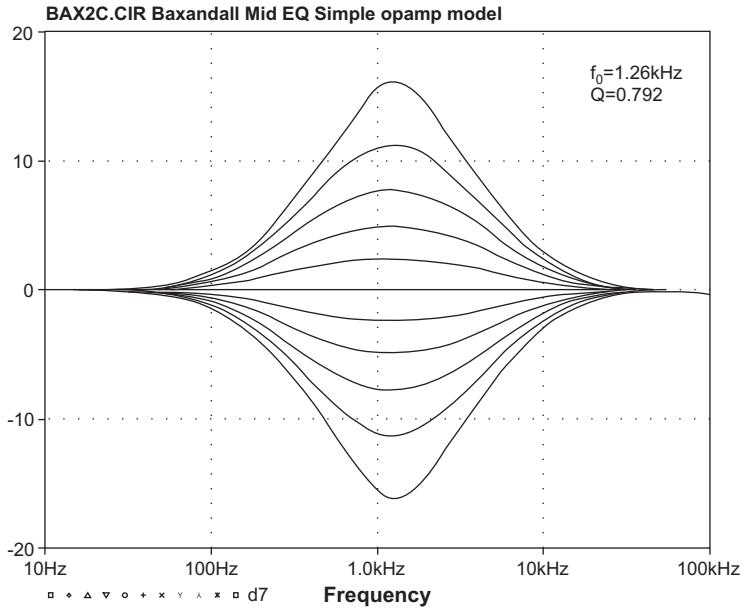


Figure 10.17: The frequency response of the Baxandall middle control in Figure 10.16

Three-Band Baxandall EQ in One Stage

The standard Baxandall tone control allows adjustment of two bands with one stage. When a three-band EQ is required, it is common practice to use one such stage for HF and LF, and a following one to implement the middle control, as in Figure 10.16. This has the advantage that the two cascaded inverting stages will leave the signal in the correct phase.

When this is not a benefit, because a phase inversion is present at some other point in the signal path, it is economical to combine HF, MID, and LF in one quasi-Baxandall stage. This not only reduces component count, but reduces power consumption by saving an op-amp.

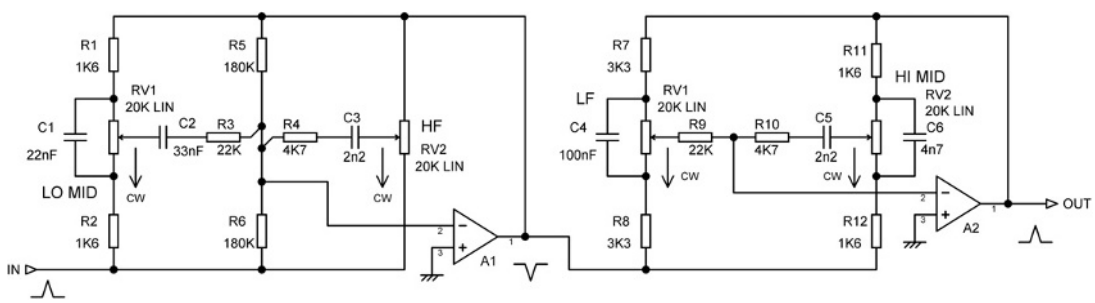


Figure 10.18: A four-band Baxandall EQ using two stages only

www.raguvaran.puzl.com

The drawback is that cramming all this functionality into one stage requires some compromises on control interaction and maximum boost/cut. The circuit shown in Figure 10.19 gives boost/cut limited to ±12 dB in each band.

The frequency responses for each band are given in Figures 10.20–10.22.

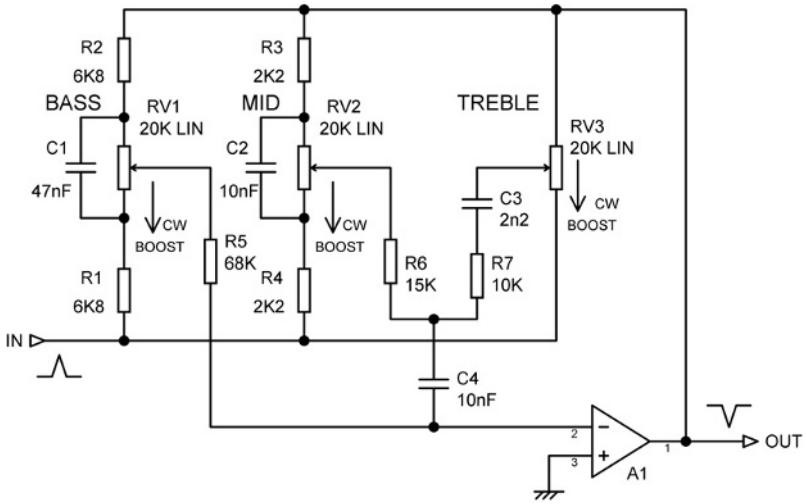


Figure 10.19: The circuit of a Baxandall three-band EQ using one stage only

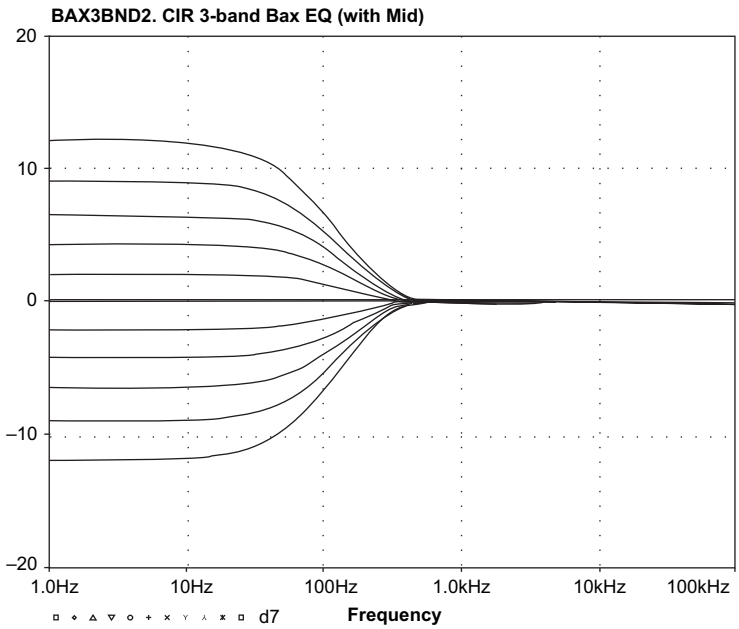


Figure 10.20: The frequency response of a Baxandall three-band EQ. Bass control

www.raguvaran.puzl.com

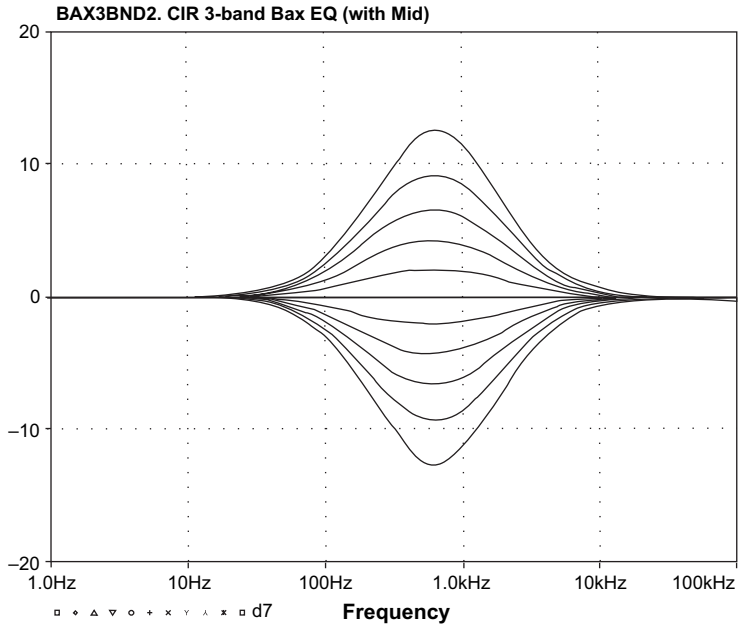


Figure 10.21: The frequency response of a Baxandall three-band EQ. Mid control

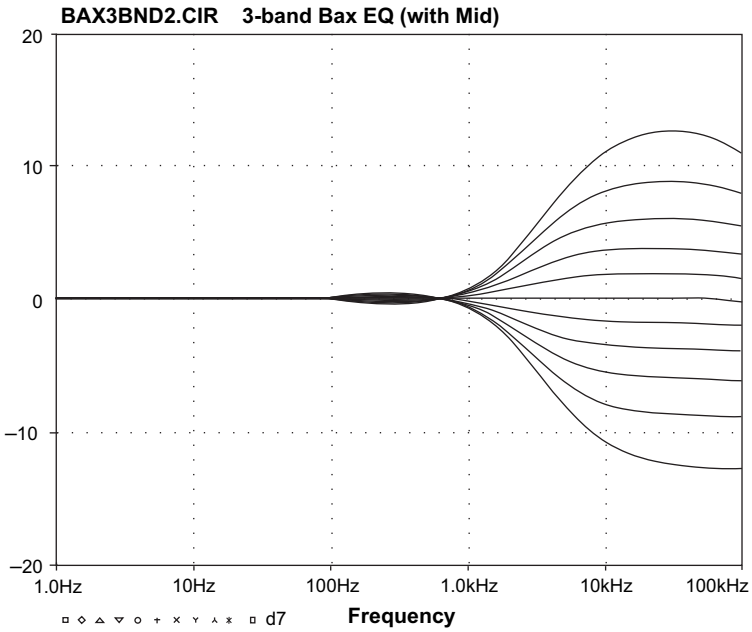


Figure 10.22: The frequency response of the Baxandall three-band EQ. Treble control

www.raguvaran.puzl.com

Wien Fixed Middle EQ

An alternative way to implement a fixed middle control is shown in Figure 10.23. Here the signal tapped off from RV1 is fed to a Wien bandpass network R1, C1, R2, C2, and returned to the op-amp non-inverting input. This is the same Wien network as used in audio oscillators.

With the values shown the center frequency is 2.26 kHz and the *Q* at max cut/boost is 1.4; it gives beautifully symmetrical response curves like those in Figure 10.17, with a maximal cut/boost of 15.5 dB.

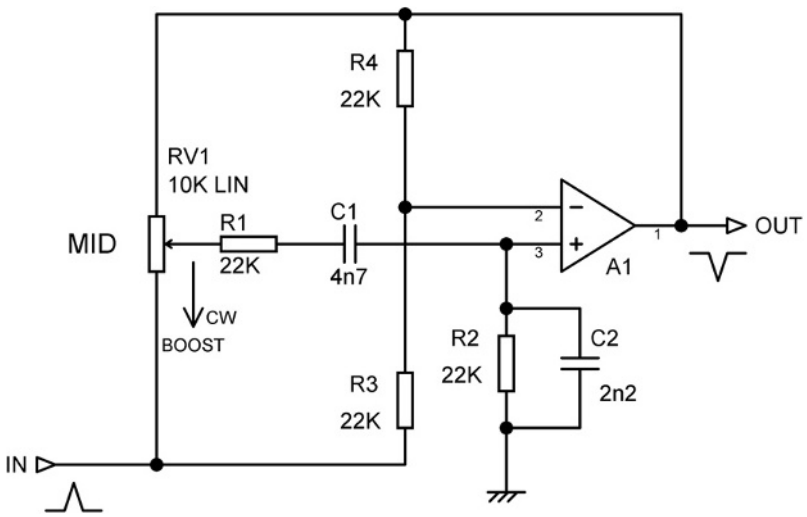


Figure 10.23: The Wien fixed-middle circuit. Center frequency is 2.26 kHz. The *Q* at max cut/boost is 1.4

Sweep Middles

A fully variable-frequency middle control is much more useful and versatile than any combination of fixed or switch middles. In professional audio this is usually called a 'sweep middle' EQ. It can be implemented very nicely by putting variable resistances in the Wien network of the stage previously described, and the resulting circuit is shown in Figure 10.24.

The variable load that the Wien network puts on the cut/boost pot RV1 causes a small amount of control interaction, which is normally considered acceptable. It could be eliminated by putting a unity-gain buffer stage between the RV1 wiper and the Wien network, but in the middle-range mixers where this circuit is commonly used, this is not normally economical.

The Wien network is carefully arranged so that the two variable resistors RV2, RV3 can have common terminals, reducing the number required from six to three. This is sometimes taken

www.raguvaran.puzl.com

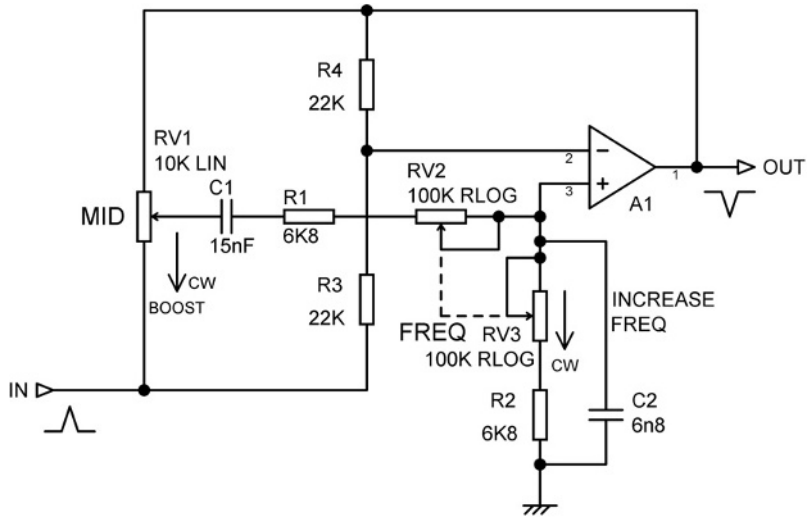


Figure 10.24: A Wien sweep circuit. The center frequency range is 150 Hz – 2.4 kHz (a 16:1 ratio)

advantage of by pot manufacturers making ganged parts specifically for this EQ application. R1, C1 are sometimes seen swapped in position but this naturally makes no difference.

The combination of a 100k pot and a 6k8 end-stop resistor gives a theoretical ratio of 15.7:1, which is about as much as can be obtained using reverse-log law C pots, without excessive cramping at the high-frequency end of the scale. This will be marked on the control calibrations as a 16:1 range.

The measured frequency responses at the control limits are shown in Figure 10.25. The frequency range is from 150 Hz to 2.3 kHz; the ratio is slightly adrift due to component tolerances.

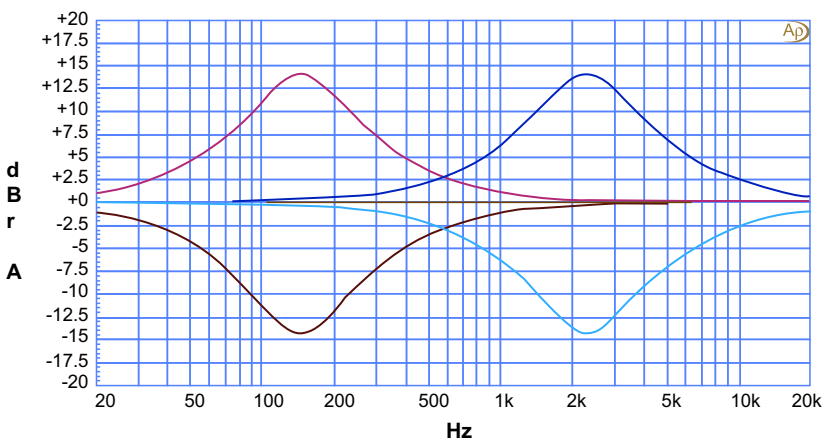


Figure 10.25: The measured response of the sweep middle circuit at control extremes. The cut/boost is slightly short of ±15 dB

www.raguvaran.puzl.com

The Single-Gang Sweep Middle

The usual type of sweep middle requires a dual-gang reverse-log pot to set the frequency. These are not hard to obtain in production quantities, but can be difficult to get in small numbers. They are always significantly more expensive than a single pot.

The problem becomes more difficult when the design requires a stereo sweep middle – if implemented in the usual way, this demands a four-gang reverse-log pot. Once again, such components are available, but only to special order, which means long lead-times and significant minimum order quantities. Four-gang pots are not possible in flat-format mixer construction where the pots are mounted on their backs, so to speak, on a single big horizontal PCB. The incentive to use a standard component is strong, and if a single-gang sweep middle circuit can be devised, a stereo EQ only requires a dual-gang pot.

This is why many people have tried to design single-gang sweep middle circuits, with varying degrees of success. It can be done, so long as you don't mind some variation of Q with center frequency; the big problem is to minimize this interaction. I too have attacked this problem, and here is my best shot so far, in Figure 10.26.

This circuit is a variation of the Wien middle EQ, the quasi-Wien network being tuned by a single control RV2, which not only varies the total resistance of the R5, R7 arm, but also the amount of bootstrapping applied to C2, effectively altering its value. This time a unity-gain buffer stage A2 has been inserted between the RV1 wiper and the Wien network; this helps to minimize variation of Q with frequency.

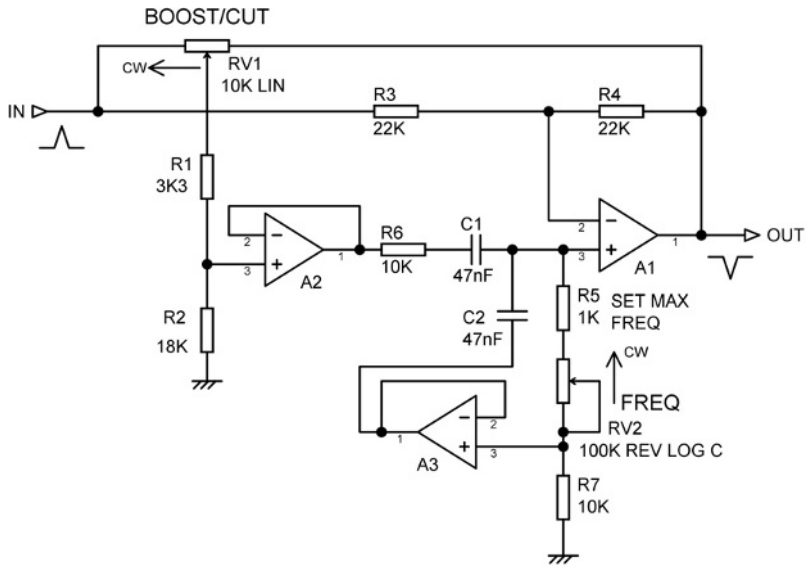


Figure 10.26: My single-gang sweep middle circuit. The center frequency range is 100 Hz – 1 kHz

www.raguvaran.puzl.com

The response is shown in Figure 10.27; note that the frequency range has been restricted to 10:1 to minimize Q variation. The graph shows only maximum cut/boost; at intermediate settings the Q variations are much less obvious.

It is possible to make a more economical version of this, if one accepts somewhat greater interaction between boost/cut, and Q and frequency. The version shown in Figure 10.28 omits

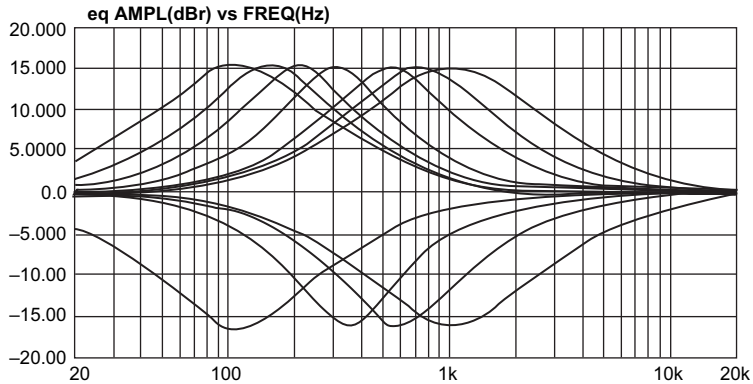


Figure 10.27: The response of the single-gang sweep middle circuit of Figure 10.26. Boost/cut is ±15 dB, and the frequency range is 100 Hz – 1 kHz The Q varies somewhat with center frequency

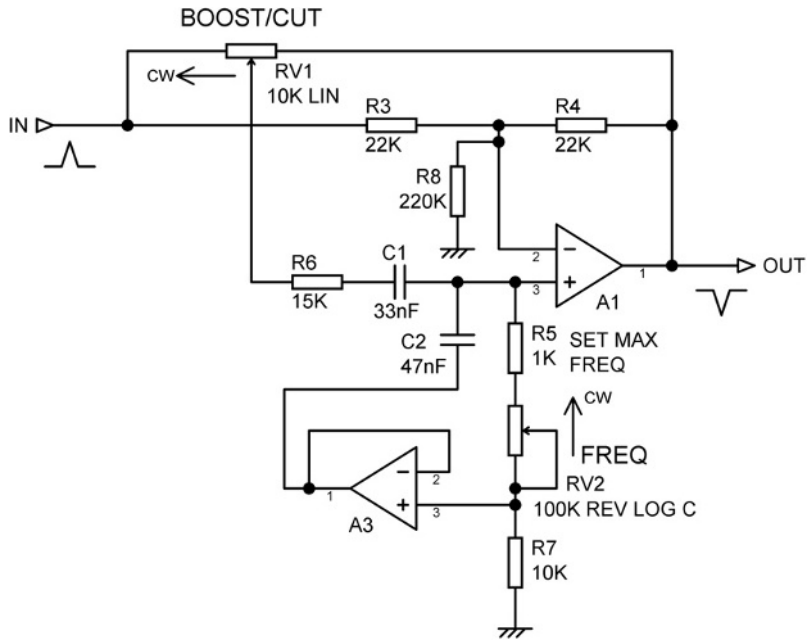


Figure 10.28: My economical single-gang sweep middle circuit. The center frequency range has been changed to 220 Hz – 2.2 kHz

www.raguvaran.puzl.com

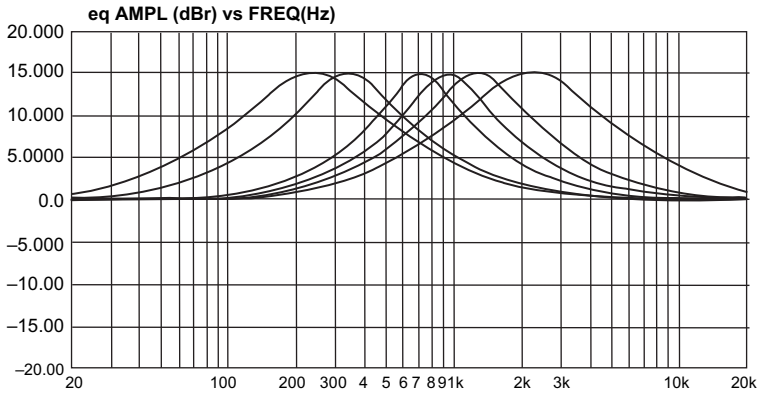


Figure 10.29: The response of the economical single-gang sweep middle circuit of Figure 10.28. The response is ±15 dB as before, but the frequency range has now been set to 220 Hz – 2 kHz. Only the boost curves are shown

the unity-gain buffer and uses unequal capacitor values to raise the *Q* of the quasi-Wien network, saving an op-amp section. The frequency range is still 10:1.

The response in Figure 10.29 shows the drawback – a higher *Q* at the center of the frequency range. Once again the *Q* variations will be much less obvious at intermediate cut/boost settings.

The question naturally arises as to whether it is possible to design a single-gang sweep middle circuit where there is absolutely no variation of *Q* with frequency. Is there an 'existence theorem', i.e. a mathematical proof that it can't be done? At the present time, I just don't know.

Switched-*Q* Variable-Frequency Wien Middle EQs

The next step in increasing EQ sophistication is to provide control over the *Q* of the middle resonance. This is often accomplished by using a full state-variable filter solution, which gives fully variable *Q* that does not interact with the other control settings, but if two or more switched values of *Q* are sufficient, there are much simpler circuits available.

One of them is shown in Figure 10.30; here the Wien bandpass network is implemented around A2, which is essentially a shunt-feedback stage, with added positive feedback via R1, R2 to raise the *Q* of the resonance. When the *Q* switch is in the LO position, the output from A2 is fed directly back to the non-inverting input of A1, because R7 is short-circuited. When the *Q* switch is in the HI position, R9 is switched into circuit and increases the positive feedback to A2, raising the *Q* of the resonance. This also increases the gain at the center frequency, and this is compensated for by the attenuation now introduced by R7 and R8.

www.raguvaran.puzl.com

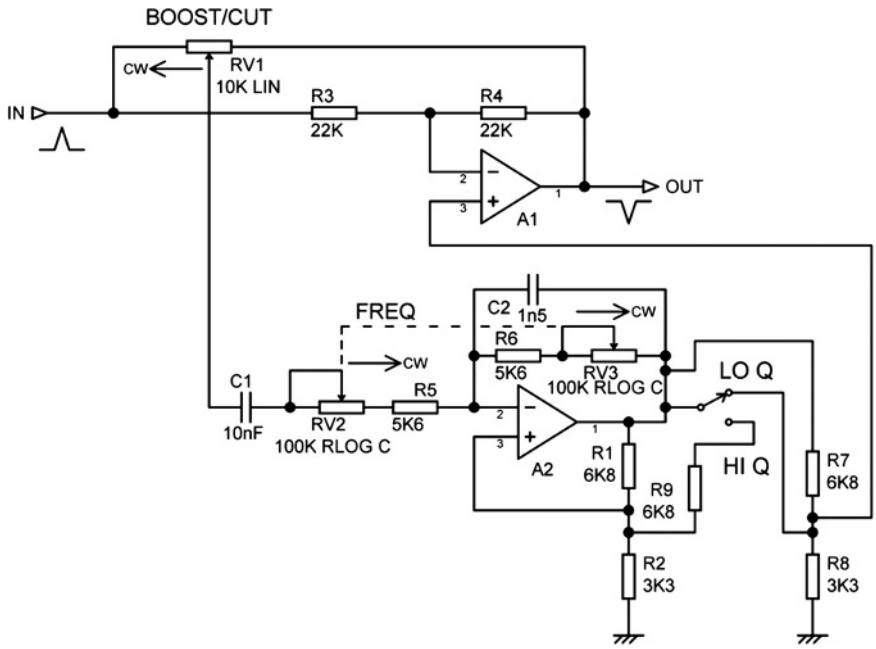


Figure 10.30: Variable-frequency switched-*Q* middle control of the Wien type

With the values shown the two *Q* values are 0.5 and 1.5. Note the cunning way that the *Q* switch is made to do two jobs at once – changing the *Q* and also introducing the compensating attenuation. If the other half of a two-pole switch is already dedicated to an LED indicator, this saves having to go to a four-pole switch. On a large mixing console with many EQ sections this sort of economy is important.

Switchable Peak/Shelving LF/HF EQs

It is frequently desirable to have the highest and lowest frequency EQ sections switchable between a peaking (resonance) mode and shelving operation. The peaking mode allows relatively large amounts of boost to be applied near the edges of the audio band without having a large and undesirable amount occurring outside it.

Figure 10.31 shows one way of accomplishing this. It is essentially a switchable combination of the variable-frequency HF shelving circuit of Figure 10.12 and the sweep middle circuit of Figure 10.24; when the switches are in the PEAK position the signal tapped off RV1 is fed via the buffer A2 to a Wien bandpass network C2, RV2, R5, C3, R6, RV3, and the circuit has a peak/dip characteristic. When the switches are in the SHLV (shelving) position, the first half of the Wien network is disconnected and C1 is switched in, and in conjunction with R6 and RV3 forms a first-order high-pass network, fed by an attenuated signal because R2 is now grounded. This switched attenuation factor is required to give equal amounts of cut/boost in

www.raguvaran.puzl.com

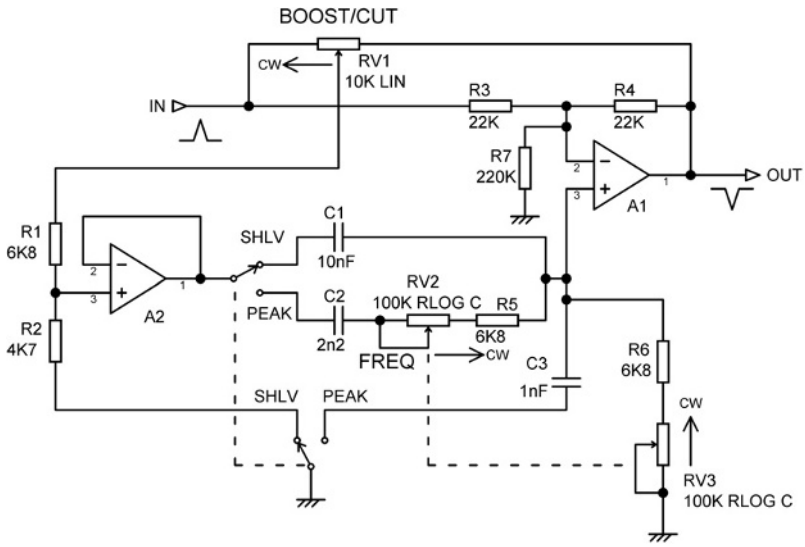


Figure 10.31: Variable-frequency peak/shelving HF EQ circuit

the two modes because the high-pass network has less loss than the Wien network. R7 allows fine-tuning of the maximum cut/boost; reducing it increases the range.

As always we want our switches to work as hard as possible, and the lower switch can be seen to vary the attenuation brought about by R1, R2 with one contact and switch in C3 with the other. Unfortunately, in this case two poles of switching are required. The response of the circuit at one frequency setting can be seen in Figure 10.32.

When the peaking is near the edges of the audio band, this is sometimes called return-to-zero (RTZ) operation as the gain returns to unity (0 dB) outside the band.

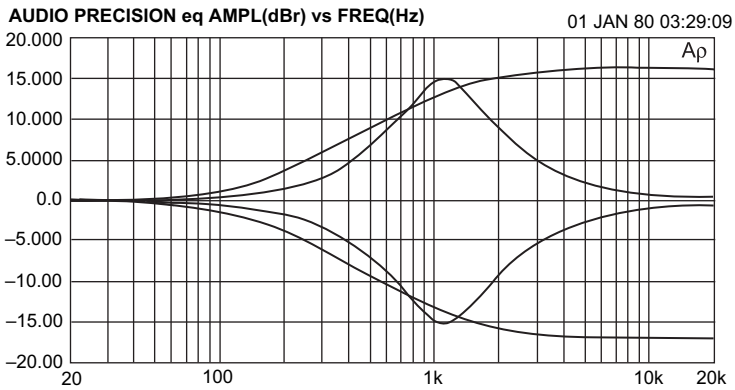


Figure 10.32: The response of the variable-frequency peak/shelving with R7 = 220k; the cut/boost range is thus set to ±15 dB

www.raguvaran.puzl.com

Parametric Middle EQs

A normal second-order resonance is completely defined by specifying its center frequency, its bandwidth or *Q*, and the gain at the peak. In mathematical language, these are the parameters of the resonance. Hence an equalizer that allows all three to be changed independently (proviso on that coming up soon) is called a parametric equalizer. Upscale mixing consoles typically have two fully parametric middle sections, and usually the LF and HF can also be switched from shelving to peaking mode, when they become two more fully parametric sections.

The parametric middle EQ shown in Figure 10.33 is included partly for its historical interest, showing how op-amps and discrete transistor circuitry were combined in the days before completely acceptable op-amps became affordable. I designed it in 1979 for a now long-gone company called Progressive Electronics, which worked in a niche market for low-noise mixing consoles. The circuitry I developed was a quite subtle mix of discrete transistor and op-amp circuitry, which gave a significantly better noise performance than designs based entirely on the less-than-perfect op-amps of the day; in time, of course, this niche disappeared. This parametric middle EQ was used, in conjunction with the usual HF and LF controls, in a channel module called the CM4.

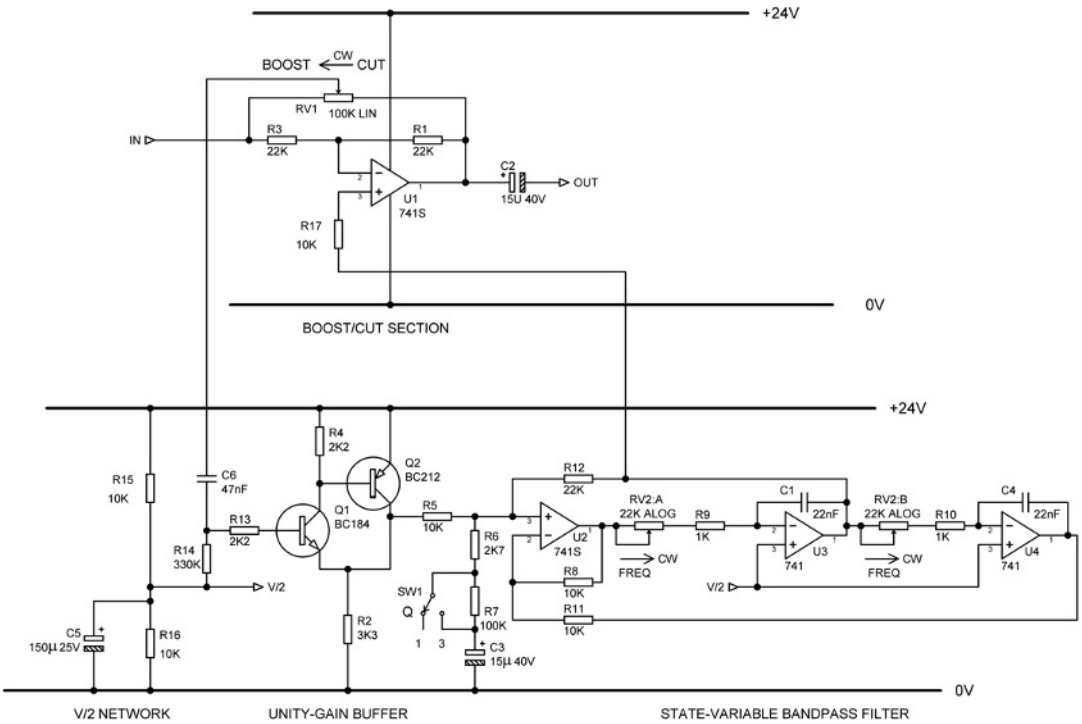


Figure 10.33: An historical parametric middle EQ dating back to 1979

www.raguvaran.puzl.com

The boost/cut section used an op-amp because of the need for both inverting and non-inverting inputs. I used a 741S, which was a completely different animal from the humble 741, with a much better distortion performance and slew rate; it was, however, markedly more expensive, and only used where its superior performance was really necessary. The unity-gain buffer Q1, Q2 that ensured a low-impedance drive to the state-variable bandpass filter was a discrete circuit block, as its function is simple to implement. Q1 and Q2 form a CFP emitter-follower. R13 was a 'base-stopper' resistor to make sure that the Q1, Q2 local feedback loop did not exhibit VHF parasitic oscillation. With the wisdom of hindsight, putting a 2k2 resistor directly in the signal path can only degrade the noise performance, and if I was doing it again I would try to solve the problem in a more elegant fashion. The high input impedance of the buffer stage (set by R14) means that C6 can be a small non-electrolytic component.

The wholly conventional state-variable bandpass filter requires a differential stage U2, which once again is best implemented with an op-amp, and another expensive 741S was pressed into service. The two integrators U3, U4 presented an interesting problem. Since only an inverting input is required, discrete amplifiers could have been used without excessive circuit complexity; a two- or possibly three-transistor circuit (see Chapter 3) would have been adequate. However, the PCB area for this approach just wasn't there, and so op-amps had to be used. To put in two more 741S op-amps would have been too costly, and so it pretty much had to be a couple of the much-despised 741 op-amps. In fact, they worked entirely satisfactorily, because they were in integrator stages. The poor HF distortion and slew rate were not really an issue because of the large amount of NFB at HF, and the fact that integrator outputs by definition do not slew quickly. The indifferent noise performance was also not an issue because the falling frequency response of the integrators filtered out most of the noise. In my designs the common-or-garden 741 was only used in this particular application. Looking at the circuit again, I have reservations about the not inconsiderable 741 bias currents flowing through the two sections of the frequency-control pot RV2, which could make them noisy, but it seemed to work alright at the time.

The filter Q was set by the resistance of R6, R7 to ground. It does not interact with filter gain or center frequency. The Q control could easily have been made fully variable by using a potentiometer here, but there was only room on the channel front panel for a small toggle switch. Note the necessity for the DC-blocking capacitor C3, because all the circuitry is biased at V/2 above ground.

The filtered signal is fed back to the boost/cut section through R17, and I have to say that at this distance of time I am unsure why that resistor was present. It could only impair the noise performance. Grounding it would give an EQ-cancel that would also stop noise from the state variable filter.

It is worth noting that the design dates back to when the use of single supply rails was customary. In part this was due to grave and widely held doubts about the reliability of

www.raguvaran.puzl.com

electrolytic coupling capacitors with no DC voltage across them, which would be the case if dual rails were used. As it happened, there proved to be no real problem with this, and things would have progressed much faster if capacitor manufacturers had not been so very wary of committing themselves to approving non-polarized operation. The use of a single supply rail naturally requires that the circuitry is biased to V/2, and this voltage was generated in the design shown by R15, R16, and C5; it was then distributed to wherever it was required in the channel signal path. The single rail was at +24 V, because 24 V IC regulators were the highest voltage versions available, and nobody wants to get involved with designing discrete power supply regulators if they can avoid it. This is obviously equivalent to a ±12 V dual rail supply, compared with the ±15 or ±17 V that was adopted when dual-rail powering became universal, and so gave a headroom that was lower by 1.9 and 3.0 dB respectively.

This design is included here because it is a good example of making use of diverse circuit techniques to obtain the best possible performance/cost ratio at a given point in time. It could be brought up to date quite quickly by replacing all the antique op-amps and the discrete unity-gain buffer with 5532s or other modern types.

Modern parametric equalizers naturally use all op-amp circuitry. Figure 10.34 shows a parametric EQ stage I designed back in 1991; it is relatively conventional, with a three-stage state-variable filter composed of A2, A3, and A4. There is, however, one improvement on the standard circuit topology. Most of the noise in a parametric equalizer comes from the filter path. In this version the filter path signal level is set to be 6 dB higher than usual, with the desired return level being restored by the attenuator R6, R7. This attenuates the filter noise as well, and the result is a parametric section approximately 6 dB quieter than the industry standard. The circuit is configured so that, despite this raised level, clipping cannot occur in the filter with any combination of control settings. If excess boost is applied, clipping can only happen at the output of A1, as usual.

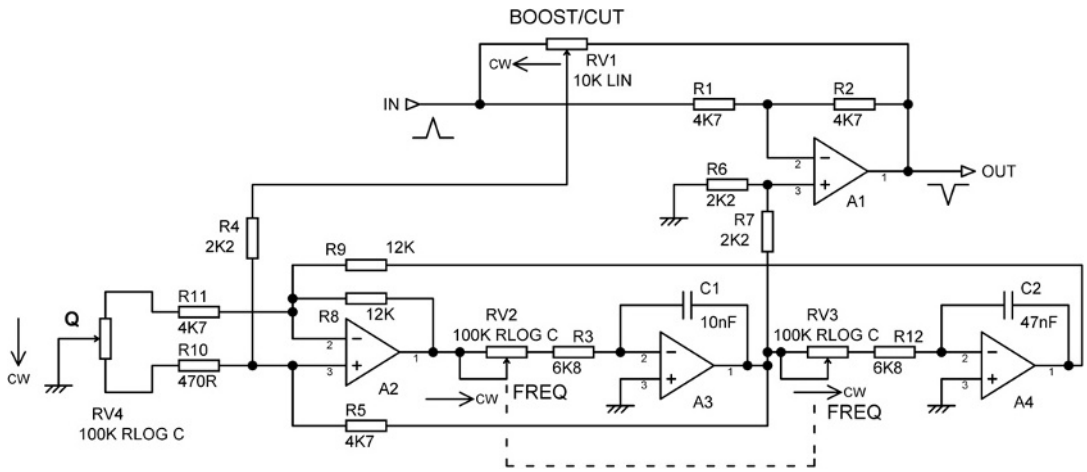


Figure 10.34: Variable-frequency variable-*Q* middle control. State-variable type

www.raguvaran.puzl.com

Other features are the Q control RV4, which is configured to give a wide parameter range without affecting the gain. Note the relatively low values for the DC feedback resistors R1 and R2, chosen to minimize Johnson noise without causing excessive op-amp loading.

This equalizer section has component values for typical low-mid use, with center frequency variable over a wide range from 70 Hz to 1.2 kHz, and Q variable from 0.7 to 5. The cut and boost range is the usual ± 15 dB.

Graphic Equalizers

Graphic equalizers are so called because their cut/boost controls are vertical sliders, the assumption being that a graph of the frequency response will pass through the slider knob positions. Graphic equalizers can have any number of bands from three to 31, the latter having bands one-third of an octave wide. This is the most popular choice for serious room equalization work as bands one-third-octave wide relate to the perceptual critical bands of human hearing.

Graphic equalizers are not normally fitted to large mixing consoles, but are often found on smaller powered mixers, usually in the path between the stereo mix and the power amplifiers. The number of bands provided is limited by the space on the mixer control surface, and is usually in the range of seven to 10.

There is more than one way to make a graphic equalizer, but the most common version is shown in its basic concept in Figure 10.35, with some typical values. L1, C1, and R3 make up an LCR series resonant circuit that has a high impedance except around its resonant frequency; at this frequency the reactances of L1, C1 cancel each other out and the impedance to ground is simply that of R3. At resonance, when the wiper of RV1 is at the R1 end, the LCR circuit forms the lower leg of an attenuator of which R1 is the upper arm; this attenuates the input signal and a dip in the frequency response is therefore created. When the RV1 wiper is at the R2 end, an attenuator is formed with R2 that reduces the feedback factor at resonance and so creates a peak in the response. It is not exactly intuitively obvious, but this process does give symmetrical cut/boost curves. At frequencies away from resonance the impedance of the LRC circuit is high and the gain of the circuit is unity.

The beauty of this arrangement is that two, or three or more LRC circuits, with associated cut/boost pots, can be connected between the two op-amp inputs, giving us an equalizer with pretty much as many bands as we want. Obviously, the more bands we have, the narrower they must be to fit together properly.

As described in Chapter 2, inductors are in general thoroughly unwelcome in a modern design, and the great breakthrough in graphic equalizers came when the LRC circuits were replaced by gyrator circuits that emulated them but used only resistors, capacitors, and a gain element. It is

www.raguvaran.puzl.com

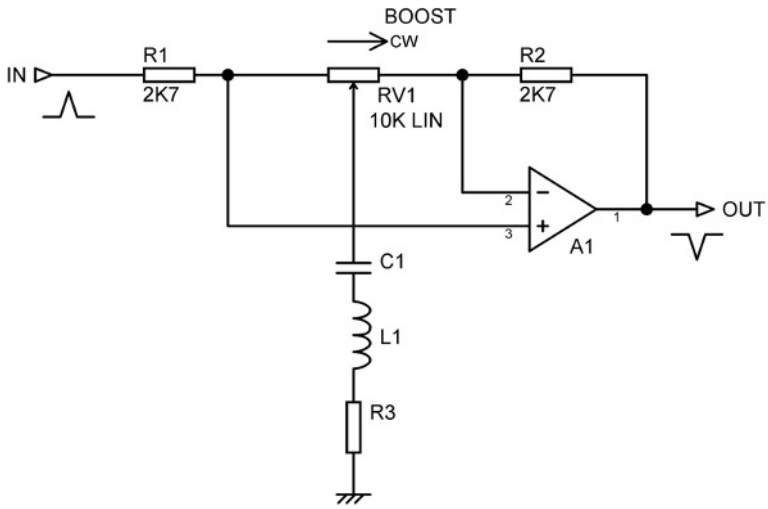


Figure 10.35: The basic idea behind graphic equalizers; gain is unity with the wiper central

not too clear just when this idea spread, but I can testify that by 1975 gyrators were the standard approach, and the use of inductors would have been thought risible.

The basic notion is shown in Figure 10.36; C1 works as a normal capacitor as in the LCR circuit, while C2 pretends to be the inductor L1. As the applied frequency rises, the attenuation of the high-pass network C2–R1 reduces, so that a greater signal is applied to unity-gain buffer A1 and it more effectively bootstraps the point X, making the impedance from it to ground increase. Therefore we have a circuit fragment where the impedance rises proportionally to frequency – which is just how an inductor behaves. There are limits to the Q values that can be obtained with this circuit because of the inevitable presence of R1 and R2.

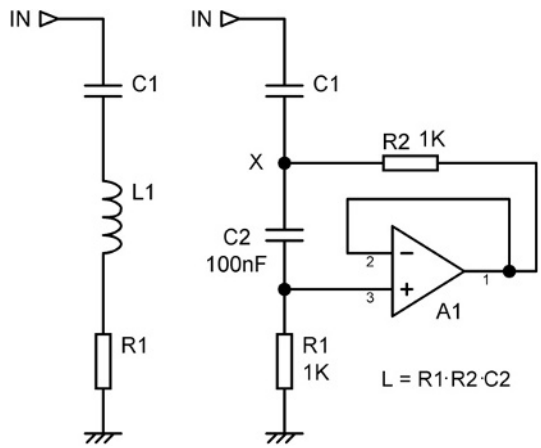


Figure 10.36: Using a gyrator to synthesize a grounded inductor in series with a resistance

www.raguvaran.puzl.com

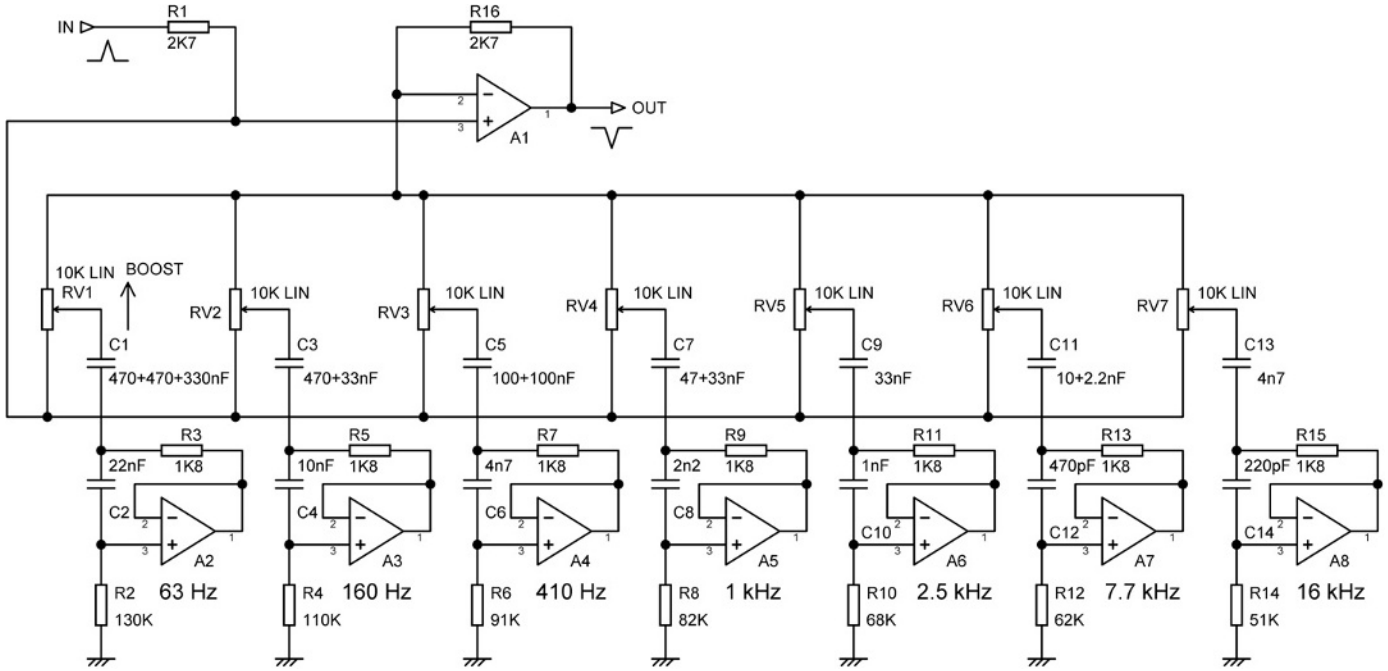


Figure 10.37: A seven-band graphic equalizer

www.raguvaran.puzl.com

The sample values in Figure 10.36 synthesize a grounded inductor of 100 mH (which would be quite a hefty component if it was real) in series with a resistance of 2 kΩ. Note the surprisingly simple equation for the inductor value. Another important point is that the op-amp is used as a unity-gain buffer, which means that the early gyrator graphic equalizers could use a simple emitter-follower in this role. The linearity was naturally not so good, but it worked and made graphic equalizers affordable.

A simple seven-band gyrator-based graphic equalizer is shown in Figure 10.37. The maximal cut/boost is ±8 dB. The band center frequencies are 63, 160 and 410 Hz, and 1, 2.5, 7.7 and 16 kHz. The Q of each band at maximum cut or boost is 0.9.

The response of each band is similar to that shown in Figure 10.21. The maximum Q value is only obtained at maximum cut or boost. For all intermediate settings the Q is lower. This behavior is typical of the straightforward equalizer design shown here, and is usually referred to as 'proportional-Q' operation, and results in a frequency response that is very different from what might be expected on looking at the slider positions.

There is, however, another mode of operation called 'constant-Q' in which the Q of each band does not decrease as the cut/boost is reduced [7]. This gives a frequency response that more closely resembles the slider positions.

The graphic equalizer described here has a symmetrical response, also known as a reciprocal response; the curves are the same for cut and for boost operations. It is also possible to design an equalizer for an asymmetric or non-reciprocal response, in which the boost curves are as shown above, but the cut response is a narrow notch. This is often considered to be more effective when the equalizer is being used to combat feedback in a sound reinforcement system.

References

[1] F. Langford-Smith (Ed.), The Radio Designer's Handbook, fourth ed. Newnes, 1953, pp. 635–677. reprinted 1999, Chapter 15.
[2] F. Langford-Smith (Ed.), The Radio Designer's Handbook, fourth ed. Newnes, 1953, p. 668. reprinted 1999, Chapter 15.
[3] H.T. Sterling, Flexible dual control system, Audio Engineering (February 1949).
[4] P. Baxandall, Negative-feedback tone control, Wireless World (October 1952) p. 402.
[5] D. Self, Self On Audio, second ed. Newnes, 2006, p. 5.
[6] R. Ambler, Tone-balance control, Wireless World (March 1970) p. 124.
[7] D.A. Bohn, Constant-Q graphic equalizers, JAES (September 1986) p. 611.

www.raguvaran.puzl.com

This page intentionally left blank

www.raguvaran.puzl.com

CHAPTER 11

A Complete Preamplifier

In this chapter a complete preamplifier is described, to bring to life the various circuit blocks described in other chapters. The design is basically that of the 'Precision Preamplifier 96' published in *Electronics World* in July–September 1996 [1]. The description here is fairly concise because most of the circuit blocks are fully described in the relevant chapters of this book.

Architecture and Philosophy

The preamplifier architecture is as established in Chapter 6. The moving-magnet (MM) phono amplifier gain is +29 dB, raising an input of 5 mVrms (all gains and levels are at 1 kHz) to a line-out level of 150 mVrms nominal, matching equipment levels that appear to have reached some sort of consensus on this value.

The separate moving-coil (MC) stage will accept inputs of 0.1 or 0.5 mV and gives an output of 5 mV that is suitable for the MM input stage. The MC stage provides flat gain, while the MM stage performs the RIAA equalization for both modes. The MM stage is followed by a third-order subsonic filter, which also implements the IEC Amendment. There are an odd number of half-5532s, so the single 5534 is placed in the MM stage, where its somewhat lower noise is put to best use. The MC input amplifier is fully described in Chapter 8, and the MM input amplifier and subsonic filter is fully described in Chapter 7. Full schematics of these circuit blocks are given there.

The phono amplifier stages feed one input of the input selector switch bank, which is followed by the unbuffered recording output, the recording input, and the monitoring switch that selects between them (see Figure 11.1). The principle of low-crosstalk input select switching is covered in Chapter 16.

The next stage is an input buffer to give low-impedance drive to the tone-control stage. It has a gain variable over a small range by the balance control; its gain is +1.1 dB with balance central. The tone-control stage that follows has fully variable turnover frequencies and a non-interrupting tone-cancel facility. The tone control phase-inverts, but this is reversed by the active gain control system (AGS) after it. The maximum gain of the AGS is +17 dB, giving a maximum output of 1 Vrms for a line input of 150 mVrms.

Small Signal Audio Design; ISBN: 9780240521770
Copyright © 2010 Elsevier Ltd; All rights of reproduction, in any form, reserved.

www.raguvaran.puzl.com

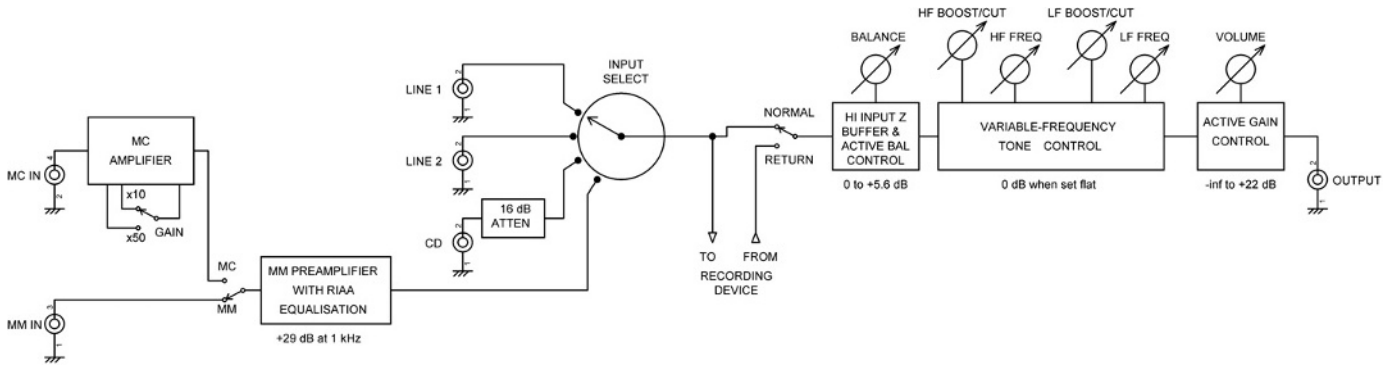


Figure 11.1: Block diagram of the complete preamplifier

www.raguvaran.puzl.com

The philosophy of this design is to achieve low noise by minimizing circuit impedances so the contribution of Johnson noise is reduced, and the conditions are set for best bipolar op-amp noise performance. The balance control, tone control, and volume control are all active gain configurations to eliminate even small compromises resulting from attenuation-then-amplification (increased noise) or amplification-then-attenuation (reduced headroom). Absolute phase is preserved for all input and outputs.

Another design aim was to make as many components in the design as possible the same value, to simplify procurement. All the pots are 10 kΩ linear; most sizeable non-electrolytics are 470 nF, most small electrolytics are 10 μF 50 V, and all large ones are 220 μF 35 V. This philosophy has to be abandoned in areas where critical parameters are set, such as in the RIAA network and tone-control stage.

The Line Inputs

The input is selected by a push-button switch bank rather than a rotary control, as this allows a four-pole configuration to be used that much improves the rejection of unselected inputs. This is fully described in Chapter 16 on signal switching.

The CD input stands out from other line sources in that its nominal level is usually high at 2 Vrms. This is possible because digital sources have rigidly defined maximum output levels, so they can be high to improve noise performance, since there is no danger of the analog output section clipping. However, this means a direct line input cannot be used without the trouble of resetting volume and recording-level controls when CD is selected. To minimize this problem a 16 dB passive attenuator, as shown in Figure 16.1b, is placed in the CD input path. The assumption is that a CD output has a low impedance, and will have no problem driving a 10 kΩ input impedance. The divider resistance values can therefore be made low to minimize the noise degradation. The output impedance of the attenuator is 1.4 kΩ, which generates −120.9 dBu of Johnson noise as opposed to −135.2 dBu from a direct 50 Ω source. This is still much less than the preamp internal noise so the noise performance is not worsened. Inter-source crosstalk is improved simply by back-grounding the CD attenuator output when not in use, so only a two-pole switch is required for good isolation of this source.

The High-Impedance Input Buffer

It can be argued that the input impedance of a preamplifier should be high to allow for interfacing to valve equipment whose output may be taken from a valve anode, where even light loading compromises distortion and available output swing. A plausible minimum input impedance for this is 100 kΩ; many preamplifier designs fall well short of this.

www.raguvaran.puzl.com

This stage therefore has to provide both a high input impedance and a variable gain for the balance control. This is a vernier balance control that alters relative channel gain by only +4.5, −1.1 dB, a difference of 6 dB; this is sufficient to move the stereo image completely from one side to the other. Since the minimum gain of this non-inverting stage is unity, the nominal gain with balance control central is +1.1 dB, and the maximum gain of the active gain stage (AGS) is chosen to allow for this. The active nature of the balance control means that the signal never receives unwanted attenuation that must be undone later with noisy amplification.

The input buffer is shown in Figure 11.2. A high input impedance is obtained simply by using a 470 kΩ biasing resistor R2, accepting that the bias current through this will give a negative output offset of 95 mV typical; this is not big enough to significantly reduce headroom. The gain law is modified by R3 to give as little gain as possible in the central control position. Maximum gain is set by R4, and DC gain held to unity by C4; R1 and C2 make up an EMC filter. The noise output of the buffer/balance stage is of the same order as the test-gear and cannot be measured directly – a subtraction method gives a noise output of about −116 dBu with balance central.

An important point is that the significant bias current of the op-amp inverting input is not allowed to flow through the balance pot; this would make it noisy. The bias current has a separate DC path through R3, while the AC feedback goes through C3.

The Tone-Control Stage

The tone-control stage is described here rather than in Chapter 10, because it is specialized in that it aims at very good noise performance. The design is not a conventional Baxandall tone control. It has break frequencies variable over a 10:1 range, which makes the facility much more useful for correcting speaker deficiencies. This extra flexibility may fly right in the face

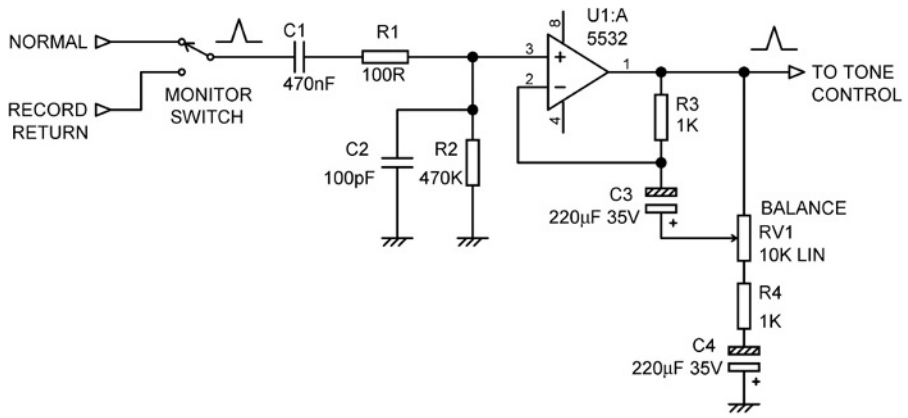


Figure 11.2: The high-input-impedance buffer, with variable gain for balance control

www.raguvaran.puzl.com

of subjectivist thinking, but I can live with that. Variable boost/cut and frequency enables any error at the top or bottom end to be corrected to at least a first approximation. It makes a major difference.

Certain features of the tone control may make it more acceptable to those prey to doubts about its sonic correctness. The tone-control range is ±10 dB, rather than the ±15 dB standard in mixing consoles. The response is built entirely from simple 6 dB/octave circuitry, with inherently gentle slopes. The stage is naturally minimum phase, and so the amplitude curves uniquely define the phase response. The maximum phase-shift does not exceed 40° at full boost. This is a return-to-flat (RTF) tone control, meaning that the curves do not plateau or shelve at their boosted or cut level, but smoothly return to unity gain outside the audio band. Boosting 10 kHz is one thing, but boosting 200 kHz is quite another, and could lead to some interesting stability or EMC problems. The RTF time-constants are fixed so the boost/cut range is necessarily less at the frequency extremes, where the RTF effect starts to overlap the variable boost/cut frequencies.

The basic principle is shown in Figure 11.3. The overall stage gives a unity-gain inversion except when the selective responses of the side-chain paths allow signal through. In the treble and bass frequency ranges where the side-chains pass signal, the boost/cut pots VR2, VR4 can

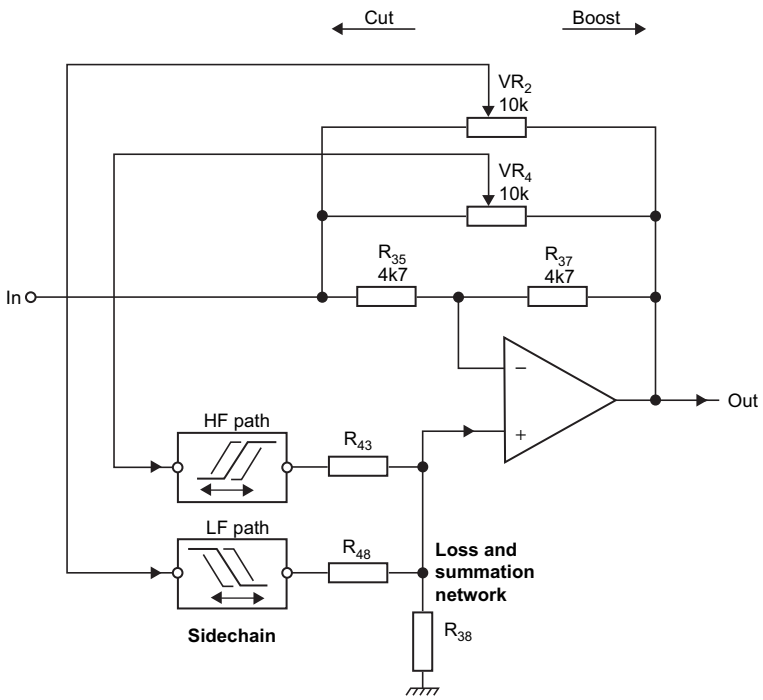


Figure 11.3: The principle of the tone-control stage. The response only deviates from unity gain at frequencies passed by the HF or LF paths

www.raguvaran.puzl.com

give either gain or attenuation. When a wiper is central, there is a null at the middle of the boost/cut pot, no signal through that side-chain, and the gain is unity.

If the pot is set so the side-chain is fed from the input then there is a partial cancellation of the forward signal; if the side-chain is fed from the output then there is a partial cancellation of the negative feedback, or to put it another way, positive feedback is introduced to counteract part of the NFB through R37. This apparently ramshackle process actually gives boost/cut curves of perfect symmetry, though this is purely cosmetic, because you can't use both sides of the curve at once, so it hardly matters if they are exact mirror images.

The tone-control stage acts in separate bands for bass and treble, so there are two parallel selective paths in the side-chain. These are simple RC time-constants, the bass path being a variable-frequency first-order low-pass filter, and the associated bass control only acting on the frequencies this lets through. Similarly the treble path is a variable-frequency high-pass filter. The filtered signals are summed and returned to the main path via the non-inverting input, and some attenuation must be introduced to limit cut and boost. Assuming a unity-gain side-chain, the necessary loss is 9 dB if cut and boost are to be limited to ±10 dB. This is implemented by R43, R48, and R38 in Figures 11.4–11.6. The side-chain has unity gain, and so has no problems with clipping before the main path does, so it is very desirable to put the 9 dB loss *after* the side-chain, where it attenuates side-chain noise. The loss attenuator is made up of the lowest value resistors that can be driven without distortion, to minimize both the Johnson noise thereof and the noise generated by op-amp IC7b.

The tone-cancel switch disconnects the entire side-chain (i.e. five out of six op-amps) from any contribution to the main path, and usefully reduces the stage output noise by about 4 dB, depending on the HF frequency setting. It leaves only IC7b in circuit, which must be retained anyway to undo the gain-control phase inversion. Unlike configurations where the entire stage is bypassed, the signal does *not* briefly disappear as the cancel switch moves between two

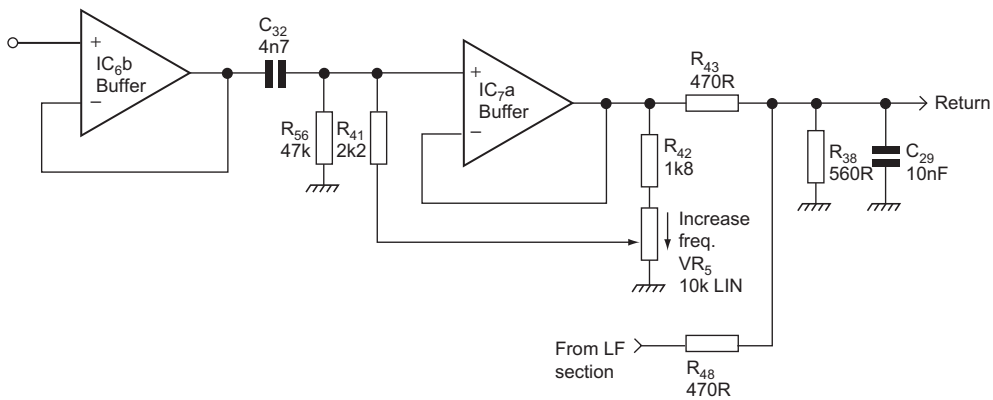


Figure 11.4: HF path of the tone control

www.raguvaran.puzl.com

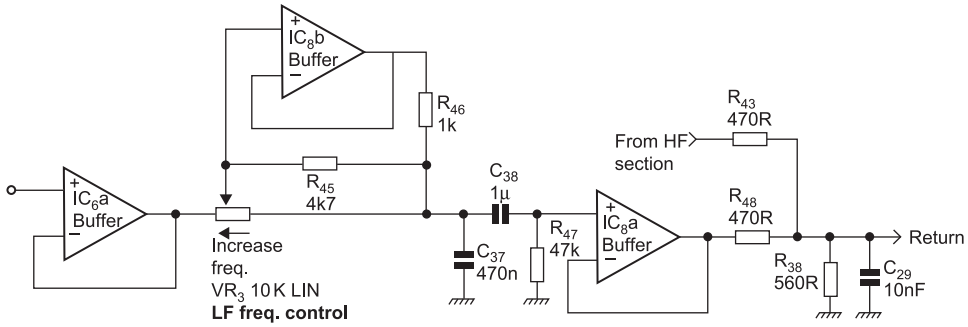


Figure 11.5: LF path of the tone control

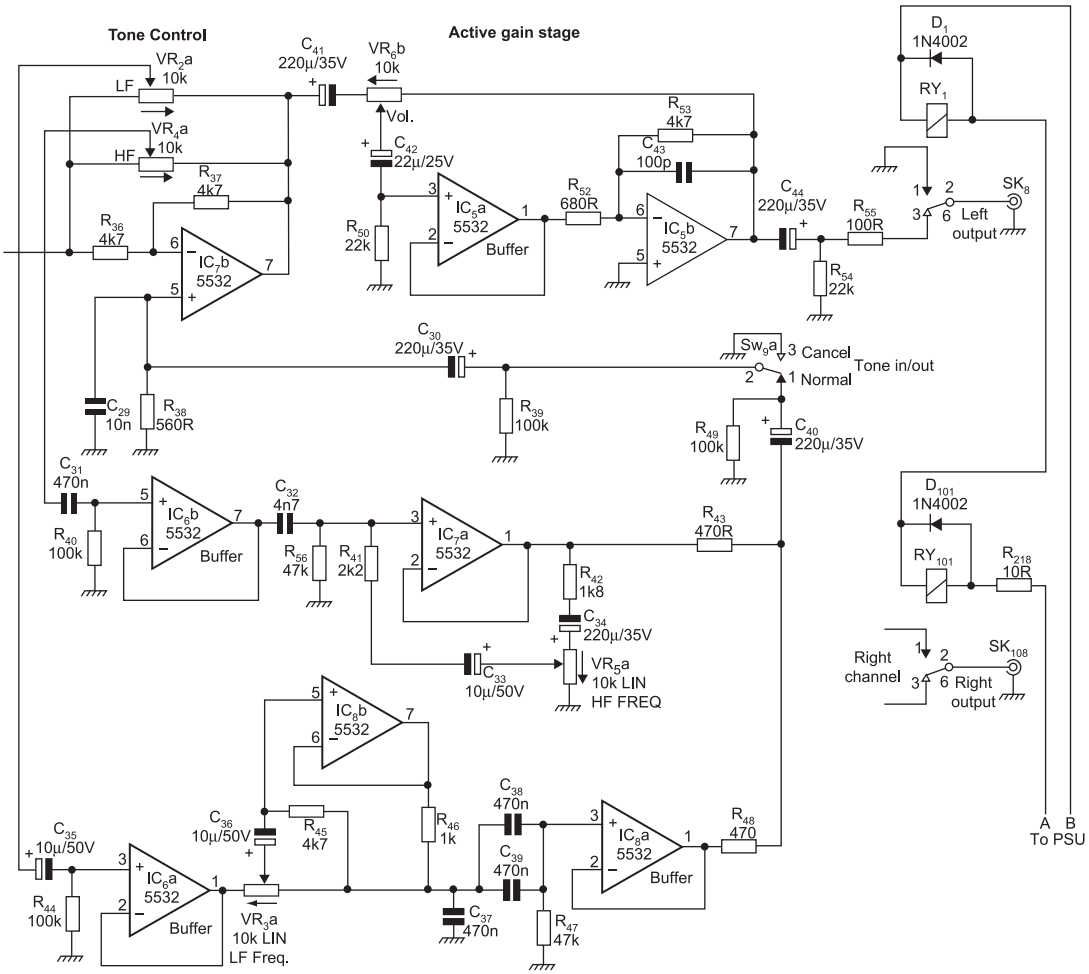


Figure 11.6: Full schematic of the tone control and active gain stage

www.raguvaran.puzl.com

contacts. This minimizes transients due to suddenly chopping the waveform and makes valid tone in/out comparisons much easier.

It is very convenient if all the pots in a design are identical, and I am quite proud of having achieved it here. I have used linear 10 kΩ controls throughout to avoid log-pot tolerances. This has required some thoughtful circuit design to give all the controls the logarithmic action that fits our perceptual processes. It is, however, desirable for the balance and the tone-control cut/boost pots to have central detents, which somewhat goes against the 'all pots identical' philosophy.

The basics of the HF path are shown in Figure 11.4. C32 and R41 are the high-pass time-constant, driven at low impedance by unity-gain buffer IC6b. This buffer prevents the high-pass frequency from changing with the boost/cut setting. The effective value of R41 is altered over a 10:1 range by varying the amount of bootstrapping it receives from IC7a, the potential divider effect and the rise in source resistance of VR5 in the center combining to give a reasonable approximation to a logarithmic frequency/rotation law, as shown in Figure 11.8 below. R42 is the frequency end-stop resistor that limits the maximum effective value of R41. C29 is the HF RTF capacitor – at frequencies above the audio band it shunts all the side-chain signal to ground, so the treble control no longer has any further effect.

The treble side-chain degrades the noise performance of the tone-control stage by 2–3 dB when connected. As would be expected, the noise contribution is greatest when the HF frequency is set to minimum, and a wider bandwidth from the side-chain is contributing to the main path.

The simplified LF path is shown in Figure 11.5. IC6a buffers the cut/boost control VR2 to prevent boost/frequency interaction. The low-pass time-constant capacitor is C37, and the associated time-constant resistance is a combination of VR3 and R45, R46. The RTF time-constant for the bass path is set by C38 and R47, which block very low frequencies and so limit the lower extent of bass control action. The bass frequency law is made approximately logarithmic by IC8b; for minimum frequency VR3 is set fully anticlockwise, so the input of buffer IC8b is the same as at the C37 end of R46, which is thus bootstrapped and has no effect. When VR3 is fully clockwise, R45, R46 are effectively in parallel with VR3 and the turnover frequency is at a maximum. The presence of R45 gives an approximately logarithmic law (see Figure 11.7). Compared with the HF path, an extra op-amp is required. However, despite this, the bass side-chain contributes very little extra noise to the tone stage. This is because most of its output is rolled off by the low-pass action of C37 at high frequencies, almost eliminating its noise contribution.

Once the active elements (here 5532s) have been chosen, and the architecture made sensible in terms of avoiding attenuation-then-amplification, keeping noise gain to a minimum, and so on,

www.raguvaran.puzl.com

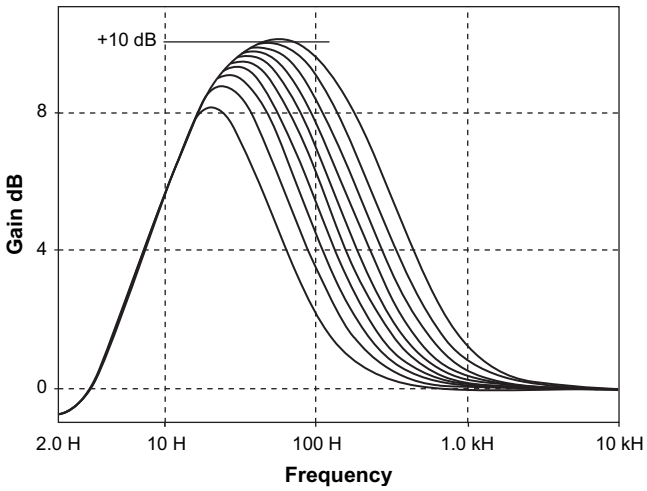


Figure 11.7: LF boost curves of the tone control

the remaining way of improving noise performance is by impedance reduction of the circuitry. The resistances are lowered in value, with capacitances scaled up to suit, by a factor limited only by the op-amp drive capability. This is another good reason to use the 5534/2. Two examples of this process as applied to the tone stage are given here. In each case the noise improvement is for the stage in isolation, set flat with HF frequency at minimum:

1. In this kind of stage R36, R37 are conventionally 22 kΩ. Using 4k7 instead dropped the noise output by 1.3 dB.

2. The summation/loss network began life with R43, R48 at 4k7, and R38 at 5k6. Reducing this by a factor of 10, to 470 and 560 Ω respectively, reduced output noise by 0.6 dB.

Neither of these improvements is startling in itself, but both cost nothing, and they all add up. The noise output of the tone stage, plus the line/balance buffer before it, with balance control central and tone cancel pressed, is −107.2 dBu (22 kHz bandwidth). With tone controls active but set flat, noise output at minimum HF frequency is −104.7 dBu, and at maximum HF frequency is −106.7 dBu.

The final tone stage and the AGS are shown in Figure 11.6. The tone section may look rather a mess of pottage, and appear to have more buffers than Clapham Junction, but this modest degree of complexity is required if control interaction is to be wholly eliminated.

The practical tone circuitry contains more capacitors than the simplified circuits in Figures 11.4 and 11.5, and this reflects a disadvantage of low-noise bipolar input op-amps. The bias currents are non-negligible and must not be allowed to flow through potentiometers if crackling noises are to be avoided when they are moved. These bias currents also tend to cause significant output offset voltages; the source resistances for the two op-amp inputs are not

www.raguvaran.puzl.com

normally the same and so there is no offset cancellation. For this reason all the variable-gain circuit sections have their gain reduced to unity at DC by blocking capacitors. C31 and C35 keep bias currents out of VR2, VR4, necessitating bias resistors R40, R44.

The treble frequency law is corrected by bootstrapping through C33, which keeps the bias current of IC7a out of VR5. Similarly C34 prevents any offset on IC7a output reaching VR5. In the bass path C36 keeps IC8b bias out of VR3, while RTF components C38, C39 and R47 inherently provide DC blocking. The final offsets at the side-chain output are blocked by C40, while the bias voltage on non-inverting input IC7b is blocked by C30. This is essential to prevent the tone-cancel switch clicking. This switch may still appear to click if it switches in or out a large amount of response modification of a non-zero signal, because the abrupt gain change generates a step in the waveform that is heard as a click. This is unavoidable with hard audio switching.

The maximal cut and boost curves over the frequency ranges are shown in Figures 11.7 and 11.8. Figure 11.9 collates the measured extremes of cut and boost at the frequency extremes.

The Active Gain Stage

The active gain stage (AGS) used here is the Baxandall configuration [2], which I have used previously in another preamplifier design [3]. The maximum gain is set to +17 dB by the ratio of R52 to R53. This active volume-control stage gives the usual advantages of lower noise at gain settings below maximum, and in the special case of the Baxandall configuration, excellent

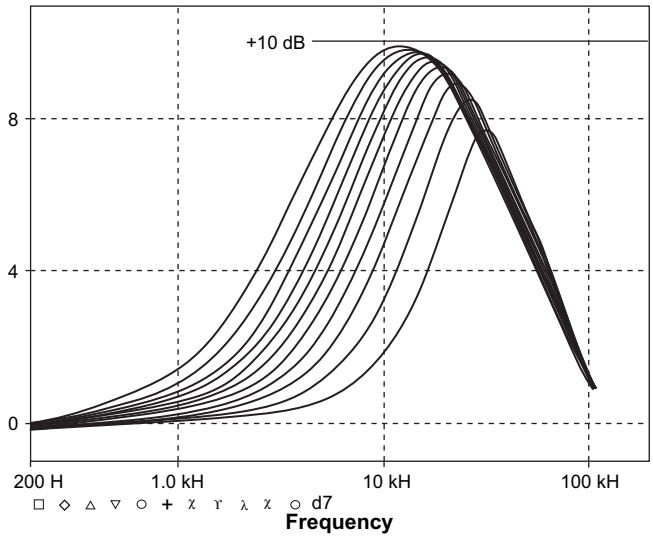


Figure 11.8: AHF boost curves of the tone control

www.raguvaran.puzl.com

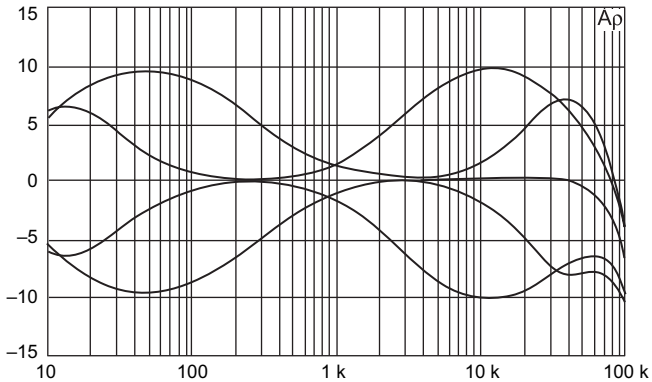


Figure 11.9: Maximum cut and boost curves of the tone control at the frequency setting extremes

channel balance that depends solely on the mechanical alignment of the dual linear pot – all mismatches of its electrical characteristics are canceled out, and there are no quasi-log dual slopes to induce anxiety. The noise performance has been improved by an impedance reduction on the gain-definition network R52, R53. The limit on this is the ability of buffer IC5a to drive R52, which has a virtual earth at its other end.

There are the usual complications to deal with the DC: C41 keeps IC7b output offset out of volume control VR6, while C42 blocks IC5a bias current from VR6 wiper. C44 gives final output DC blocking. Details of the volume-control law and other aspects of active gain stage design are given in Chapter 9.

Figure 11.6 includes relay muting on the main outputs. The control circuitry is not shown.

Performance

The THD of the MC stage alone, at 50 times gain, is below 0.001% from 20 Hz to 3 kHz, rising to 0.002% at 20 kHz; this is at a 2 Vrms output, which is of course massive compared with the actual operating level of 5 mVrms. The THD of the MM disc input is well below 0.001% at 8 Vrms output, rising to 0.002% at 20 kHz. The entire rear section has THD below 0.001% at 8 Vrms output at all frequencies. These levels are all much higher than those in real use; distortion is not a problem here.

The crosstalk performance attained depends very much on physical layout. Capacitative crosstalk can be minimized by spacing components well apart, or by simple screening. Resistive crosstalk depends on the thickness of the various ground paths. The crosstalk results depend somewhat on control settings, but under most conditions a prototype built on a double-sided PCB in ground-plane format gave about −100 dB flat across 20 Hz – 20 kHz, with noise contributing to the reading.

www.raguvaran.puzl.com

TABLE 11.1 Noise output from preamplifier rear section (dBu)

	Tone cancel	Tone in, flat
AGS zero gain	−114.5	−114.5
AGS unity gain	−107.4	−105.3
AGS fully up	−90.2	−86.4

The measured noise output of the MM stage was −92.3 dBu (unweighted) with the input loaded with a Shure M75ED. The RIAA-equalized noise output from the disc stage in MC mode was −93.9 dBu for 10 times gain, and −85.8 dBu for 50 times. In the 10 times case the noise in MC mode is actually 1.7 dB lower than in MM mode.

For the rear section of the preamp (i.e. that after the recording monitor switch) the noise performance depends on the control settings. Table 11.1 gives the noise outputs; with the tone control in circuit the HF frequency control is at minimum, the worst case.

The figures for maximum gain may look unimpressive, but remember this is with +17 dB of gain; at normal volume settings the noise output is below −100 dBu. I think this is reasonably quiet.

When the original design was published [1], I invited correspondence on whether it was the quietest preamp yet built, expecting to be challenged at once – I wasn't. Never, in fact.

References

[1] D. Self, Precision Preamplifier 96, Electronics World (July/August and September 1996).
[2] P. Baxandall, Audio gain controls, Wireless World (November 1980) pp. 79–81.
[3] D. Self, A precision preamplifier, Wireless World (October 1983) p. 31.

www.raguvaran.puzl.com

CHAPTER 12

Mixer Architectures

A large mixing console arguably represents the most demanding area of audio design. The steady advance of digital media demands that every part of the chain that takes music from performer to consumer must be near-perfect, as the comfortable certainty that everything will be squeezed through the quality bottleneck of either analog tape or vinyl disc now looks very old-fashioned.

The technical problems that must be overcome in a professional mixing console are many and varied, and often unique to console design. A large number of signals flow in a small space and they must be kept strictly apart until the operator chooses to mix them; crosstalk must be exceedingly low.

There may be 72 input channels or more, each with many stages that all have the potential to add distortion and noise to the precious signal. Even summing these signals together, while sounding trivially easy, is in practice a major challenge. The quality requirements, especially for recording consoles, are much more demanding than those for the most expensive hi-fi equipment, because degradation introduced at this stage can never be retrieved.

Performance Factors

Two primary requirements of modern consoles are very low noise and minimal distortion. Since a comprehensive console must pass the audio through a large number of circuit stages (perhaps over 100 from microphone to final mixdown), great attention to detail is essential at each stage to prevent a build-up of noise and distortion; one of the most important trade-offs is the impedance of the circuitry surrounding the op-amp, for if this is too high Johnson noise will be increased, while if it is too low an op-amp will exhibit non-linearity in struggling to drive it.

The choice of op-amp is also critical, for cost considerations discourage the global use of expensive ICs. In a console with many stages of signal processing, this becomes a major concern; nonetheless it is possible to keep the right-through THD below 0.002% at 20 dB above the normal operating level, so that at normal levels it is unmeasurable.

Small Signal Audio Design; ISBN: 9780240521770
Copyright © 2010 Elsevier Ltd; All rights of reproduction, in any form, reserved.

www.raguvaran.puzl.com

I will use the terms mixer, mixing console, and mixing desk interchangeably where it will make the text less clunky. In the USA it is sometimes called simply 'the board'. I hope I can write 'mixer' without anyone thinking I am referring to the front end of a radio or a kitchen appliance, or indeed anything concrete-related. Throughout this chapter the various sections of the mixer – the channels, groups, master, etc. – will be referred to as 'modules', even if they are not constructed as separate physical modules.

In hi-fi, the preservation of absolute phase has only recently come to be considered necessary; with normal listening material it appears highly unlikely that it actually makes any audible difference, but it unquestionably looks good on the spec. In mixer design the meticulous preservation of phase is a rigid requirement, to avoid the possible cancellation of signals. Both summing amplifiers and most EQ sections use shunt feedback and thus invert the signal, and it is sometimes necessary to include inverting stages that do nothing but get the phase straight again. It is a mark of good design to keep such stages to a minimum.

Mixer Internal Levels

The internal signal levels of any mixing console are always a compromise between noise and headroom. A higher level means that the signal is less compromised by circuit noise, but makes inadvertent overload more likely, and vice versa. The levels chosen depend on the purpose of the console. If you are recording you only have to get it right once, but you do have to get it exactly right, i.e. with the best possible signal-to-noise ratio, so the internal level is relatively high, very often −2 dBu (615 mVrms), which gives a headroom of about 22 dB. In broadcast work to air you only have one chance to get it right, and a mildly impaired signal-to-noise ratio is much preferable to a crunching overload, especially since the FM or DAB channel has limited noise performance anyway, so the internal levels are considerably lower. The Neve 51 Series broadcast consoles used −16 dBu (123 mVrms), which gives a much increased headroom of 36 dB at the cost of noise performance. At least one manufacturer has used balanced mixing buses to reduce the impact of the lower internal levels on signal-to-noise ratio.

As a sidelight on this issue, I might mention a mixer design I worked on some years ago. The decision was taken (not by me) to use an internal level of −6 dBu instead of the usual −2 dBu to improve headroom, the assumption being that a 4 dB deterioration in the internal noise would go unnoticed. It most certainly did not, and vociferous customer complaints led to a rapid redesign to convert to the old −2 dBu level. Fortunately we were ready for this one, and the design had been arranged so that the internal level could be changed simply by changing resistor values.

An internal level of −2 dBu does have one unique advantage: doubling the level, as is done by most forms of balanced output stage, gets you directly to the professional +4 dBu level.

www.raguvaran.puzl.com

Mixer Architecture

A mixing console has a number of input channels that can be summed into groups, and these groups again can be summed into a stereo mix as the final result. Small mixers often have no groups and mix directly to stereo.

The structure and signal flow of a mixing console depends very much on its purpose. A recording console needs track return sections to allow the creation of a monitor mix so that as a track is recorded, those already done can be played back at roughly the same levels they will have in the final mix; this is essential so new material can be synchronized with that already in existence. While it is quite possible to use a recording-oriented console for PA work, the reverse is much more difficult, and usually impractical.

A sound-reinforcement (PA) console does not need monitor sections, so it can be somewhat simpler, but it typically is fitted with a large number of effects returns. The output from a PA console at a large event is not likely to be in stereo – there will be multiple speakers, each of which requires an individual feed with different equalization and delay requirements; often there is a center output.

In PA work there are often two mixers. The front-of-house (FOH) mixer usually resides in the middle of the audience or in a control room at the back. If the mixer is in the audience, the less room it takes up the better, as it is occupying seats that could have been paid for. Size is also very much an issue for mixers intended for installation in a control room, as access to these is often cramped and difficult.

The second type of PA mixer is the so-called monitor mixer. PA work makes much use of 'monitors', i.e. small speakers, often wedge-shaped, that are placed close to the musician so that he/she can hear their output clearly without straining to pick it out of the main mix. Creating these foldback feeds on a separate mixer, where basically each channel consists of a bank of aux sends (from 12 to 32 may be available), gives much greater flexibility than relying on the relatively few aux sends available on the FOH desk. The inputs to the monitor mixers are taken from the inputs to the FOH console by a splitter box, usually transformer based, at the stage end of the 'snake' cable to the FOH console. Traditionally monitor mixers were placed just out of sight at the side of the stage, so that disgruntled performers could signal to the hapless monitor engineer that the foldback mix was not to their liking. The widespread adoption of intercoms has given more flexibility in monitor desk placement.

Broadcast consoles are more specialized in their layout and no one expects to be able to use them for recording or PA applications.

Recording consoles come in two different formats. The first and more traditional is called a split, separate-group, or European-style arrangement and has input channels arranged

www.raguvaran.puzl.com

in one bank, typically to the left of the master section, with the groups banked to the right of it; on a large console there may be a further set of channels to the right of the groups.

The second format is called the in-line arrangement, and has the channels and groups built into the same module. Normally there is a row of channel/group modules to both the left and right of a central master section. In-line consoles are more compact and make more efficient use of the electronic facilities, but they are conceptually more complex and somewhat harder to use. They are not popular in PA work where you need to do things quickly in real time without juggling a block diagram in your head.

In the sections that follow we will look at the internals of generic mixers that do not follow any particular design but are configured to bring out as many instructive points as possible.

The Split Mixing Architecture

Figure 12.1 shows the basic arrangement of a split (separate-group) mixer. Its design is oriented to recording, but like many small and medium-sized mixers it could also be used effectively for PA work.

Each channel has both a microphone and line-level input, with a wide-ranging gain control. Once it has been raised to the nominal internal level of the console, the signal is processed by the equalization section, controlled in level by the channel fader, and then routed to a group or pair of groups, or directly to the stereo L–R bus. The position of the channel signal in the overall stereo image is set by the panoramic potentiometer (universally known as a panpot) that controls the proportion of the signal sent to left and right. The channel also has auxiliary send controls that feed the channel signal to the auxiliary buses; sends taken off before the fader are called prefade and are normally used for foldback, i.e. helping the musician hear the sound they are producing. Sends taken off after the fader, so their level is dependent on the fader position, are called postfade sends and are normally used for adding effects such as reverberation that are expected to vary in level with the original signal.

The final feed from the channel to the buses is the prefade listen (PFL); when the PFL switch is pressed the prefade channel signal is sent to the master module via the PFL bus, where it automatically replaces the normal stereo mix feed to the control-room monitor (CRM) loudspeakers.

The group/monitor modules contain both the group section and one or two monitor sections. The group consists of a summing amplifier that collects together all the signals sent to its group bus, and passes it to the outside world via a fader and an output amplifier. When recording, this group signal is sent to one track of the recording machine. The monitor sections allow tracks

www.raguvaran.puzl.com

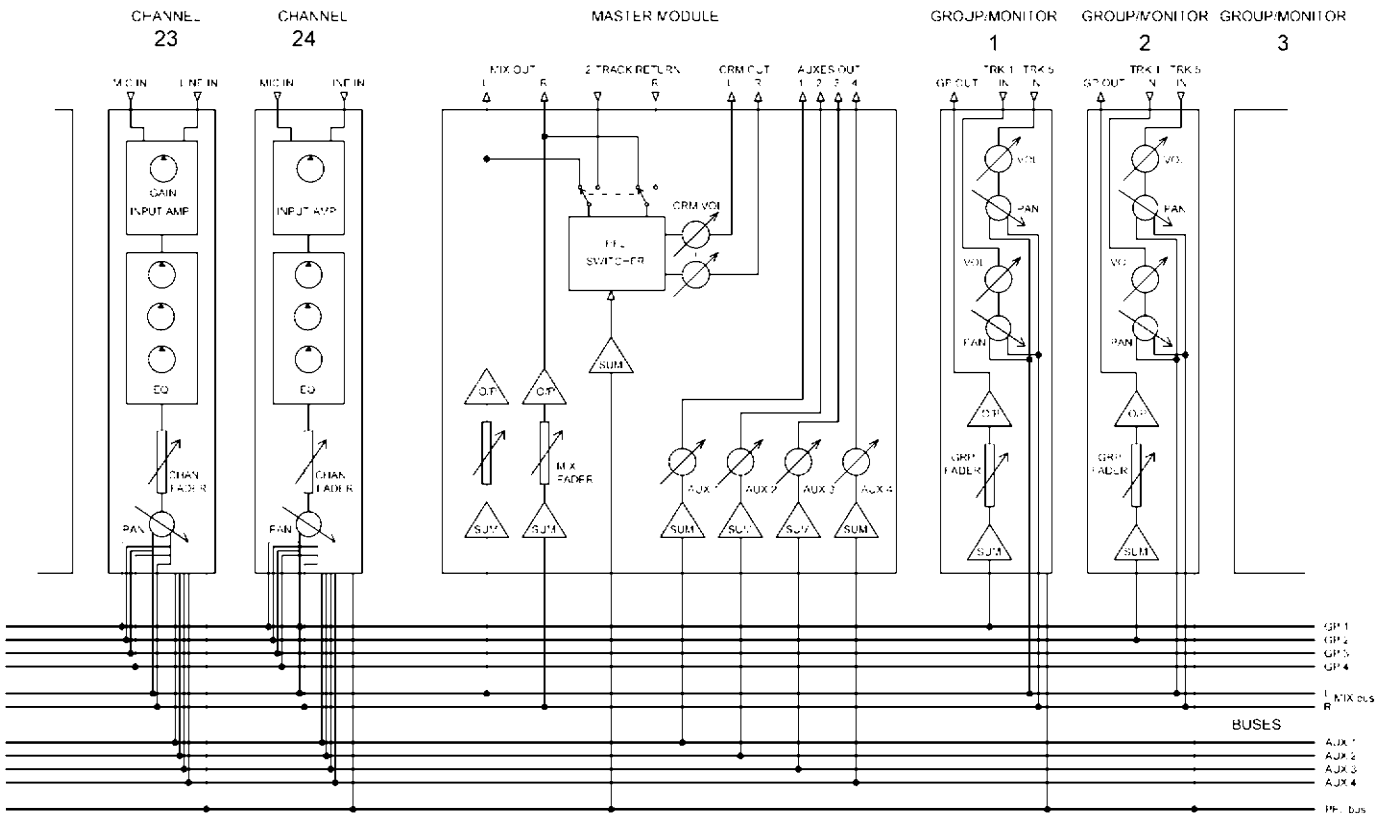


Figure 12.1: Block diagram of a split-format mixer with four groups and four aux sends

www.raguvaran.puzl.com

that are already recorded to be played back to the stereo mix bus via a level control and a panpot, so that a rough idea of what the final mix will sound like can be set up as early in the recording process as possible. One of the monitor sections will be switchable between the group output and the track return, so the monitor mix can be made up of tracks that are being recorded as well as those that have already been laid down. This group/track switching is not shown in Figure 12.1 to aid clarity, but is fully shown in Figure 12.4 below. Lack of space means that the monitor return sections usually have only simple EQ or none at all, but they almost always have a prefade aux send for foldback purposes. A postfade send is also extremely useful as the rough monitor mix is much more convincing if reverberation can be added as appropriate; this is called 'wet' monitoring. The monitor sections often each have a PFL switch.

When recording is finished, and mixdown begins, the full facilities of the channels are required, and so the track return signal is now sent to the channel line input and the channel routed to the stereo mix bus. This is often a connection normalled through the line-input jack, so that the input amplifier receives the track return signal at all times unless a jack is inserted. Back in the days of tape machines this facility was called 'tape normalling' but now 'track return normalling' describes it rather better. The monitor sections in the group/monitor modules are now no longer needed, but since they are routed to the stereo mix bus they can be used as extra effects returns.

The master section contains the summing amplifiers for the stereo mix bus, which feed a stereo fader and suitable output amplifiers. These provide the final feed to the two-track recording device at mixdown. Normally this stereo output is also sent to a pair of meters and the control-room monitor (CRM) loudspeakers, the latter via a volume control, but the meters and CRM speakers can also be switched to monitor the return from the two-track recorder. When any PFL switch on the console is pressed, this is detected and the PFL signal (in mono) is fed to the meters and CRM speakers instead, so the signal quality and level from any channel can be assessed.

The master section also contains the summing and level controls for the aux send buses, and may also incorporate talkback facilities and a line-up oscillator; these features are dealt with in more detail later in this chapter.

The In-Line Mixing Architecture

Figure 12.2 shows the basic arrangement of an in-line mixer, essentially intended for recording. There are now no group/monitor modules, their functions being performed by what might be called channel/group/monitor modules. Only four group buses are shown for clarity, but in practice there would normally be a much higher number, to make use of the in-line format.

www.raguvaran.puzl.com

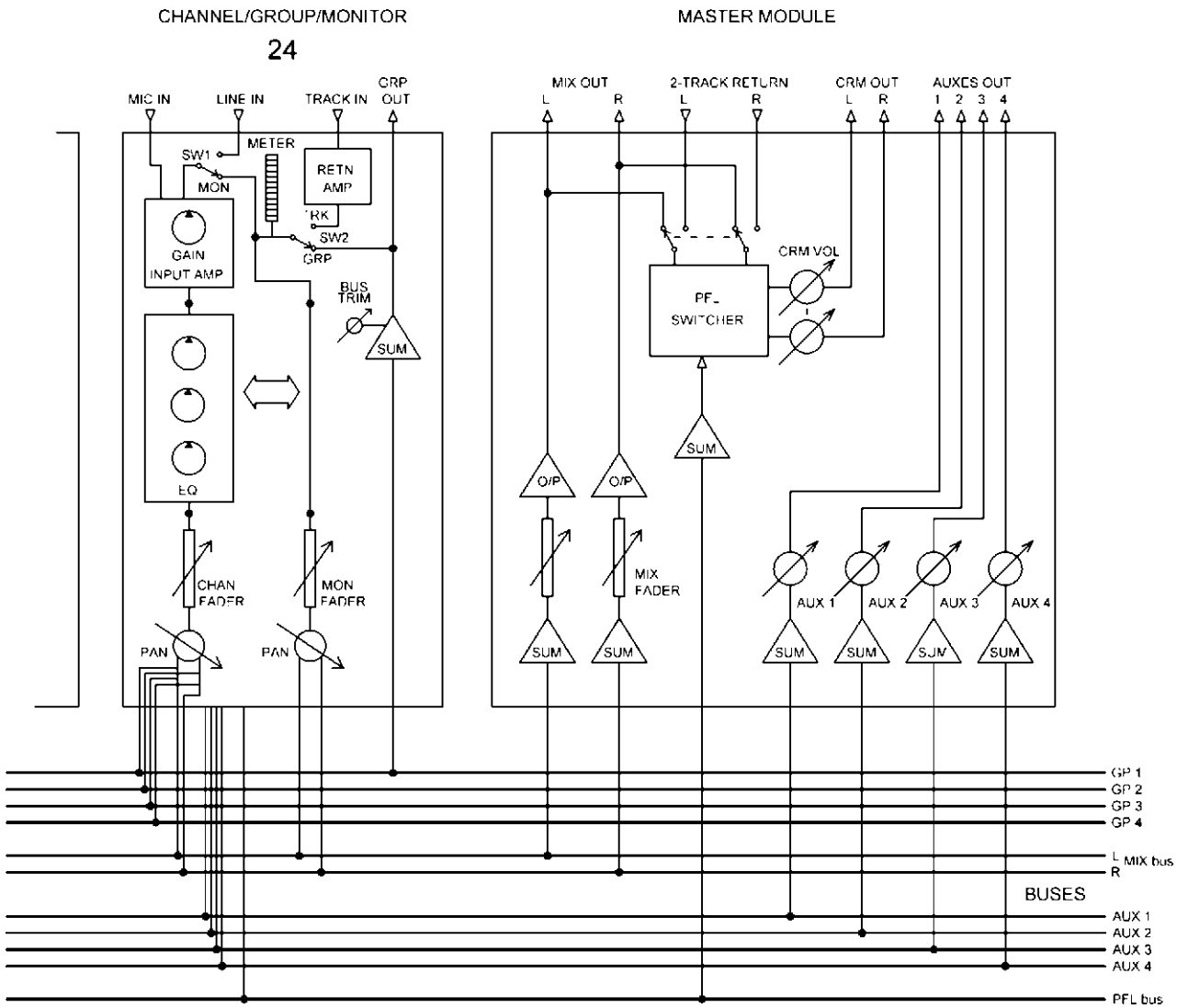


Figure 12.2: Block diagram of an in-line mixer with *N* groups and four aux sends

www.raguvaran.puzl.com

The group summing amplifier is now part of the channel. It no longer has its own fader; instead the summing amp gain is controlled by a rotary 'bus trim' control so overload can be avoided. This is normally placed out of the way at the very top of the module. The group signal is sent out to the recording machine, and the track return comes back in, usually via a balanced amplifier. A track/group switch SW2 selects either the group or the track return signal for the metering and the monitor path; as shown in the figure it consists only of a monitor fader and a monitor panpot, which send the signal to the stereo mix bus to create a monitor mix. The monitor fader is normally a short fader mounted on the channel, as opposed to the long channel fader, which is right at the front of the console. Sometimes a rotary monitor fader is fitted.

There are two slightly differing approaches to the operation of an in-line module:

In the first approach, at mixdown switch SW1 now selects the track return signal rather than the line input, and the signal reaches the stereo mix bus via the full facilities of the channel. Sometimes additional switching allows the monitor path to be used as an extra effects return, as in split consoles. Typically the line-input jack is pressed into service as the input connector for this purpose.

An alternative design approach is not to switch the source of the channel path to the track return at mixdown, but instead to move the whole EQ section from the channel to the monitor path. This means that the short monitor fader would have to be used for level adjustment, which is not ideal when there is a long fader right in front of you. This has particular force when fader automation is fitted, for it is normally only applied to the long fader. A switch called 'fader flip' is therefore fitted, which swaps over the function of the two faders, allowing the long fader to be used at mixdown. This sort of approach involves a lot of reconfiguration when switching from record to mixdown mode, and on the more sophisticated consoles this is handled entirely by electronic switching, with global control from the master section, so that only one switch need be pressed to change the console mode. This second method always seems to me to be unnecessarily complex, but it has seen extensive use.

The master section shown in Figure 12.2 is identical to that shown for the split console in Figure 12.1.

A Closer Look at Split-Format Modules

We will now look in more detail at the channel, group, and master modules that make up a split mixing console.

The Channel Module (Split Format)

Figure 12.3 shows a typical input channel for a relatively simple split-format mixing console. The input stage provides switchable balanced mic and line inputs; the mic input has an

www.raguvaran.puzl.com

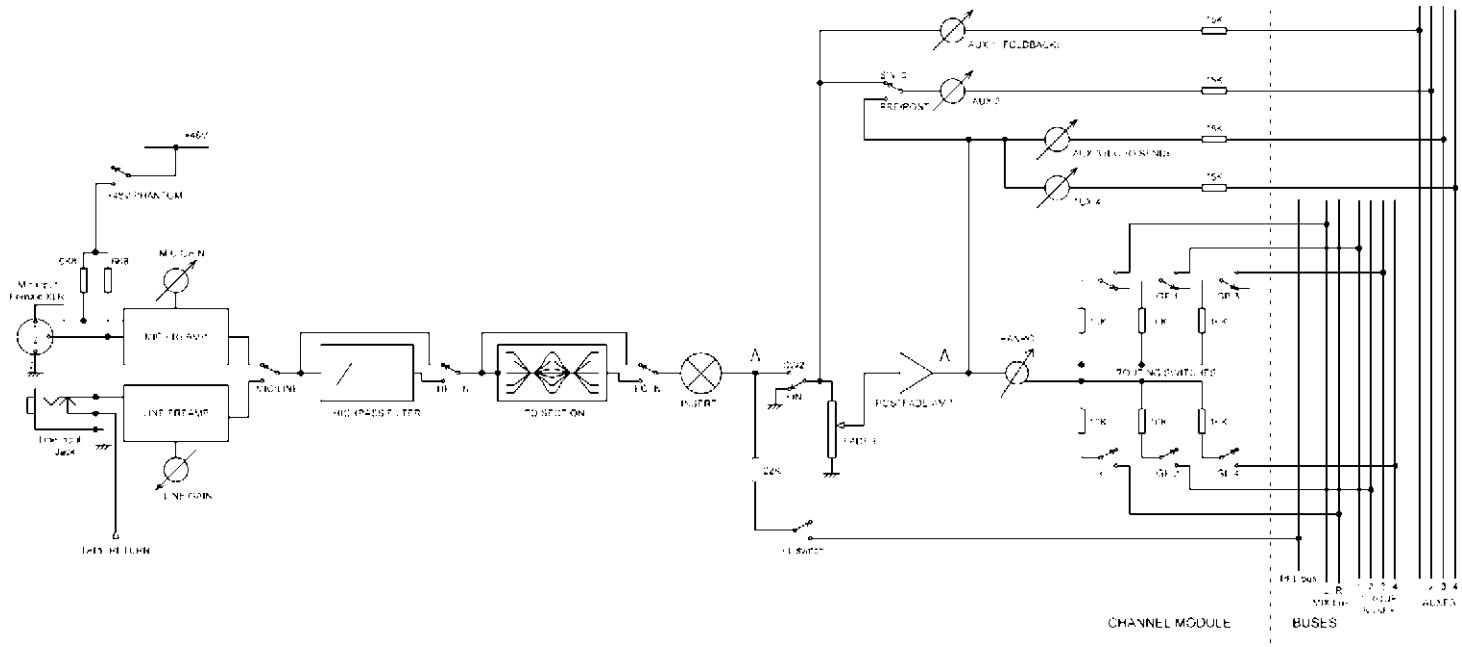


Figure 12.3: Block diagram of a typical channel module for a small mixer

www.raguvaran.puzl.com

impedance of 1–2 kΩ, which provides appropriate loading for a 200 Ω mic capsule. If a +48 V phantom power facility to power microphones is provided, it is independently switchable on each channel. The line input will have a bridging impedance of not less than 10 kΩ. The mic preamplifier in particular will have a wide range of gain, such as 0–70 dB, while the line input tends to have a more restricted range, such as +20 to −10 dB. The track return from the recording machine is shown connected through to the line input via the normalling contact of the line jack socket, so that mixdown with the full channel facilities can begin as soon as the mic/line switch is set to 'line', so long as no jack has been inserted into the line input to break the normalling contact. The line input is shown as unbalanced in Figure 12.3 for clarity, but in practice it would usually be a balanced input using the tip and ring connections of the jack socket.

The input stage is followed immediately by a switchable high-pass filter (usually −3 dB at 100 Hz) to remove low-frequency disturbances picked up by the microphone as soon as possible; if not filtered out these can eat up headroom. The filter is usually second or third order, giving a roll-off of 12 or 18 dB/octave respectively.

The tone-control section (universally known in the pro audio business as 'EQ' or equalization) typically includes one or more mid-band resonance controls as well as the usual shelving Baxandall-type high and low controls. On all but the simplest mixers the EQ section can be switched out to allow before/after comparisons.

The larger and more sophisticated consoles incorporate dynamics sections into each channel. This is not shown in Figure 12.3 to aid clarity. The dynamics facilities available vary but usually include compression, limiting, and sometimes noise gating; some consoles have been produced with just the noise-gate function, as it is easier to fit the required electronics into a limited space. A perennial problem with this sort of thing is finding panel room for the extra controls required; the permissible length of a module is limited by the reach of the human arm.

Next comes the insert point, though in some designs it may be configured so it can be placed ahead of the EQ and dynamics section instead. This is a jack or pair of jacks that allow external processing units to be plugged into the signal path. When nothing is plugged in, the 'normalling' contacts on the jack socket allow the signal to flow through unchanged.

The PFL switch routes the post-insert signal to the master module and the monitor speakers independently of all other controls; a PFL detect bus signals the master module to switch the studio monitoring speakers from the normal stereo mix bus to the PFL bus.

Next in the chain comes the channel on switch. This may be either a simple mechanical switch or an electronic mute block. Note the PFL feed is taken off *before* the On switch so the channel signal is always accessible. The channel level in the mix is controlled by a linear fader with a postfade amplifier typically giving 10 dB of gain, and the panpot sets the stereo positioning, odd group numbers being treated as left and even as right.

www.raguvaran.puzl.com

The channel shown has three auxiliary sends. The auxiliary sends of a console represent an extra mixing system that works independently of the main groups; the number and configuration of these sends have a large effect in determining the overall versatility of the console. Each send control provides a feed to a console-wide bus; this is centrally summed and then sent out of the console. Sends come essentially in two kinds. Prefade sends are taken from before the main channel fader and are therefore independent of its setting. Postfade sends take their feed from after the fader, so that the send level falls or rises according to the fader setting.

Prefade sends are normally used for 'foldback', i.e. sending the artist a headphone feed of what he/she is perpetrating, which is important if electronic manipulation is part of the creative process, and essential if the artist is adding extra material that must be in time with that already recorded. In the latter case, the existing tracks are played back to the artist via the prefade sends on the monitor sections.

Postfade sends are used as effects sends; their source is after the fader, so that the effect will be faded down at the same rate as the untreated signal, maintaining the same ratio. The sum of all feeds to a given bus is sent to an external effects unit and the output of this returned to the console. This allows many channels to share one expensive device, such as a high-quality digital reverb, and for this sort of purpose is much more appropriate than patching processors into the channel insert points.

There may be anything from one to 12 or more sends available. In the example shown, the first send (AUX 1) is a dedicated prefade send, which is typically used for foldback. The third (AUX 3) is a dedicated postfade send, typically used for effects such as reverb. The second send (AUX 2) can be switched so it is either pre- or postfade. In a recording console the greatest need for foldback sends is during track laying, while effect sends are in greatest demand at mixdown; it is therefore common for some or all of the sends on a channel to be switchable pre- or postfade.

On more sophisticated consoles, it is often possible to switch every auxiliary send between pre and post to give maximum flexibility. Traditionally, this meant pressing a switch on every input module, since it is most unlikely that a mixture of pre and post sends on the same bus would be useful for anything. On a 64-input console this is a laborious process. More advanced designs minimize the effort by setting pre/post selection for each auxiliary bus with a master switch that controls solid-state pre/post switching in each module. An example of this approach is the Soundcraft 3200 recording console.

Effects Return Modules

In complex mixdowns it may be necessary to return a large number of effects to the mix. While, as described above, it is often possible to press unused monitor sections or channel modules into service as effect returns, sometimes this just does not provide enough, and specialized 'effects return' modules may be fitted. These usually have facilities intermediate

www.raguvaran.puzl.com

between channels and monitor sections, and it is common to fit two, and sometimes even four, into the space occupied by a channel module. The returned effect, which may well now be in stereo, such as the output of a digital reverb unit for example, is usually added to the stereo mix bus via level and pan controls.

The Group Module

Figure 12.4 shows a typical group module for a recording desk. Each mix bus has its own summing amplifier; the summed group signal is then sent to an insert so external processing can be applied. The insert send is very often arranged for ground-canceling operation, for the simple reason that the summing amp inherently phase-inverts the signal, which must be returned to the correct polarity by another amplifier stage before it meets the outside world. This second stage can be arranged to be ground-canceling at minimal expense. The operation of ground-canceling outputs is described in Chapter 15.

The signal from the insert return goes to the group fader. Like that of the channel, it is a linear fader with a postfade amplifier typically giving 10 dB of gain. The signal then goes to an output amplifier suitable for driving the recording device, possibly via lengthy cables. On all but the cheapest designs this is a balanced output.

The second major function of the typical group module is to allow the creation of a monitor mix so that as a track is recorded, those already done can be played back at roughly the same levels they will have in the final mix; this is obviously essential to allow new material to be synchronized with that already in existence. Back in the days of tape recording machines, the monitor playback had to be done by the record head, as the playback head was physically displaced and so would have given a delayed signal; this meant the quality of the monitor mix was compromised as the record head was optimized for recording and not playback. With digital recording this is of course not a problem.

The monitor section typically consists of a balanced line input to prevent ground-loop problems with the recorder, some form of switching between the group and recorder return, so the track being recorded can be incorporated into the monitor mix, and level and panpot controls. EQ is not often fitted for reasons of space, but aux sends for foldback and effects are standard, because adding reverb to the monitor mix makes it much more realistic. A PFL switch for each monitor section is usually provided.

Normally relatively few tracks are recorded at a time, so it is sensible to have more monitor sections than groups. For example, a four-group mixer can be very effectively used with an eight-track recorder if group 1 is connected to track 1 and track 5, group 2 is connected to track 2 and track 6, and so on. The track to be recorded is selected at the recorder. To make use of this each group module will contain one actual group section and two monitor sections, which explains why there is not normally room for even basic EQ in the monitor sections.

www.raguvaran.puzl.com

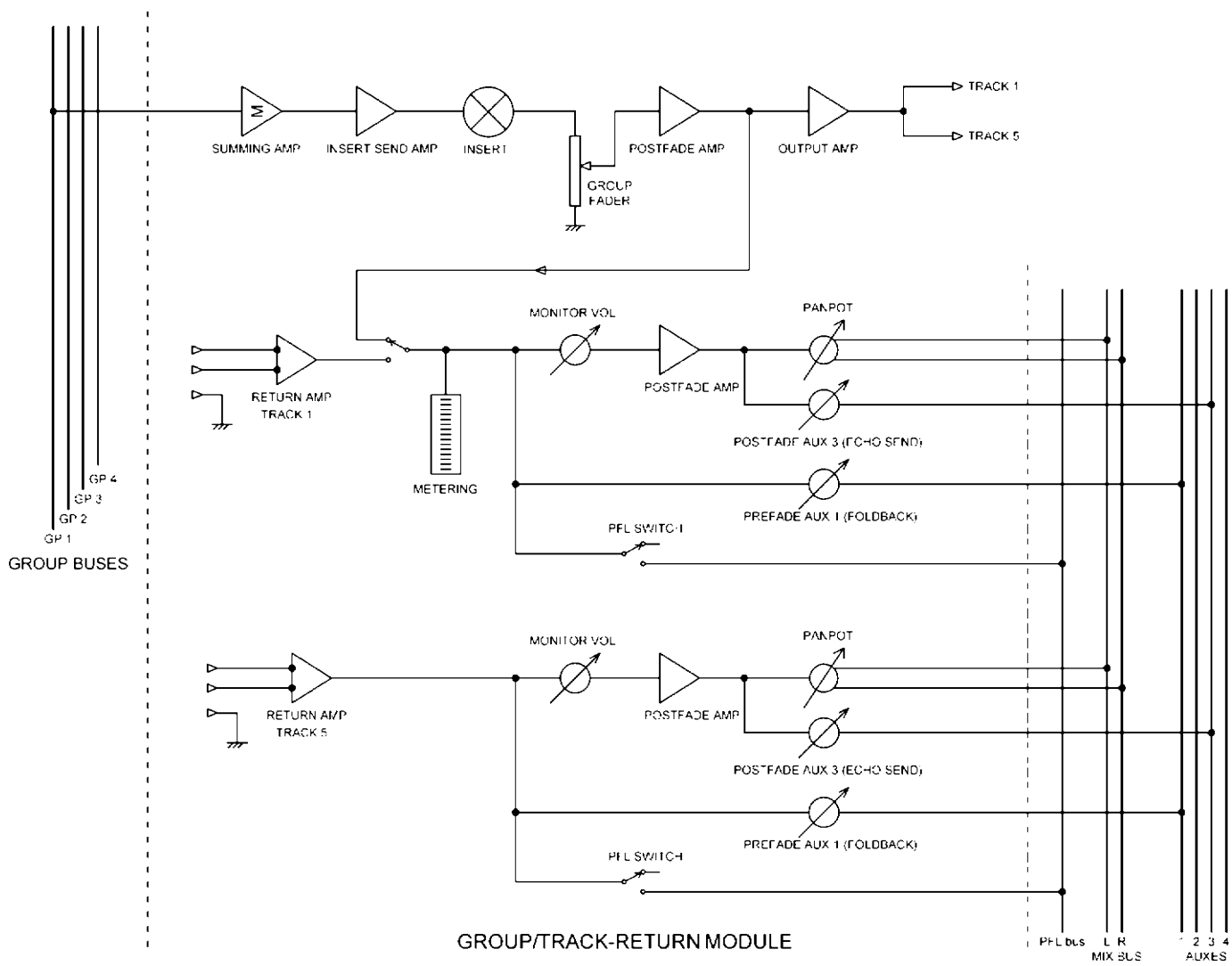


Figure 12.4: Block diagram of a typical group module for a small recording mixer

www.raguvaran.puzl.com

The Master Module

Figure 12.5 shows the basics of a typical master section. It contains the summing amplifiers for the stereo mix bus, with their associated insert send amplifiers and insert points. This is followed by a stereo fader (sometimes implemented as two mono faders with the knobs

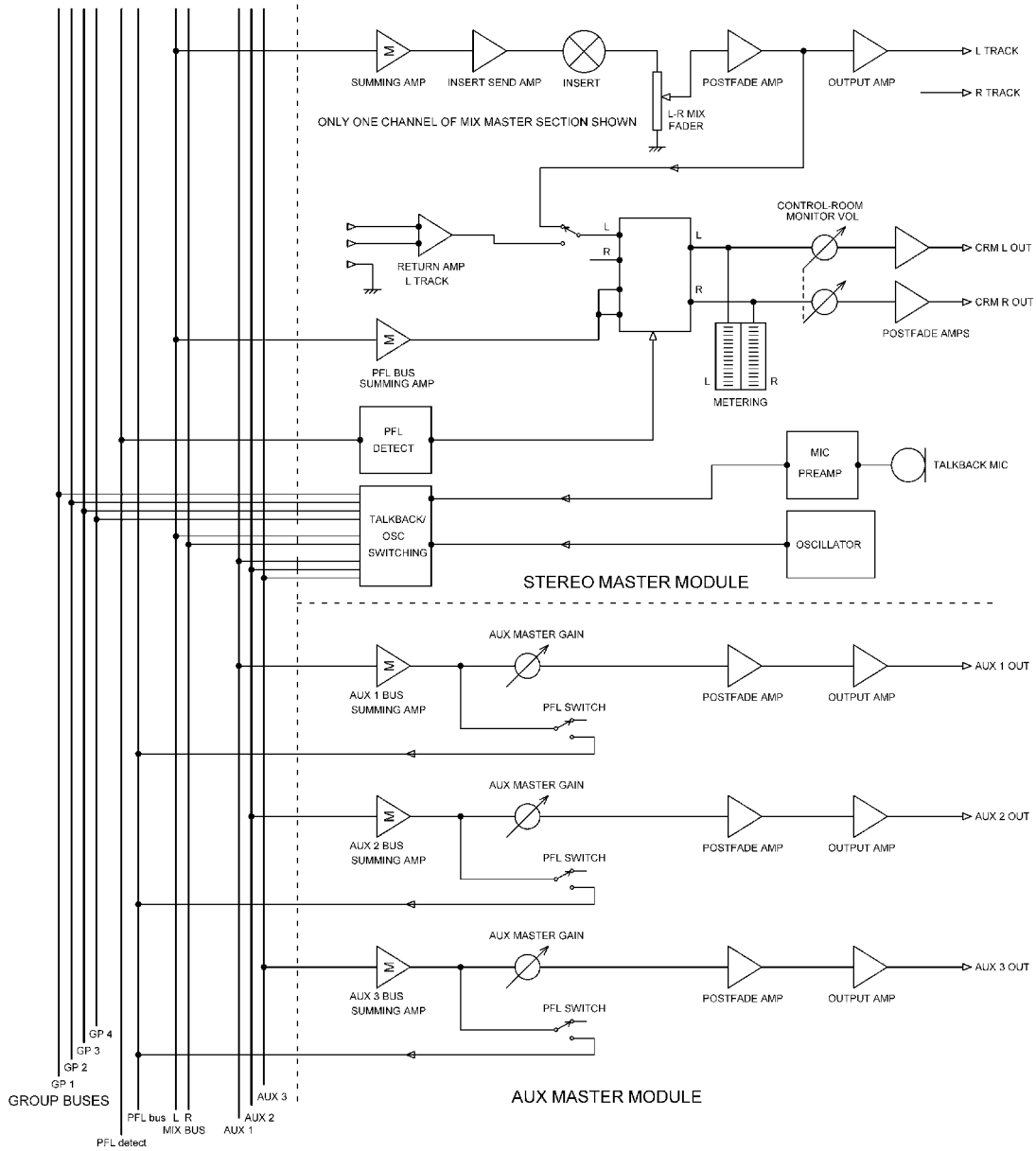


Figure 12.5: Block diagram of a typical master module for a small mixer

www.raguvaran.puzl.com

mounted immediately next to each other), a +10 dB post amp, and a balanced or ground-canceling output stage, which feeds the two-track recorder during mixdown.

The stereo mix bus normally drives the L–R meters and the CRM loudspeakers, but a manual source-select switch allows this feed to be replaced by the return signal from the two-track recorder for quality checking of the final stereo recording. Whenever a PFL switch anywhere on the console is pressed, the PFL detect system responds and activates two solid-state switches that replace the stereo monitor signal with the PFL signal. The L–R meter feeds are taken off after the PFL switch, so those meters can be used to check the level at whichever point in the system the PFL switch in question happens to be.

The master section also contains the summing and level controls for the aux send buses. The auxes may have dedicated meters or just a PFL switch each.

The master module will also carry any master status switches for globally changing things such as group/recorder switching for the monitor sections, pre/post operation of auxes, and so on.

Talkback and Oscillator Systems

The talkback system allows the mixing engineer to talk to the musicians in the studio, or add spoken comments to the recording, using a microphone mounted on the console. Back in the day it was considered cool to mount the talkback microphone on a flexible gooseneck, but these get in the way and the modern approach is to have a small electret microphone mounted flush with the master module panel.

For talkback to the studio the microphone feed is routed to the aux buses, on the assumption that one or more of these will be in use for foldback and so will be connected to a loudspeaker in the studio area. Routing may also be provided to a dedicated talkback loudspeaker. In many cases there is a facility to route to the first two auxes only, as these will almost certainly be used for foldback purposes (see the typical system in Figure 12.6).

For recording identification the microphone feed is routed to the group buses. This facility, which allows the engineer to identify a recording by saying something like ‘Spinal Tap – take 147’, is often called a ‘Slate’ facility, by analogy with the film industry.

When the talkback facility is used, there is a danger of acoustical feedback. If a microphone in the studio is active and routed through channel and group to the control-room monitors, the monitor speaker output will be picked up by the talkback microphone, fed to the studio … and away she goes. It is therefore normally arranged that pressing any of the talkback switches attenuates or cuts completely the monitor speaker signal. On larger consoles the amount of attenuation (or ‘dim’ as it is usually called in this context) is adjustable.

www.raguvaran.puzl.com

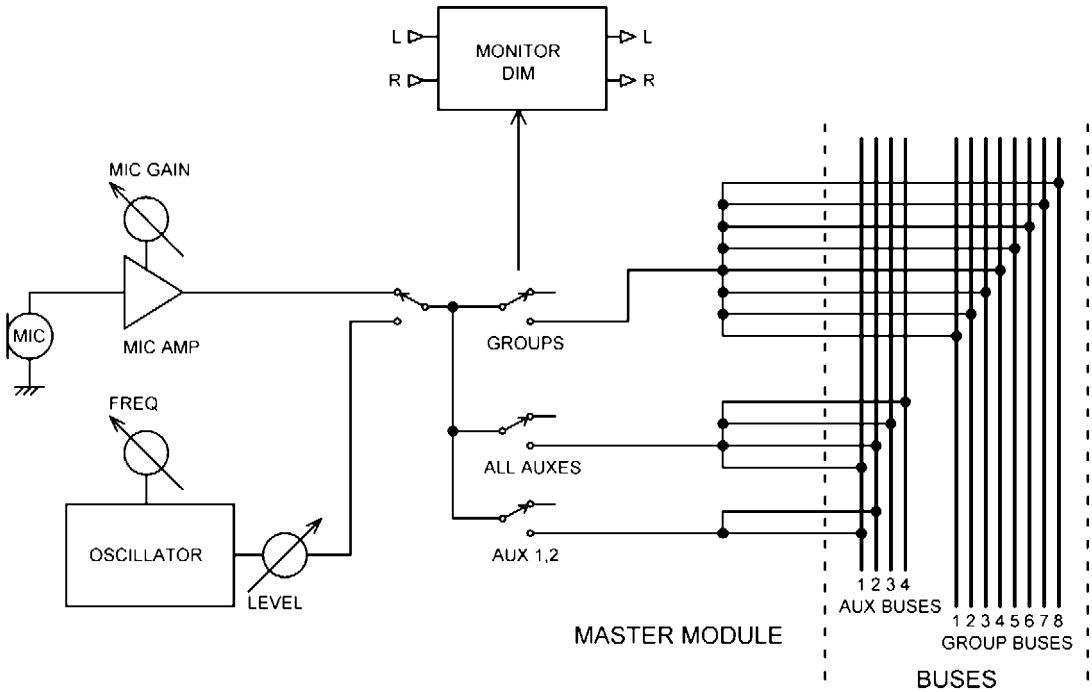


Figure 12.6: Talkback and oscillator system with routing to aux and group buses

If the talkback microphone is mounted in the master panel, the distance between it and the engineer will vary as he moves. More sophisticated consoles sometimes incorporate a talkback compressor to reduce the level variations.

Many of the larger recording consoles include a 'Listen' facility whereby the control room can hear messages from the studio even if no microphone channels are faded up. A microphone is mounted somewhere out of the way (often suspended from the ceiling) and is routed to the monitors when the 'Listen' switch is pressed.

Most mixers of medium size and above incorporate a line-up oscillator that can be routed to all the groups and the stereo mix. This allows the mixer metering to be lined up with external level indication on the recorder. A fixed-frequency oscillator running at 1 kHz with a variable output level is the minimum facility; more advanced models have switched or fully variable frequency controls.

In the typical talkback/oscillator system of Figure 12.6, the line-up oscillator is shown sharing with the talkback system the bank of mix resistors that inject signal into the buses. On a large console this is quite a lot of resistors, and it would not be sensible to design them in twice. The only compromise is that it is not possible to use talkback and the oscillator at the same time,

www.raguvaran.puzl.com

but then why would you want to? The default selection is talkback, because this is used much more frequently than the oscillator.

The In-Line Channel Module

Figure 12.7 shows a typical channel/group/monitor module for an in-line mixing console. This uses the first approach to in-line operation described above, where to enter mixdown mode the source of the channel path is switched to the track return, rather than bodily moving the EQ and other facilities into the monitor path.

The group fader no longer appears. More usefully, the summing amplifier gain can be varied by a rotary 'bus trim' control to prevent overload; it is highly desirable that this control alters the gain of the actual summing amplifier, rather than having a low-gain summing amplifier with a level control and post amplifier following it, because the former gives greater protection against clipping, and superior noise performance at low gain settings. The bus trim is rarely altered so it can be conveniently placed out of the way at the very top of the in-line module. The group signal from the summing amplifier is sent out to the recording machine, usually by means of a balanced or ground-canceling output stage to prevent ground loops occurring with the multi-track recorder. The track return usually comes back via a balanced amplifier for the same reason. A track-return/group switch selects either the group or the track return signal for the metering and the monitor path; as shown in the figure this path consists of a monitor fader (the 'short fader') and a monitor panpot, which send the signal to the stereo mix bus to create a monitor mix. Note that every in-line module will have a meter, though it is not usual to provide meters for every channel on a split console. A monitor PFL switch is provided so that an individual track return or group can be conveniently listened to.

As in split consoles, it is convenient to connect the track return to the channel line input through normalling contacts on the line-input jack, so switching to mixdown requires the minimum number of operations. As described earlier, in this form of in-line configuration the monitor path is redundant, as the channel fader and panpot are used to control the signal coming back from the multi-track recorder, and so there is usually a way in which the monitor path can be used as an extra effects return, as is commonly done in split consoles. Typically the line-input jack is used as the input connector for this purpose; alternatively an extra insert point may be provided in the monitor path, into which an external line signal can be fed. These arrangements are not shown in Figure 12.7, which is quite complicated enough already; it is one of the drawbacks of the in-line system that it is conceptually more complex than the split console format.

The auxiliary send system is slightly more complex in an in-line module, in the interests of maximum flexibility of working. In the example shown, the sends can be switched in pairs to

www.raguvaran.puzl.com

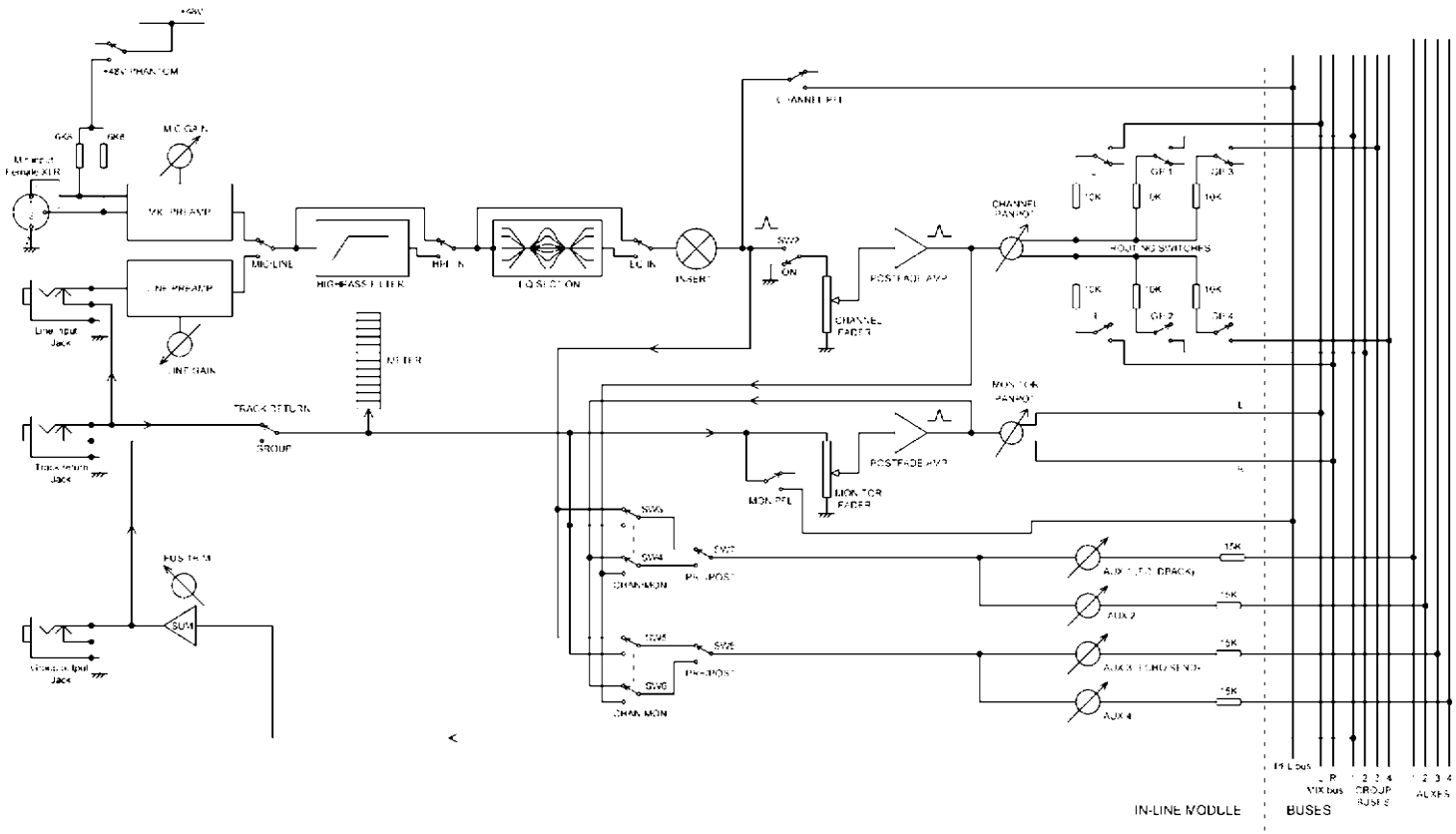


Figure 12.7: A channel/group/monitor module for an in-line recording console

www.raguvaran.puzl.com

take their signal either pre- or postfade from either the channel or monitor path. This is implemented by means of the switches SW3–SW8. For example, during recording the effect sends can be switched to be monitor postfade, allowing wet monitoring. Many variations on this aux send are possible; on larger consoles there will be six or eight sends, and sometimes a stereo aux send with its own panpot is provided.

www.raguvaran.puzl.com

This page intentionally left blank

www.raguvaran.puzl.com

CHAPTER 13

Microphone Preamplifiers

Microphone Preamplifier Requirements

A microphone preamplifier is a serious design challenge. It must provide a gain variable from 0 to 70 dB or more to amplify deafening drum-kits or discrete dulcimers, generate minimal internal noise, and have a high common-mode rejection ratio (CMRR) balanced input to cancel noise pickup in long cables. It must also be able to withstand +48 V DC of phantom power suddenly applied to its input while handling microvolt signals.

It is now rare to use input transformers to match a low-impedance (150–200 Ω) microphone to the preamplifier, since the cost and weight penalty is serious, especially when linearity at low frequencies and high levels is important. Both dynamic and capacitor microphones have a low output impedance of this order, the former because of the low number of turns used in the coil, the latter because active buffering of the extremely high capsule impedance is essential.

The low-noise requirement rules out the direct use of op-amps, since their design involves compromises that make them at least 10 dB noisier than discrete transistors when faced with low source impedances. The answer, for at least the last 30 years, has been to use hybrid input stages that combine discrete input transistors to give low noise, combined with op-amps to provide raw open-loop gain for linearization, and load-driving capability.

The requirements of a microphone preamplifier can be summarized as follows:

1. Variable gain is necessary, usually from 0 to +70 dB. Some designs have a gain range extending to +80 dB. The bottom 20 dB section of the range is often accessed by switching in a 20 dB input attenuator.

2. Minimal noise must be produced. Taking the source impedance of a microphone as 200 Ω, the Johnson noise from that resistance is −129.6 dBu at 25°C, 20 kHz bandwidth. This puts an immediate limit on the noise performance; if you set up 70 dB of gain then the noise output from the preamplifier will be −59.6 dBu even if it is itself completely noiseless. Most mic preamplifiers approach this at maximum gain, often having a noise figure as low as 1 or 2 dB, but depart from it further and further as gain is reduced. In other words, the noise output does not fall as fast as it would if the preamp was noiseless, because reducing

Small Signal Audio Design; ISBN: 9780240521770
Copyright © 2010 Elsevier Ltd; All rights of reproduction, in any form, reserved.

www.raguvaran.puzl.com

the gain means increasing the effective resistance of the gain control network, so it creates more Johnson noise.

3. The input must have a high CMRR to reject interference and ground noise. The CMRR should ideally be high, be flat with frequency, and remain high as the input gain is altered over the whole range. In practice, CMRR tends to worsen as gain is reduced. Because of the need for a balanced input, microphone inputs are almost always female XLR connectors.

4. The input must have a constant resistive input impedance of 1–2 kΩ, which provides appropriate loading for a 200 Ω dynamic microphone capsule. This is also a suitable load for the internal head amplifiers of capacitor microphones.

5. The input must be proofed against the sudden application (or removal) of +48 V DC phantom power. It must withstand this for many repeated cycles over the life of the equipment.

Transformer Microphone Inputs

For a long time transformer microphone inputs were the only real option. The cost and weight of a microphone input transformer for every channel are considerable, especially at the quality end of the market, because a large transformer core is needed to handle high levels at low frequency without distortion. Because of the low signal levels, mu-metal screening cans were normally used to minimize magnetic interference, and this added to the cost and weight.

Step-up ratios from 1:5 to 1:10 were used, the higher ratios giving a better noise performance but more difficulties with frequency response because of the greater capacitance of a larger secondary winding. The impedance reflected to the input depends on the square of the step-up ratio, so with a 1:10 transformer the secondary had to be loaded with 100 kΩ to get the desired 1 kΩ at the primary input.

The arrangement in Figure 13.1 has an amplifier gain range from 0.3 to 60 dB, to which is added 20 dB from the transformer ratio of 1:10, giving +20 to +80 dB of overall gain. R6 in conjunction with the reverse-log pot shapes the control law so it is roughly logarithmic. R4 is the 100 kΩ secondary loading resistor, and C1, R3 form a Zobel network to damp the secondary resonance. In the mid-1970s, when op-amps were still a dubious proposition for quality audio, the input amplifiers were very often discrete transistor stages using four or more transistors. Note R5, which prevents the gain from being reduced to exactly unity. Its presence is testimony to the fact that the discrete preamplifiers were difficult to stabilize at HF if the output and non-inverting input were directly connected together. R1 and R2 feed in phantom power when the switch is in the On position.

www.raguvaran.puzl.com

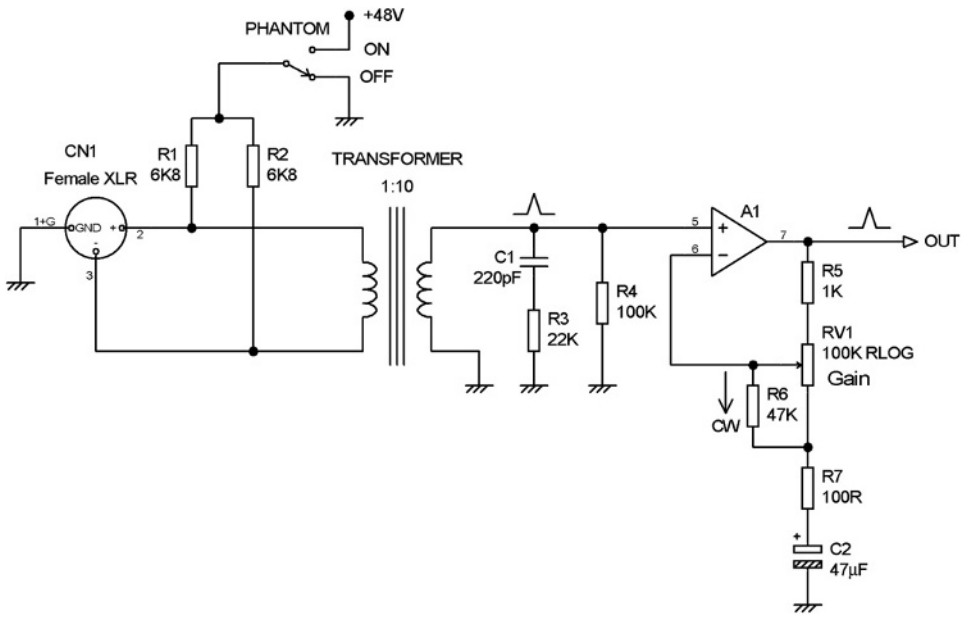


Figure 13.1: A transformer microphone preamplifier. Gain range +20 to +80 dB

Microphone transformer technology had, and still has, its own advantages. A balanced input with a good and constant CMRR was inherent. The step-up ratio, which gave 'gain for free' in electronic if not financial terms, and the associated impedance matching, meant that it was easy to design a quiet input stage. Transformers also gave good RF rejection, and isolated the input preamplifier from phantom power voltages. For EMC reasons in particular, transformer microphones continued in use in broadcast mixing consoles long after they had disappeared from recording and PA equipment.

The Simple Hybrid Microphone Preamplifier

The cost incentive to develop an effective low-noise transformerless microphone input was considerable, and after much experimentation the arrangement shown here in Figure 13.2 became pretty much standard for some years, being extensively used in mixers around the period 1978–1984 (in this period the more expensive consoles stuck to transformer microphone amplifiers). The difficulties of getting the noise low enough and the linearity good enough were at first formidable, but after a good deal of work (some of which I did) such stages have become almost universal. Q1 and Q2 work as common-emitter stages, with the gain-control network R5, RV1, C3 connected between the emitters. As the resistance of this network is reduced, the differential gain increases but the common-mode gain remains low. The two signals at the collectors are then summed by the op-amp. This does not have to be a low-noise type because the gain in the transistors means it is they which determine the noise

www.raguvaran.puzl.com

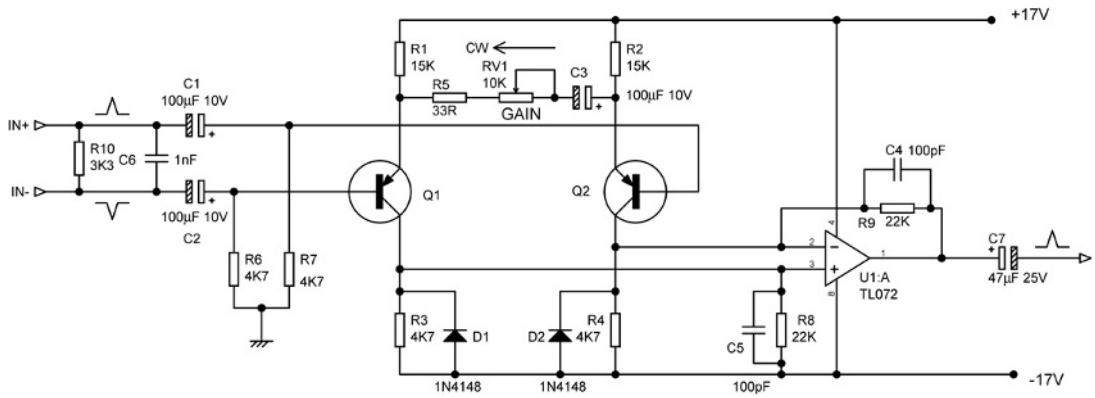


Figure 13.2: The simple hybrid microphone preamplifier

performance, and the TL072 was almost universally used in this position, being at the time much cheaper than the 5532. With appropriate choice of transistors (a type with low R_b) and collector current, preamplifiers of this type can give an equivalent input noise (EIN) of −128 dBu at maximum gain. The EIN rises as gain is reduced, because the resistance of the gain-control network and its Johnson noise is increased. The noise output, however, still falls as gain is reduced.

The gain law is very non-linear with R_g; a reverse-log D-law pot helps but there is still some cramping of the calibration at the high gain end, and preamplifier gain ranges of more than 50 dB are not really practicable with this configuration.

Today this approach is considered obsolete for anything except budget mixer purposes because of the mediocre distortion performance. You will note that the input transistor pair have no overall feedback loop closed around them, and their non-linearity creates significant distortion, especially at high gains.

The two reverse-biased diodes in the transistor collector circuits are to prevent the op-amp being damaged by having its input driven below the −17 V rail when phantom power is applied or removed. Note that the input coupling capacitors C1, C2 shown are not intended to cope with the phantom voltage if used. On arriving at the input XLR, the microphone signal first encounters the phantom feed resistors, DC-blocking capacitors (usually rated at 63 V), a switchable 20 dB attenuator, and possibly a phase-invert switch that swaps over the inputs, before it reaches the preamplifier. This is illustrated below in Figure 13.4.

The Balanced-Feedback Hybrid Microphone Preamplifier

The microphone preamp architecture described in the previous section has the merit of simplicity, but because there is no global feedback loop around the input transistors, its

www.raguvaran.puzl.com

distortion performance falls far short of the rest of a mixer, which typically consists of pure op-amp circuitry with very low distortion. It is obviously undesirable practically, aesthetically, and in every other way for the very first stage in the signal chain to irretrievably mess up the signal, but this is what happened for several years. In the mid-1980s I decided to do something about this. Many methods of applying feedback to the input stage were tried, but foundered on the fact that if feedback was applied to one of the transistors, the current variations in the other were excessive and created distortion. The use of two feedback paths in anti-phase, i.e. balanced feedback, was the solution to this problem, but this meant that one feedback path would have to be via an inverting amplifier with extra phase-shifts that would imperil HF stability.

The basic concept is shown in Figure 13.3. Direct negative feedback to Q2 goes through R12, while the inverted feedback passes through the unity-gain inverting stage U1:B and goes through R11 to Q1. The gain-control network R5, RV1, and C3 is connected between the two feedback points at the Q1, Q2 emitter, and sets the closed-loop gain by controlling the NFB factor. Distortion is pretty much eliminated, even at maximum gain.

The solution to the stability problem is to make sure that the direct HF feedback through C10 dominates that through C12, which has been through the inverter U1:A; this is aided by adding C4 to the inverter to control its HF response. This feedback is therefore not symmetrical at HF, but this has no effect on the functioning of the circuit with audio signals.

$$\text{Gain} = \frac{R_{NFB} + \left(\left(\frac{R_g}{2}R_e\right) / \left(\frac{R_g}{2} + R_e\right)\right)}{2\left(\frac{R_g}{2}R_e\right) / \left(\frac{R_g}{2} + R_e\right)} \qquad \text{(Equation 13.1)}$$

The closed-loop gain is given by Equation 13.1, which looks complicated, but becomes clearer when you appreciate that the terms in brackets simply represent the parallel combination of the emitter resistor R_e (R1, R2 in the figure) and half the resistance of the gain control network R_g. R_{NFB} is the value of R11, R12. The 2 on the bottom comes from the fact that we are only using one output of the amplifier – there is an inverted output from the inverter, which could be phase-summed to give twice the gain. We don't do that because it is actually desirable to have as low a minimum gain as possible. With the values shown the gain range is +22 to +71 dB, which in conjunction with a switchable 20 dB pad gives a useful total range of about 0–70 dB.

In order to get enough maximum gain, the feedback resistors R11, R12 have to be quite high at 82 kΩ. This means that the control of the DC level at the output is not very good. To solve this the DC servo integrator U2:A was added; R13 is connected to the preamplifier output and the integrator acts via the non-inverting input of the inverter op-amp to keep the final output at 0 V.

www.raguvaran.puzl.com

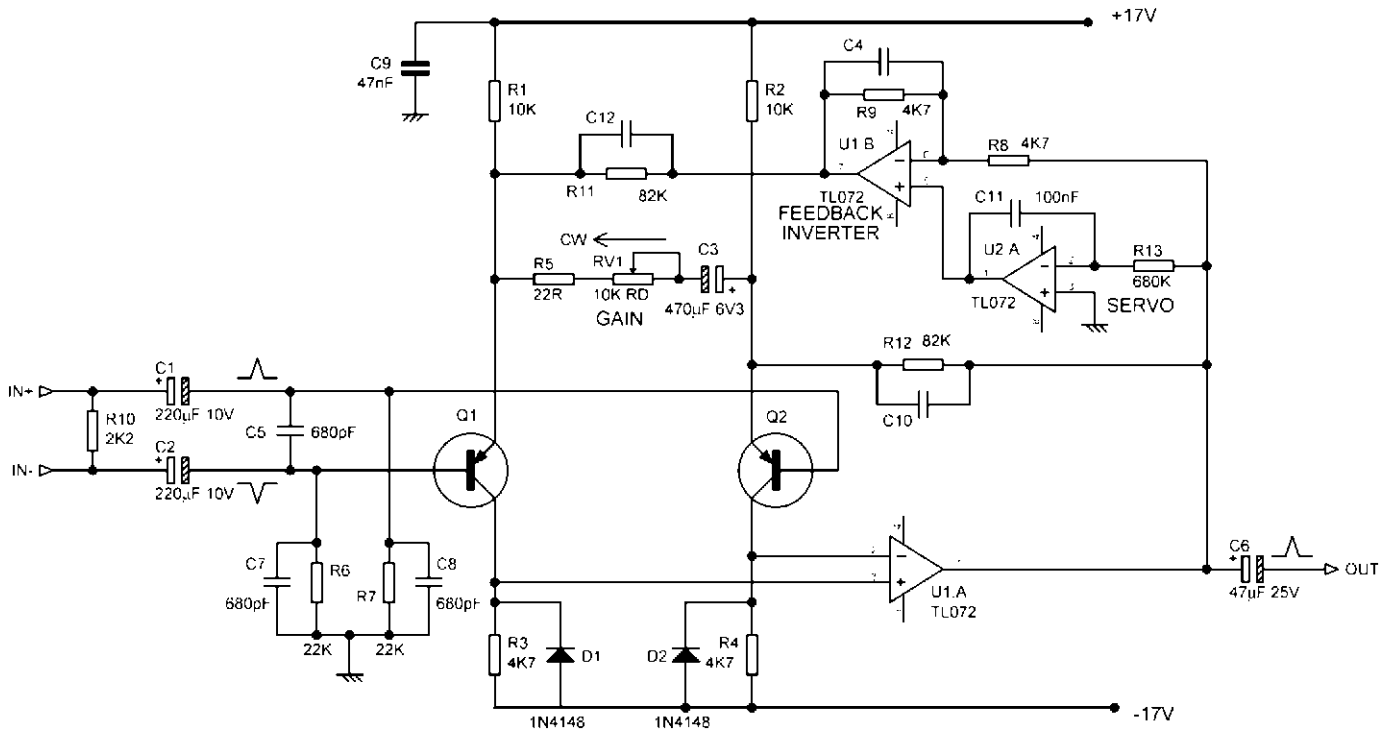


Figure 13.3: The balanced-feedback microphone amplifier, known to its friends as the BFMA

www.raguvaran.puzl.com

This technology was used on several consoles, such as the Soundcraft TS12, but it had a relatively short life as I came up with something better – the padless microphone preamplifier described later in this chapter.

Microphone and Line-Input Pads

Microphone pads, or attenuators, are used when the output is too high for the mixing console input to cope with; this typically happens when you put a microphone inside a kick-drum. Attenuators are also used when for reasons of economy it is desirable that the microphone input doubles up as a line input.

A typical arrangement is shown in Figure 13.4. The preamplifier has a typical gain range of +20 to +70 dB. There is an input XLR and phantom feed resistors R1, R2. C1 and C2 are DC-blocking capacitors to stop the phantom voltage from getting into the preamplifier; these should be as large as possible to preserve LF CMRR. Next comes a 20 dB balanced attenuator made up of R3–R6; note that the loading of the preamplifier input resistor R10 must be taken into account when designing the attenuator resistor values; one of the functions of this resistor is to prevent the preamplifier input from being open-circuited when the pad switch SW1 or the mic/line switch SW2 is between contacts. C3 and C4 are two further DC-blocking capacitors that prevent the input terminals of the preamplifier, which are not in general at 0 V, from causing clicks in the switching. C5, C6, and C7 increase EMC immunity and also keep the preamplifier from oscillating if the microphone input is left open-circuit at maximum gain. Such oscillation is not an indication that the preamplifier itself is unstable – it normally happens because the insert jacks, which carry the output signal from the preamplifier, are capacitatively crosstalking to the microphone input, forming a feedback loop. Ideally the microphone input should be open-circuit stable not only with the input gain at maximum, but also with full treble boost set up on the EQ section. This can be challenging to achieve, but it is possible. I have done it many times.

In line mode, the microphone gain is of course much too high, and the usual practice is to use a 30 dB attenuator on the line input, which allows a high input impedance to be set by R7, R9, while R8 provides a low source impedance to minimize preamplifier noise. This pad alters the gain range to −10 to +40 dB, which is actually too wide for a line input, and some consoles have another switch section in the mic/line switch, which reduces the gain range of the preamplifier so that the overall range is a more useful −10 to +20 dB. The larger and more expensive consoles usually have separate line-input stages, which avoid the compromises inherent in using the microphone input as a line input.

There is an important point to be made about the two attenuators. You will have noticed that the microphone attenuator uses four resistors and has its center connected to ground, whereas the line attenuator uses a more economical three resistors with no ground connection.

www.raguvaran.puzl.com

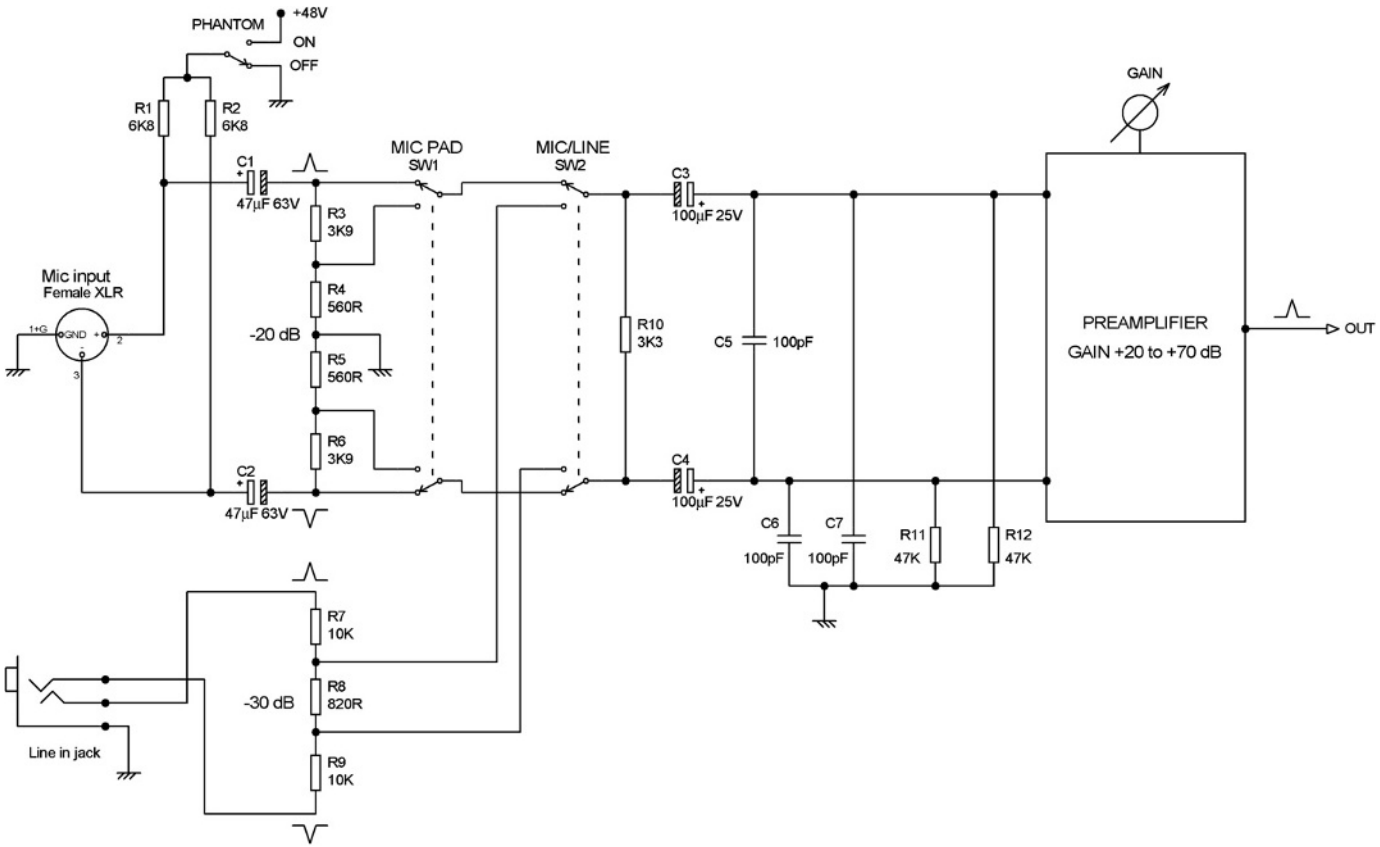


Figure 13.4: Mic and line attenuators at the input of a preamplifier

www.raguvaran.puzl.com

The disadvantage of the three-resistor version is that the wanted differential signal is attenuated, but the unwanted common-mode signal is not, and so the CMRR is much worsened. This does not happen with the four-resistor configuration because the ground connection means that both differential and common-mode signals are attenuated equally. There is no reason why the line attenuator here could not have been designed with four resistors – I just wanted to make the point.

The microphone amplifiers described have a high CMRR, and a problem with attenuators like this is that both types degrade the overall CMRR quite seriously because of their resistor tolerances, even if 1% components are used.

The Padless Microphone Preamplifier

The ideal microphone preamplifier would have a gain range of 70 dB or thereabouts on a single control, going down to unity gain without the inconvenience of a pad switch. It was mentioned in the previous section that resistive pads degrade the overall CMRR, and also the noise performance, as an inevitable consequence of following a 20 dB pad with an amplifier having 20 dB of gain. In addition, space on a channel front panel is always in short supply and losing a switch would be very welcome. I therefore invented the padless microphone preamplifier. Looking at the mixer market today (2009), the idea seems to have caught on.

The concept is based on the balanced-feedback mic amp described above, but now the total gain is spread over two stages to give a smooth 0–70 dB gain range with the rotation of a single knob.

The first stage shown in Figure 13.5 is based on the BFMA circuit in Figure 13.3, but with the feedback resistors reduced to 2k7 to reduce the gain range. The gain-control network R5, RV1, and C3 has also been halved in resistance to reduce Johnson noise, and the net result is a gain range of +1.5 to +49 dB; as before a reverse-log D-law pot is used. The lower feedback resistors mean that no servo is required to correct the DC conditions. Note that the greater amount of NFB means that under overload conditions it is possible for the common-mode range of the op-amp to be exceeded, leading to the well-known TL072 phase reversal and latch-up. This is prevented by R11, D3, and D4, which have no effect on linearity in normal operation.

The second stage of the padless mic amp is shown in Figure 13.6. This consists essentially of a variable-gain balanced input stage as described in Chapter 14, configured for a gain range of 0 to +20 dB. The gain pot is once again a reverse-log D-law pot and the combination of the gain laws of the two stages gives a very reasonable law over the almost 70 dB range, though there is still a little cramping at the high-gain end.

www.raguvaran.puzl.com

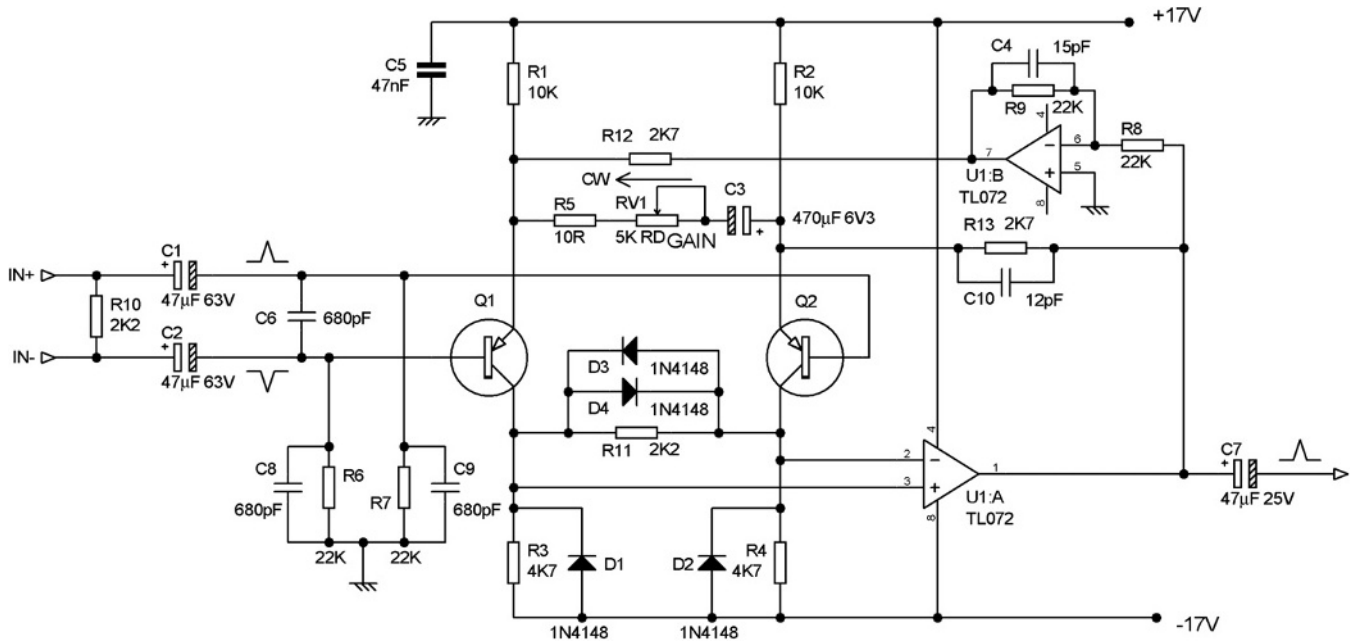


Figure 13.5: The padless balanced-feedback microphone preamp: mic input stage

www.raguvaran.puzl.com

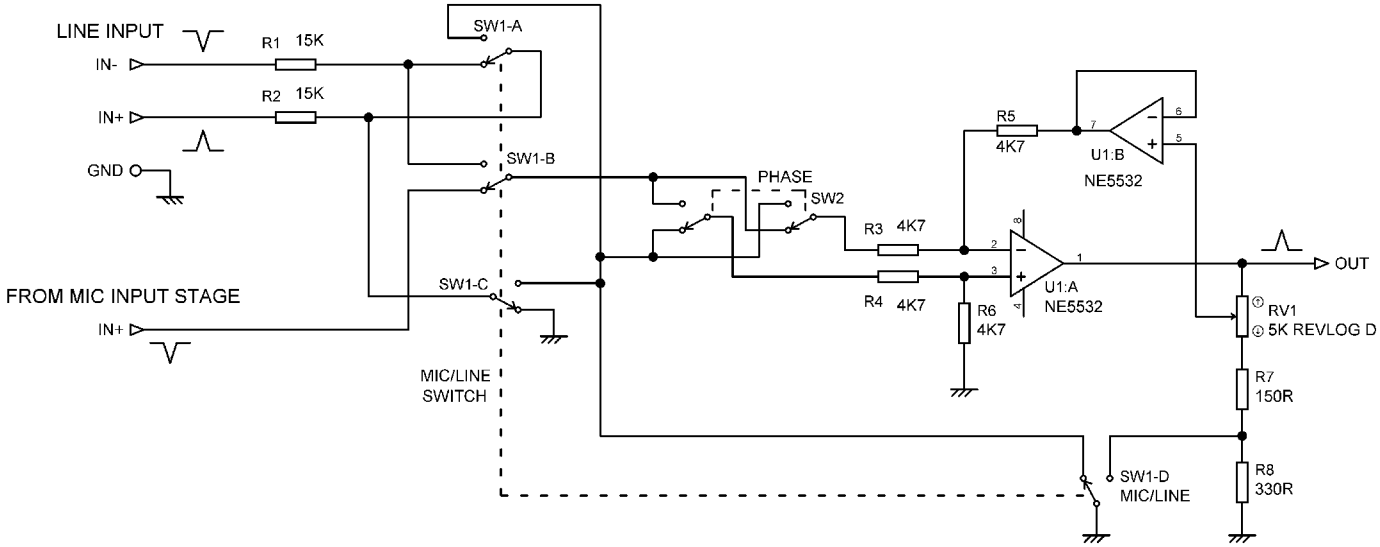


Figure 13.6: The padless balanced-feedback microphone preamp: mic/line switching and second stage

www.raguvaran.puzl.com

The second stage is also used as a line-input stage with a gain range of −10 to +20 dB. The mic/line switching used to do this may look rather complex but it does a bit more than simply change sources. In Figure 13.6 switch SW1 is shown in the 'mic' position, and the first stage reaches the inverting input of the second via SW1-B; the output of the mic amp is phase-inverted simply by swapping over its inputs. Line-input resistors R1, R2 give reduced gain for line input working, and in mic mode they are shorted together by SW1-A to prevent crosstalk from line to mic, which is an important issue when track normalling is incorporated (see Chapter 12). In several of my designs these resistors were placed on the input connector PCB rather than in the channel, to keep their hot ends away from other circuitry and further reduce line to mic crosstalk. SW1-C shorts the junction of R1, R2 to ground, further improving mic/line crosstalk if the line input is not balanced. SW1-D shorts the unused second-stage non-inverting input to ground.

In line mode R1, R2 are connected to the second stage via SW1-B and SW1-C. In mic mode, SW1-D shorts R8 to ground so that the gain range of the second stage is increased to the required 30 dB. SW2 is a phase-invert switch that simply swaps over the input connections to the second stage.

The padless mic amp gives both a good law gain control and lower noise at low gain settings. The noise performance versus gain for a typical example can be seen in Table 13.1; as described earlier, the EIN and noise figures worsen as the gain is turned down, due to the increased resistance of the gain control network. A noise figure of 30 dB may appear to be pretty dire, but the corresponding noise output is only −98 dBu, and this will soon be submerged in the noise from the following stages.

This effect can be reduced by reducing the impedance of the feedback and gain-control network, but this increases the power required to drive them, and because of the square root in the Johnson noise equation, a reduction by a factor of 10, which would need some serious

TABLE 13.1 Noise performance versus gain of padless mic amp

Gain (dB)	Noise out (dBu)	EIN (dBu)	Noise figure (dB)
1.4	−98.2	−99.6	30.0
6.6	−95.4	−102.0	27.6
15.4	−90.5	−105.9	23.7
31.7	−83.1	−114.8	14.8
44.1	−77.5	−121.6	8.0
51.8	−72.7	−124.5	5.1
60.6	−65.7	−126.3	3.3
69.3	−58.7	−128.0	1.6

www.raguvaran.puzl.com

electronics, would only give a 10 dB improvement, and that at the low gain end where it is least needed. Specialized outboard mic amps with low-resistance feedback networks driven by what are in effect small power amplifiers have been developed but do not seem to have caught on.

Another advantage of the padless approach is that one pair of DC-blocking capacitors suffices, rated at 63 V as in Figure 13.5, and this improves the CMRR at low frequencies. The padless microphone preamplifier concept was protected by patent number GB 2242089 in 1991, and was used extensively over many ranges of mixing console.

Capacitor Microphone Head Amplifiers

A capacitor capsule has an extremely high output impedance, equivalent to a very small capacitor of a few picofarads. It is in fact the highest impedance you are ever likely to encounter in the audio business, and certainly the highest I have ever had to deal with. Special circuit techniques are required to combine low noise and high impedance, working with a strictly limited amount of power. A while ago I designed the electronics for a new capacitor microphone by one of the well-known manufacturers, and the circuitry described here is a somewhat simplified version of that.

The first point is that the microphone capsule had an impedance of about 5 pF, so to get a −3 dB point of 10 Hz the total load impedance has to be no more than 3.2 GΩ (yes, that's 3200 MΩ). The capsule needs to be fed with a polarizing voltage through a resistor, and the head amplifier needs a biasing resistor. In this design both were 10 GΩ, which means that the input impedance of the amplifier itself had to be not less than 8.9 GΩ. Resistors with these astronomical values are exotic components that come in a glass encapsulation that must be manipulated with tweezers – one touch of a finger and the insulation properties of the glass are fatally compromised.

Figure 13.7 shows my capacitor mic head amp. R1 supplies the capsule polarizing voltage and R2 biases the first stage, a unity-gain JFET source-follower augmented by op-amp U1:A, which provides the gain for a high NFB factor to linearize Q1. The drain of Q1 is bootstrapped via C3 to prevent local feedback through the gate-drain capacitance of the JFET from reducing the input impedance. R3, R5, R6 set the DC conditions for Q1.

The second stage is a low-noise amplifier with gain of +4 dB, defined by the ratio of feedback resistor R17 to R15 and R16. Like the first stage, it is a hybrid design that combines the low noise of low-R_b transistor Q2 with the open-loop gain and load-driving capability of an op-amp.

The stage also acts as a unity-gain follower, making up a second-order Butterworth Sallen-and-Key high-pass filter when C6, C7 or C8, C9 are switched in; the resistive elements are R10 and R14.

www.raguvaran.puzl.com

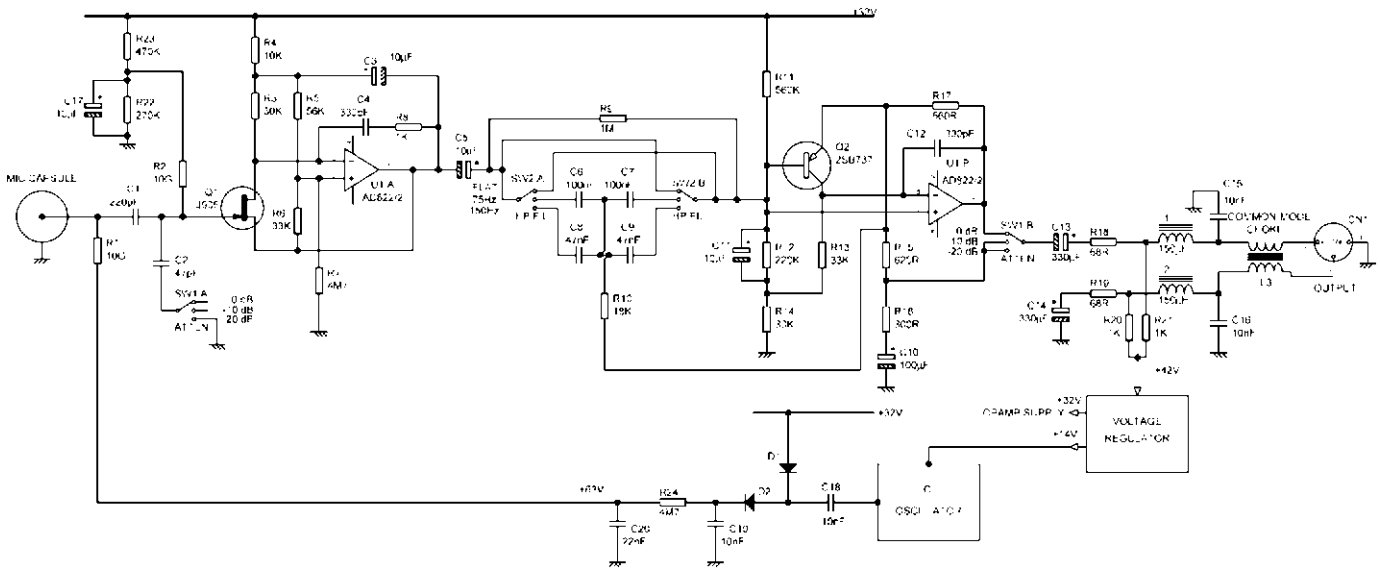


Figure 13.7: A typical head amp system for a capacitor microphone with phantom powering. All of this circuitry is fitted on to the body of the microphone

www.raguvaran.puzl.com

The system has two steps of attenuation: −10 and −20 dB. The first 10 dB step is obtained by using SW1-B to take the output from the junction of R15 and R16 instead of the normal stage output. The second 10 dB step results from switching C2 across the mic capsule, forming a capacitative attenuator that reduces the input to the first stage and prevents overload. The gains available are thus +4, −6, and −16 dB. The maximum sound pressure handling is +146 dB SPL, or +155 dB SPL with the −20 dB pad engaged.

The incoming phantom power is tapped off by R20, R21 and fed to a discrete BJT regulator that gives +32 V to power the op-amp, and +14 V to run a small LC oscillator that pumps the +32 V rail up to the +63 V required to polarize the capsule. Total current consumption is 2.2 mA.

The noise output is −120.7 dBu (A-weighted), which may appear high compared with the microphone amplifiers described above, but remember that a capacitor microphone puts out a high signal voltage so that the signal-to-noise ratio is actually very good. The mic capsule, being a pure reactance, generates no noise of its own, and its noise output comes only from the Brownian motion of the air against the capsule diaphragm and from the electronics. Noise measurements of this technology require special methods. The impedances are so high that meaningful results can only be obtained by putting the circuitry inside a completely closed metal screening enclosure.

www.raguvaran.puzl.com

This page intentionally left blank

www.raguvaran.puzl.com

CHAPTER 14

Line Inputs

External Signal Levels

There are several standards for line signal levels. The −10 dBv standard is used for a lot of semi-professional recording equipment as it gives more headroom with unbalanced connections – the professional levels of +4 and +6 dBu assume balanced outputs that inherently give twice the output level for the same supply rails as it is measured between two pins with signals of opposite phase on them. See Table 14.1

Signal levels in dBu are expressed with reference to 0 dBu = 775 mVrms; the origin of this odd value is that it gives a power of 1 mW in a purely historical 600 Ω load. The unit of dBm refers to the same level but takes the power rather than the voltage as the reference – a distinction of little interest nowadays. Signals in dBv (or dBV) are expressed with reference to 0 dB = 1.000 Vrms.

These standards are well established, but that does not mean all equipment follows them. To take a current example, the Yamaha P7000S power amplifier requires +8 dBu (1.95 Vrms) to give its full output of 750 W into 8 Ω.

Internal Signal Levels

In any audio system it is necessary to select a suitable nominal level for the signal passing through it. This level is always a compromise – the signal level should be high so it suffers minimal degradation by the addition of circuit noise as it passes through the system, but not so high that it is likely to suffer clipping before the gain control, or generate undue distortion

TABLE 14.1 Nominal signal levels

	Vrms	dBu	dBv
Semi-professional	0.316	−7.78	−10
Professional	1.228	+4.0	+1.78
German ARD	1.55	+6.0	+3.78

Small Signal Audio Design; ISBN: 9780240521770
Copyright © 2010 Elsevier Ltd; All rights of reproduction, in any form, reserved.

www.raguvaran.puzl.com

below the clipping level. (This last constraint is not normally a problem with modern circuitry, which gives very low distortion right up to the clipping point.)

It must always be considered that the gain control may be maladjusted by setting it too low and turning up the input level from the source equipment, making input clipping more likely. The internal levels chosen are usually in the range −6 to 0 dBu (388 mVrms to 775 Vrms) but in some specialized equipment such as broadcast mixing consoles, where levels are unpredictable and clipping distortion less acceptable than a bit more noise, the nominal internal level may be as low as −16 dBu (123 mVrms). If the internal level is in the normal −6 to 0 dBu range, and the maximum output of an op-amp is taken as 9 Vrms (+21.3 dBu), this gives 20–27 dB of headroom before clipping occurs.

If the incoming signal does have to be amplified, this should be done as early as possible in the signal path, to get the signal well above the noise floor as quickly as possible. If the gain is implemented in the first stage (i.e. the input amplifier, balanced or otherwise), the signal will be able to pass through later stages at a high level and so their noise contribution will be less significant. On the other hand, if the input stage is configured with a fixed gain, it will not be possible to turn it down to avoid clipping. Ideally the input stage should have variable gain. It is not straightforward to combine this feature with a balanced input, but several ways of doing it are shown later in this chapter.

Input Amplifier Functions

Firstly, RF filtering is applied at the very front end to prevent noise breakthrough and other EMC problems. It must be done before the incoming signal encounters any semiconductors where RF demodulation could occur, and can be regarded as a 'roofing filter'. At the same time, the bandwidth at the low end is given an early limit by the use of DC-blocking capacitors, and in some cases overvoltage spikes are clamped by diodes. The input amplifier should present a reasonably high impedance to the outside world, not less than 10 kΩ, and preferably more. It must have a suitable gain – possibly switched or variable – to scale the incoming signal to the nominal internal level. Balanced input amplifiers also accurately perform the subtraction process that converts differential signals to single-ended ones, so noise produced by ground loops and the like is rejected. It's quite a lot of work for one stage.

Unbalanced Inputs

The simplest unbalanced input feeds the incoming signal directly to the first stage of the audio chain. This is often impractical; for example, if the first stage was a Baxandall tone-control circuit then the boost and cut curves would be at the mercy of whatever source impedance was feeding the input. In addition, the input impedance would be low, and variable with frequency

www.raguvaran.puzl.com

and control settings. Some sort of buffer amplifier that can be fed from a significant impedance without ill effect is needed.

Figure 14.1 shows an unbalanced input amplifier, with the added components needed for interfacing to the real world. The op-amp U1:A acts as a unity-gain voltage-follower; it can be easily altered to give gain by adding two series feedback resistors. A 5532 bipolar type is used here for low noise; with the low source impedances that are likely to be encountered here, an FET-input op-amp would be 10 dB or more noisier. R1 and C1 are a first-order low-pass filter to remove incoming RF before it has a chance to reach the op-amp and demodulate into the audio band; once this has occurred any further attempts at RF filtering are, of course, pointless. R1 and C1 must be as close to the input socket as physically possible to prevent RF from being radiated inside the box before it is shunted to ground and so come before all other components in the signal path.

Selecting component values for input filters of this sort is always a compromise, because the output impedance of the source equipment is not known. If the source is an active preamplifier stage, then the output impedance will probably be around 50 Ω, but it could be as high as 500 Ω or more. If the source is an oxymoronic 'passive preamplifier' – i.e. just an input selector switch and a volume potentiometer – then the output impedance will be a good deal higher. (At least one passive preamplifier uses a transformer with switched taps for volume control – see Chapter 9.) If you really want to use a piece of equipment that embodies its internal contradictions in its very name, then a reasonable potentiometer value is 10 kΩ, and its maximum output impedance (when it is set for 6 dB of attenuation) will be 2.5 kΩ, which is very different from the 50 Ω we might expect from a good active preamplifier. This is in series with R1 and affects the turnover frequency of the RF filter. Effective RF filtering is very desirable, but it is also important to avoid a frequency response that sags significantly at 20 kHz. Valve equipment is also likely to have a high output impedance.

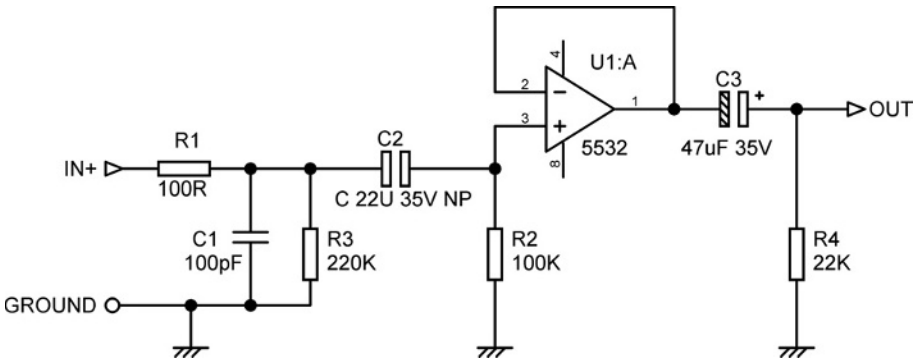


Figure 14.1: A typical unbalanced input amplifier with associated components

www.raguvaran.puzl.com

Taking 2.5 kΩ as a worst-case source impedance and adding R1, then 2.6 kΩ and 100 pF together give us −3 dB at 612 kHz; this gives a 20 kHz loss of only 0.005 dB, so possibly C1 could be usefully increased; for example, if we made it 220 pF then the 20 kHz loss is still only 0.022 dB, but the −3 dB point is 278 kHz, much improving the rejection of what used to be called the Medium Wave. If we stick with C1 at 100 pF and assume an active output with a 50 Ω impedance in the source equipment, then together with the 100 Ω resistance of R1 the total is 150 Ω, which in conjunction with 100 pF gives us −3 dB at 10.6 MHz. This is rather higher than desirable, but it is not easy to see what to do about it, and we must accept the compromise. If there was a consensus that the output impedance of a respectable piece of audio equipment should not exceed 100 Ω, then things would be much easier.

Our compromise seems reasonable, but can we rely on 2.5 kΩ as a worst-case source impedance? I did a quick survey of the potentiometer values that passive preamplifiers currently employ, and while it confirmed that 10 kΩ seems to be the most popular value, one model had a 20 kΩ potentiometer and another had a 100 kΩ pot. The latter would have a maximum output impedance of 25 kΩ, and would give very different results with a C1 value of 100 pF – the worst-case frequency response would now be −3 dB at 63.4 kHz and −0.41 dB at 20 kHz, which is not helpful if you are aiming for a ruler-flat response in the audio band.

To put this into perspective, filter capacitor C1 will almost certainly be smaller than the capacitance of the interconnecting cable. Audio interconnect capacitance is usually in the range 50–150 pF/meter, so with our assumed 2.5 kΩ source impedance and 150 pF/meter cable, and ignoring C1, you can only permit yourself a rather short run of 3.3 meters before you are −0.1 dB down at 20 kHz, while with a 25 kΩ source impedance you can hardly afford to have any cable at all; if you use low-capacitance 50 pF/meter cable you might just get away with a meter. This is just one of many reasons why 'passive preamplifiers' really are *not* a good idea.

Another important consideration is that the series resistance R1 must be kept as low as practicable to minimize Johnson noise, but lowering this resistance means increasing the value of shunt capacitor C1, and if it becomes too big then its impedance at high audio frequencies will become too low. Not only will there be too low a roll-off frequency if the source has a high output impedance, but there might be an increase in distortion at high audio frequencies because of excessive loading on the source output stage.

Replacing R1 with a small inductor to make an LC low-pass filter will give much better RF rejection at increased cost. This is justifiable in professional audio equipment, but it is much less common in hi-fi, one reason being that the unpredictable source impedance makes the filter design difficult, as we have just seen. In the professional world one *can* assume that the source impedance will be low. Adding more capacitors and inductors allows a three- or four-pole LC filter to be made. If you do use inductors then it is essential to check the frequency response to make sure it is what you expect and there is no peaking at the turnover frequency.

www.raguvaran.puzl.com

C2 is a DC-blocking capacitor to prevent voltages from ill-conceived source equipment getting into the circuitry. It is a non-polarized type as voltages from the outside world are of unpredictable polarity, and it is rated at not less than 35 V so that even if it gets connected to defective direct-coupled equipment with an op-amp output jammed hard against one of the supply rails, no harm will result. R3 is a DC drain resistor that prevents the charge put on C2 by the aforesaid external equipment from remaining there for a long time and causing a thud when connections are replugged; as with all input drain resistors, its value is a compromise between discharging the capacitor reasonably quickly and keeping the input impedance acceptably high. The input impedance here is R3 in parallel with R2, i.e. 220 kΩ in parallel with 100 kΩ, giving 68 kΩ. This is a good high value and should work well with just about any source equipment you can find, including valve technology.

R2 provides the biasing for the op-amp input; it must be a high value to keep the input impedance up, but bipolar input op-amps draw significant input bias current. The Fairchild 5532 data sheet quotes 200 nA typical and 800 nA maximum, and these currents would give a voltage drop across R2 of 20 and 80 mV respectively. This offset voltage will be reproduced at the output of the op-amp, with the input offset voltage added on; this is only 4 mV maximum and so will not affect the final voltage much, whatever its polarity. The 5532 has NPN input transistors, and the bias current flows into the input pins, so the voltage at pin 3 and hence the output will be negative with respect to ground by anything up to 84 mV.

Such offset voltages are not so great that the output voltage swing of the op-amp is significantly affected, but they are enough to generate unpleasant clicks and pops if the input stage is followed by any sort of switching, and enough to make potentiometers crackly. Output DC blocking is therefore required in the shape of C3, while R4 is another DC drain resistor to keep the output at zero volts. It can be made rather lower in value than the input drain resistor R3 as the only requirement is that it should not significantly load the op-amp output. FET-input op-amps have much lower input bias currents, so that the offsets they generate as they flow through biasing resistors are usually negligible, but they still have input offsets of a few millivolts, so DC blocking will still be needed if switches downstream are to work silently.

This input stage, with its input terminated by 50 Ω to ground, has a noise output of only −119.0 dBu over the usual 22 Hz – 22 kHz bandwidth. This is very quiet indeed, and is a reflection of the fact that R1, the only resistor in the signal path, has the low value of 100 Ω and so generates a very small amount of Johnson noise, only −132.6 dBu. This is swamped by the voltage noise of the op-amp, which is basically all we see; its current noise has negligible effect because of the low circuit impedances.

An unbalanced input stage with a 470 kΩ input impedance and gain variable over a 6 dB range for stereo balance control is described in Chapter 11.

www.raguvaran.puzl.com

Balanced Interconnections

Balanced inputs are used to prevent noise and crosstalk from affecting the input signal, especially in applications where long interconnections are used. They are standard on professional audio equipment, and are slowly but steadily becoming more common in the world of hi-fi. Their importance is that they can render ground loops and other connection imperfections harmless. Since there is no point in making a wonderful piece of equipment and then feeding it with an impaired signal, making an effective balanced input is of the first importance.

The basic principle of balanced interconnection is to get the signal you want by subtraction, using a three-wire connection. In some cases a balanced input is driven by a balanced output, with two anti-phase output signals – one signal wire (the hot or in-phase) sensing the in-phase output of the sending unit, while the other senses the anti-phase output.

In other cases, when a balanced input is driven by an unbalanced output, as shown in Figure 14.2, one signal wire (the hot or in-phase) senses the single output of the sending unit, while the other (the cold or phase-inverted) senses the unit's output-socket ground, and once again the difference between them gives the wanted signal. In either of these two cases, any noise voltages that appear identically on both lines (i.e. common-mode signals) are in theory completely canceled by the subtraction. In real life the subtraction falls short of perfection, as the gains via the hot and cold inputs will not be precisely the same, and the degree of discrimination actually achieved is called the common-mode rejection ratio (CMRR), of which more later.

It is deeply tedious to keep referring to non-inverting and inverting inputs, and so these are usually abbreviated to 'hot' and 'cold' respectively. This does *not* necessarily mean that the hot terminal carries more signal voltage than the cold one. For a true balanced connection, the voltages will be equal. The 'hot' and 'cold' terminals are also often referred to as In+ and In−, and this latter convention has been followed in the diagrams here.

The subject of balanced interconnections is a large one, and a big book could be written on this topic alone; one of the classic papers on the subject is by Muncy [1]. To make a start, let us look at the pros and cons of balanced connections.

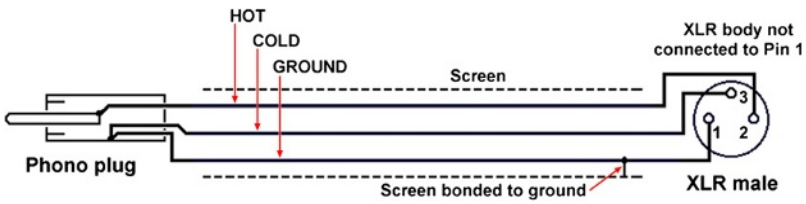


Figure 14.2: Unbalanced output to balanced input interconnection

www.raguvaran.puzl.com

The Advantages of Balanced Interconnections

- Balanced interconnections discriminate against noise and crosstalk, whether they result from ground currents, or electrostatic or magnetic coupling to signal conductors.
- Balanced connections make ground loops much less intrusive, and usually inaudible, so people are less tempted to start 'lifting the ground' to break the loop, with possibly fatal consequences. This tactic is only acceptable if the equipment has a dedicated ground-lift switch that leaves the external metalwork firmly connected to mains safety earth. In the absence of this switch the foolhardy and optimistic will break the mains earth (not quite so easy now that moulded mains plugs are standard) and this practice is of course highly dangerous, as a short-circuit from mains to the equipment chassis will result in live metalwork but dead people.
- A balanced interconnection incorporating a true balanced output gives 6 dB more signal level on the line, which should give 6 dB more dynamic range. However, this is true only with respect to *external* noise – as is described later in this chapter, a standard balanced input using 10 kΩ resistors is about 14 dB noisier than the unbalanced input shown in Figure 14.1.
- Balanced connections are usually made with XLR connectors. These are a professional three-pin format, and are far superior to the phono (RCA) type normally used for unbalanced connections. More on this below.

The Disadvantages of Balanced Interconnections

- Balanced inputs are inherently noisier than unbalanced inputs by a large margin, in terms of the noise generated by the input circuitry itself rather than external noise. This may appear paradoxical but it is all too true, and the reasons will be fully explained in this chapter.
- More hardware means more cost. Small-signal electronics is relatively cheap; unless you are using a sophisticated low-noise input stage, of which more later, most of the extra cost is likely to be in the balanced input connectors.
- Balanced connections do not of themselves provide any greater RF immunity than an unbalanced input. For this to happen both legs of the balanced input would have to demodulate the RF in equal measure for common-mode cancellation to occur. This is highly unlikely, and the chances of it happening over a wide frequency range are zero. It remains vital to provide the usual passive RF filtering in front of any electronics to avoid EMC troubles.
- There is the possibility of introducing a phase error. It is all too easy to create an unwanted phase inversion by confusing hot and cold when wiring up a connector, and

www.raguvaran.puzl.com

this can go undiscovered for some time. The same mistake on an unbalanced system interrupts the audio completely and leaves no room for doubt.

Balanced Cables and Interference

In a balanced interconnection two wires carry the signal, and the third connection is the ground wire, which has two functions. Firstly, it joins the grounds of the interconnected equipment together. This is not always desirable, and if galvanic isolation is required a transformer balancing system will be necessary because the large common-mode voltages are likely to exceed the range of an electronic balanced input. A good transformer will also have a very high CMRR, which will be needed to get a clean signal in the face of large CM voltages.

Secondly, the presence of the ground allows electrostatic screening of the two signal wires, preventing both the emission and pickup of unwanted signals. This can mean:

1. a lapped screen, with wires laid parallel to the central signal conductor. The screening coverage is not total, and can be badly degraded as the screen tends to open up on the outside of cable bends. Not recommended unless cost is the dominating factor;
2. a braided screen around the central signal wires. This is much more expensive, as it is harder to make, but opens up less on bending than a lap screen. Even so, screening is not 100%. It has to be said that it is a pain to terminate in the usual audio connectors. Not recommended;
3. an overlapping foil screen, with the ground wire (called the drain wire in this context for some reason) running down the inside of the foil and in electrical contact with it. This is usually the most effective as the foil is a solid sheet and cannot open up on the outside of bends. It should give perfect electrostatic screening, and it is much easier to work with than either lap screen or braided cable. However, the higher resistance of aluminum foil compared with copper braid means that RF immunity may not be so good.

There are three main ways in which an interconnection is susceptible to hum and noise.

1 Electrostatic Coupling

An interfering signal at significant voltage couples directly to the inner signal line, through stray capacitance. The stray capacitance between imperfectly screened conductors will be a fraction of a pF in most circumstances, as electrostatic coupling falls off with the square of distance. This form of coupling can be serious in studio installations with unrelated signals running down the same ducting.

The three main lines of defence against electrostatic coupling are effective screening, low impedance drive, and a good CMRR maintained up to the top of the audio spectrum. As regards screening, an overlapped foil screen provides complete protection.

www.raguvaran.puzl.com

Driving the line from a low impedance, of the order of 100 Ω or less, is also helpful because the interfering signal, having passed through a very small stray capacitance, is a very small current and cannot develop much voltage across such a low impedance. This is convenient because there are other reasons for using a low output impedance, such as optimizing the interconnection CMRR, minimizing HF losses due to cable capacitance, and driving multiple inputs without introducing gain errors. For the best immunity to crosstalk the output impedance must remain low up to as high a frequency as possible. This is definitely an issue as op-amps invariably have a feedback factor that begins to fall from a low, and quite possibly sub-audio, frequency and this makes the output impedance rise with frequency as the negative feedback factor falls, as if an inductor were in series. Some line outputs have physical series inductors to improve stability or EMC immunity, and these should not be so large that they significantly increase the output impedance at 20 kHz. From the point of view of electrostatic screening alone, the screen does not need to be grounded at both ends, or form part of a circuit [2]. It must of course be grounded at some point.

If the screening is imperfect, and the line impedance non-zero, some of the interfering signal will get into the hot and cold conductors, and now the CMRR must be relied upon to make the immunity acceptable. If it is possible, rearranging the cable-run away from the source of interference and getting some properly screened cable is more practical and more cost-effective than relying on very good common-mode rejection.

Stereo hi-fi balanced interconnections almost invariably use XLR connectors. Since the usual three-pin XLR can only handle one balanced channel, two separate cables are almost invariably used and interchannel capacitative crosstalk is not an issue. Professional systems, on the other hand, use multi-way connectors that do not have screening between the pins and there is an opportunity for capacitative crosstalk here, but the use of low source impedances should reduce it to below the noise floor.

2 Magnetic Coupling

If a cable runs through an AC magnetic field, an EMF is induced both in signal conductors and in the screen, and according to some writers the screen current must be allowed to flow freely or its magnetic field will not cancel out the field acting on the signal conductors, and therefore the screen should be grounded at both ends, to form a circuit [3]. In practice the magnetic field cancellation will be very imperfect and reliance is better placed on the CMRR of the balanced system to cancel out the hopefully equal voltages induced in the two signal wires. The need to ground both ends to possibly optimize the magnetic rejection is not usually a restriction, as it is rare that galvanic isolation is required between two pieces of audio equipment.

The equality of the induced voltages can be maximized by minimizing the loop area between the hot and cold signal wires, for example by twisting them tightly together in manufacture. In practice most audio foil-screen cables have parallel rather than twisted signal conductors,

www.raguvaran.puzl.com

but this seems adequate almost all of the time. Magnetic coupling falls off with the square of distance, so rearranging the cable-run away from the source of magnetic field is usually all that is required. It is unusual for it to present serious difficulties in a hi-fi application.

3 Ground Voltages

These are the result of current flowing through the ground connection, and is often called 'common-impedance coupling' in the literature [1]. This is the root of most ground-loop problems. The existence of a loop in itself does no harm, but it is invariably immersed in a 50 Hz magnetic field that induces mains-frequency currents plus harmonics into it. This current produces a voltage drop down non-negligible ground-wire resistances, and this effectively appears as a voltage source in each of the two signal lines. Since the CMRR is finite a proportion of this voltage will appear to be a differential signal, and will be reproduced as such.

Balanced Connectors

Balanced connections are most commonly made with XLR connectors, though it can be done with stereo (tip-ring-sleeve) jack plugs. XLRs are a professional three-pin format, and are a much better connector in every way than the usual phono (RCA) connectors used for unbalanced interconnections. Phono connectors have the great disadvantage that if you are connecting them with the system active (inadvisable, but then people are always doing inadvisable things) the signal contacts meet before the grounds and thunderous noises result. The XLR standard has pin 2 as hot, pin 3 as cold, and pin 1 as ground.

Stereo jack plugs are often used for line-level signals in a recording environment, and are frequently found on the rear of professional power amplifiers as an alternative to an adjacent XLR connector; both full-size and 3.5 mm sizes are used. Balanced jacks are wired with the tip as hot, the ring as cold, and the sleeve as ground. Sound reinforcement systems often use large multi-way connectors that carry dozens of three-wire balanced connections.

Balanced Signal Levels

Many pieces of equipment, including preamplifiers and power amplifiers designed to work together, have both unbalanced and balanced inputs and outputs. The general consensus in the hi-fi world is that if the unbalanced output is say 1 Vrms, then the balanced output will be created by feeding the in-phase output to the hot output pin, and also to a unity-gain inverting stage, which drives the cold output pin with 1 Vrms phase-inverted. The total balanced output voltage between hot and cold pins is therefore 2 Vrms, and so the balanced input must have a gain of 0.5 or −6 dB relative to the unbalanced input to maintain consistent internal signal levels.

www.raguvaran.puzl.com

Electronic vs. Transformer Balanced Inputs

Balanced interconnections can be made using either transformer or electronic balancing. Electronic balancing has many advantages, such as low cost, low size and weight, superior frequency and transient response, and no low-frequency linearity problems. It may still be regarded as a second-best solution in some quarters, but the performance is more than adequate for most professional applications. Transformer balancing does have some advantages of its own, particularly for work in very hostile RF/EMC environments, but serious drawbacks. The advantages are that transformers are electrically bulletproof, retain their high CMRR performance forever, and consume no power even at high signal levels. They are essential if galvanic isolation between ground is required. Unfortunately transformers generate LF distortion, particularly if they have been made with minimal core sizes to save weight and cost. They are liable to have HF response problems due to leakage reactance and distributed capacitance, and compensating for this requires a carefully designed Zobel network across the secondary. Inevitably they are heavy and expensive. Transformer balancing is therefore relatively rare, even in professional audio applications, and the greater part of this chapter deals with electronically balanced inputs.

Common-Mode Rejection

Figure 14.3 shows a balanced interconnection reduced to its bare essentials: hot and cold line outputs with source resistances R_{out+}, R_{out-} and a standard differential amplifier at the input end. The output resistances are assumed to be exactly equal, and the balanced input in the receiving equipment has two exactly equal input resistances to ground R1, R2. The ideal balanced input amplifier senses the voltage difference between the points marked In+ (hot) and In− (cold), and ignores any common-mode voltages that are present on both. The amount by which it discriminates is called the common-mode rejection ratio or CMRR, and is usually measured in dB. Suppose a differential voltage input between In+ and In− gives an output

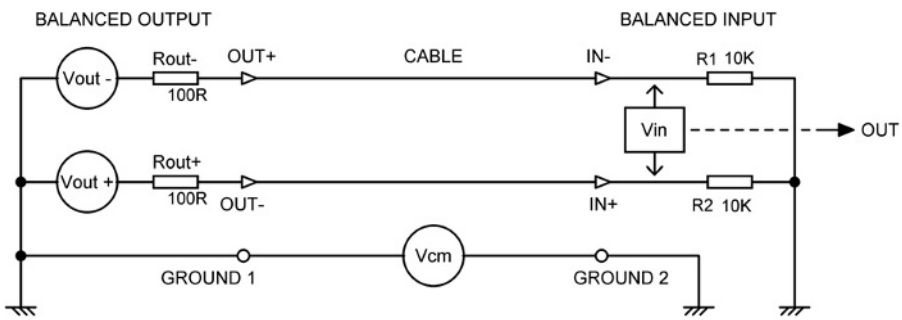


Figure 14.3: A theoretical balanced interconnection showing how the output and input impedances influence CMRR

www.raguvaran.puzl.com

voltage of 0 dB; now reconnect the input so that In+ and In− are joined together and the same voltage is applied between them and ground. Ideally the result would be zero output, but in this imperfect world it won't be, and the output could be anywhere between −20 dB (for a bad balanced interconnection, which probably has something wrong with it) and −140 dB (for an extremely good one). The CMRR when plotted may have a flat section at low frequencies, but it very commonly degrades at high audio frequencies and may also deteriorate at very low frequencies. More on that later.

In one respect balanced audio connections have it easy. The common-mode signal is normally well below the level of the unwanted signal, and so the common-mode range of the input is not an issue. In other areas of technology, such as electrocardiogram amplifiers, the common-mode signal may be many times greater than the wanted signal.

The simplified conceptual circuit of Figure 14.3, under SPICE simulation, demonstrates the need to get the resistor values right for a good CMRR, before you even begin to consider the rest of the circuitry. The differential voltage sources V_{out+}, V_{out-} representing the actual balanced output are set to zero, and V_{cm}, which represents the common-mode voltage drop down the cable ground, is set to 1 V to give a convenient result in dBv. The output resulting from the presence of this voltage source is measured by a mathematical subtraction of the voltages at In+ and In− so there is no actual input amplifier to confuse the results with its non-ideal performance.

Let us begin with R_{out+} and R_{out-} set to 100 Ω and R1 and R2 set to 10 kΩ. These are typical real-life values as well as being nice round figures. When all four resistances are exactly at their nominal value, the CMRR is in theory infinite, which my SPICE simulator rather curiously reports as exactly −400 dB. If one of the output resistors or one of the input resistors is then altered in value by 1%, then the CMRR drops like a stone to −80 dB. If the deviation from equality is 10%, things are predictably worse and the CMRR degrades to −60 dB, as shown in Table 14.2. That would be quite a good figure in reality, but since we have not yet even thought about op-amp imperfections or other circuit imbalances, and have only altered one resistance out of the four that will in real circuitry all have their own tolerances, it underlines the need to get things right at the most basic theoretical level before we dig deeper into the circuitry. The CMRR is naturally flat with frequency because our simple model has no frequency-dependent components.

The essence of the problem is that we have two resistive dividers, and to get an infinite CMRR they must have exactly the same attenuation. If we increase the ratio between the output and input resistors, by reducing the former or increasing the latter, the attenuation factor becomes closer to unity, so variations in either resistor value have less effect on it. If we increase the input impedance to 100 kΩ, which is quite practical in real life (we will put aside the noise implications of this for the moment), things are 10 times better, as the R_{in}/R_{out} ratio has improved from 100 to 1000 times. We now get a CMRR of −100 dB with a 1% resistance deviation, and −80 dB with a 10% deviation. An even higher input impedance of 1 MΩ, which

www.raguvaran.puzl.com

TABLE 14.2 How resistor tolerances affect the theoretical CMRR of the theoretical circuit in Figure 14.3

R_{out+}	R_{out-}	R_{out} deviation	R1	R2	R1, R2 deviation	R_{in}/R_{out} ratio	CMRR (dB)
100	100	0	10k	10k	0	100	Infinity
100	101	1%	10k	10k	0	100	−80.2
100	110	10%	10k	10k	0	100	−60.2
100	100	0	10k	10.1k	1%	100	−80.3
100	100	0	10k	11k	10%	100	−61.0
100	100	0	100k	101k	1%	1000	−100.1
100	100	0	100k	110k	10%	1000	−80.8
100	100	0	1M	1.01M	1%	10,000	−120.1
100	100	0	1M	1.1M	10%	10,000	−100.8
68	68	0	20k	20.2k	1%	294	−89.5
68	68	0	20k	22k	10%	294	−70.3

is perhaps a bit less practical, raises R_{in}/R_{out} to 10,000 and gives −120 dB for a 1% resistance deviation and −100 dB for a 10% deviation.

We can attack the other aspect of the attenuation problem by reducing the output impedances to 10 Ω, ignoring for the moment the need to secure against HF instability caused by cable capacitance, and also return the input impedance resistors to 100 kΩ. R_{in}/R_{out} is 10,000 once more, and, as you might suspect, the CMRR is once more −120 dB for a 1% deviation and −100 dB for a 10% deviation. Ways to make stable output stages with very low output impedances are described in Chapter 15; a fraction of an ohm at 1 kHz is quite easy to achieve.

In conventional circuits, the combination of 68 Ω output resistors and a 20 kΩ input impedance is often encountered, 68 Ω being about as low as you want to go if HF instability is to be absolutely guarded against with the long lines used in professional audio. The 20 kΩ common-mode input impedance is what you get if you make a basic balanced input amplifier with four 10 kΩ resistors. I strongly suspect that this value is so popular because it looks as if it gives standard 10 kΩ input impedances; in fact it does nothing of the sort, and the common-mode input impedance, which is what matters here, is 20 kΩ on each leg – more on that later. It turns out that 68 Ω output resistors and a 20 kΩ input impedance give a theoretical CMRR of −89.5 dB for a 1% deviation of one resistor, which is quite encouraging. These results are summarized in Table 14.2.

The conclusion is simple: we need the lowest possible output impedances and the highest possible input impedances to get the maximum common-mode rejection. This is highly

www.raguvaran.puzl.com

convenient because low output impedances are already needed to drive multiple amplifier inputs and cable capacitance, and high input impedances are needed to minimize loading and maximize the number of amplifiers that can be driven.

The Basic Electronic Balanced Input

Figure 14.4 shows the basic balanced input amplifier using a single op-amp. To achieve balance R1 must be equal to R3 and R2 equal to R4. The amplifier in Figure 14.5 has a gain of R2/R1 (= R4/R3). The standard one-op-amp balanced input or differential amplifier is a very familiar circuit block, but its operation often appears somewhat mysterious. Its input impedances are *not* equal when it is driven from a balanced output; this has often been commented on [4], and some confusion has resulted.

The source of the confusion is that a simple differential amplifier has interaction between the two inputs, so that the input impedance seen on the cold input depends on the signal applied to the hot input. Input impedance is measured by applying a signal and seeing how much current flows into the input, so it follows that the apparent input impedance on each leg varies

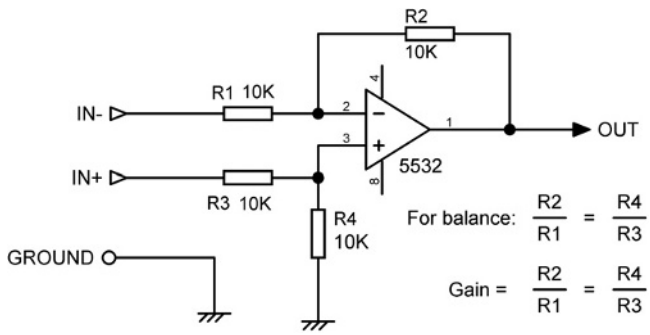


Figure 14.4: The basic balanced input amplifier

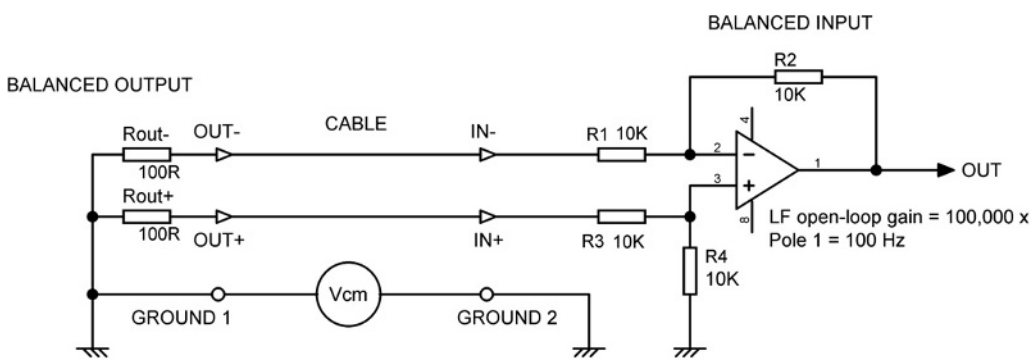


Figure 14.5: A simple balanced interconnection for SPICE simulation to show the effect that op-amp properties have on the CMRR

www.raguvaran.puzl.com

according to how the cold input is driven. If the amplifier is made with four 10 kΩ resistors, then the input impedances on hot and cold are as shown in Table 14.3.

Some of these impedances are not exactly what you might expect, and require some explanation.

Case 1. The balanced input is being used as an unbalanced input by grounding the cold input and driving the hot input only. The input impedance is therefore simply R3 + R4. Resistors R3 and R4 reduce the signal by a factor of a half, but this loss is undone as R1 and R2 set the amplifier gain to 2 times, and the overall gain is unity. If the cold input is not grounded then the gain is 0.5 times. The attenuate-then-amplify architecture, plus the Johnson noise from the resistors, makes this configuration much noisier than the dedicated unbalanced input of Figure 14.1, which has only a single 100 Ω resistor in the signal path.

Case 2. The balanced input is again being used as an unbalanced input, but this time by grounding the hot input, and driving the cold input only. This gives a phase inversion and it is unlikely you would want to do it except as an emergency measure to correct a phase error somewhere else. The important point here is that the input impedance is now only 10 kΩ, the value of R1, because shunt negative feedback through R2 creates a virtual earth at pin 2 of the op-amp. Clearly this simple circuit is not as symmetrical as it looks. The gain is unity, whether or not the hot input is grounded; grounding it is desirable because it not only prevents interference being picked up on the hot input pin, but also puts R3 and R4 in parallel, reducing the resistance from op-amp pin 3 to ground and so reducing Johnson noise.

Case 3. This is the standard balanced interconnection. The input is driven from a balanced output with the same signal levels on hot and cold, as if from a transformer with its center-tap grounded, or an electronically balanced output using a simple inverter to drive the cold pin. The input impedance on the hot input is what you would expect; R3 + R4 add up to 20 kΩ. However, on the cold input there is a much lower input impedance of 6.66 kΩ. This at first sounds impossible as the first thing the signal encounters is a 10 kΩ series resistor, but the crucial point is that the hot input is being driven simultaneously with a signal of the opposite phase, so the inverting op-amp input is moving in the opposite direction to the cold input due to negative feedback, and what you might call anti-bootstrapping reduces the effective value of the 10 kΩ resistor to 6.66 kΩ. These are the differential input impedances we are examining,

TABLE 14.3 The input impedances for different input drive conditions

Case	Pins driven	Hot input res. (Ω)	Cold input res. (Ω)
1	Hot only	20k	Grounded
2	Cold only	Grounded	10k
3	Both (balanced)	20k	6.66k
4	Both common-mode	20k	20k
5	Both floating	10k	10k

www.raguvaran.puzl.com

the impedances seen by the balanced output driving them. Common-mode signals see a common-mode impedance of 20 kΩ, as in Case 4 below.

You will sometimes see the statement that these unequal differential input impedances 'unbalance the line'. From the point of view of CMRR, this is not the case, as it is the CM input impedance that counts. The line is, however, unbalanced in the sense that the cold input draws three times the current from the output that the hot one does. This current imbalance might conceivably lead to inductive crosstalk in some multi-way cable situations, but I have never encountered it. The differential input impedances can be made equal by increasing the R1 and R2 resistor values by a factor of 3, but this degrades the noise performance markedly and makes the common-mode impedances to ground unequal, which is a much worse situation as it compromises the rejection of ground voltages, and these are almost always the main problem in real life.

Case 4. Here both inputs are driven by the same signal, representing the existence of a common-mode voltage. Now both inputs shown an impedance of 20 kΩ. It is the symmetry of the common-mode input impedances that determines how effectively the balanced input rejects the common-mode signal. This configuration is of course only used for CMRR testing.

Case 5. Now the input is driven as from a floating transformer with the center-tap (if any) unconnected, and the impedances can be regarded as equal; they must be, because with a floating winding the same current must flow into each input. However, in this connection the line voltages are *not* equal and opposite: with a true floating transformer winding the hot input has all the signal voltage on it while the cold has none at all, due to the negative feedback action of the balanced input amplifier. This seemed very strange when it emerged in SPICE simulation, but a sanity check with real components proves it true. The line has been completely unbalanced as regards crosstalk to other lines, although its own common-mode rejection remains good.

Even if absolutely accurate resistors are assumed, the CMRR of the stage in Figure 14.4 is not infinite; with a TL072 it is about −90 dB, degrading from 100 Hz upwards, due to the limited open-loop gain of the op-amp. We will now examine this effect.

Common-Mode Rejection: The Basic Balanced Input and Op-Amp Effects

In the earlier section on CMRR we saw that in a theoretical balanced line, choosing low output impedances and high input impedances would give very good CM rejection even if the resistors were not perfectly matched. Things are a bit more complex (i.e. worse) if we replace the mathematical subtraction with a real op-amp. We quickly find that even if perfectly matched resistors everywhere are assumed, the CMRR of the stage is not infinite, because the two op-amp inputs are not at exactly the same voltage. The negative feedback error voltage between the inputs depends on the open-loop gain of the op-amp, and that is neither infinite nor

www.raguvaran.puzl.com

flat with frequency into the far ultraviolet. Far from it. There is also the fact that op-amps themselves have a CMRR; it is high, but once more it is not infinite.

As usual, SPICE simulation is instructive, and Figure 14.5 shows a simple balanced interconnection, with the balanced output represented simply by two 100 Ω output resistances connected to the source equipment ground, here called Ground 1, and the usual differential op-amp configuration at the input end, where we have Ground 2.

A common-mode voltage V_{cm} is now injected between Ground 1 and Ground 2, and the signal between the op-amp output and Ground 2 measured. The balanced input amplifier has all four of its resistances set to precisely 10 kΩ, and the op-amp is represented by a very simple model that has only two parameters: a low-frequency open-loop gain, and a single pole frequency that says where that gain begins to roll-off at 6 dB/octave. The op-amp input impedances and the op-amp's own CMRR are assumed infinite, as in the world of simulation they so easily can be. Its output impedance is set at zero.

For the first experiments, even the pole frequency is made infinite, so now the only contact with harsh reality is that the op-amp open-loop gain is finite. That is, however, enough to give distinctly non-ideal CMRR figures, as Table 14.4 shows.

With a low-frequency open-loop gain of 100,000, which happens to be the typical figure for a 5532 op-amp, even perfect components everywhere will never yield a better CMRR than −94 dB. The CMRR is shown as a raw ratio in the third column so you can see that the CMRR is inversely proportional to the gain, and so we want as much gain as possible.

Op-Amp Frequency Response Effects

To examine these we will set the low-frequency gain to 100,000, which gives a CMRR 'floor' of −94 dB, and then introduce the pole frequency that determines where it rolls off. The CMRR now worsens at 6 dB/octave, starting at a frequency set by the interaction of the low-frequency gain and the pole frequency. The results are summarized in Table 14.5, which shows that, as you might expect, the lower the open-loop bandwidth of the op-amp, the lower the frequency at which the CMRR begins to fall off. Figure 14.6 shows the situation diagrammatically.

TABLE 14.4 The effect of finite op-amp gain on CMRR for the circuit of Figure 14.5

Open-loop gain	CMRR (dB)	CMRR ratio
10,000	−74.0	19.9×10^{-5}
30,000	−83.6	66.4×10^{-6}
100,000	−94.0	19.9×10^{-6}
300,000	−103.6	6.64×10^{-6}
1,000,000	−114.1	1.97×10^{-6}

www.raguvaran.puzl.com

Table 14.6 gives the open-loop gain and pole parameters for a few op-amps of interest. Both parameters, but especially the gain, are subject to considerable variation; the typical values from the manufacturers' data sheets are given here.

Some of these op-amps have very high open-loop gains, but only at very low frequencies. This may be good for DC applications, but in audio line-input applications, where the lowest frequency of CMRR interest is 50 Hz, they will be operating above the pole frequency and so the gain available will be less – possibly considerably so, in the case of op-amps like the OPA2134. This is not, however, a real limitation, for even if a humble TL072 is used the

TABLE 14.5 The effect of op-amp open-loop pole frequency on CMRR for the circuit of Figure 14.5

Pole frequency	CMRR break-point frequency
10 kHz	10.2 kHz
1 kHz	1.02 kHz
100 Hz	102 Hz
10 Hz	10.2 Hz

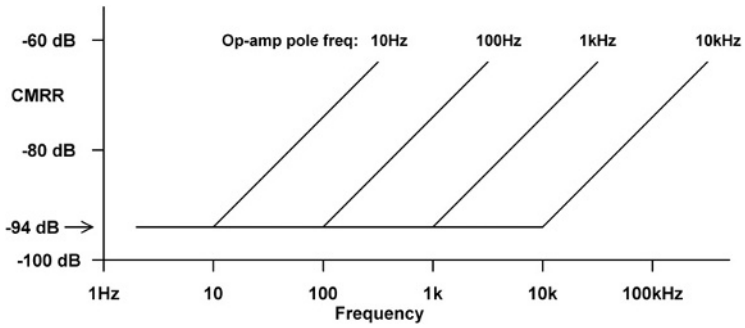


Figure 14.6: How the CMRR degrades with frequency for different op-amp pole frequencies. All resistors are assumed to be perfectly matched

TABLE 14.6 Typical LF gain and open-loop pole frequency for some op-amps commonly used in audio

Name	Input device type	LF gain	Pole frequency (Hz)	Op-amp LF CMRR (dB)
NE5532	Bipolar	100,000	100	100
LM4562	Bipolar	10,000,000	Below 10	120
LT1028	Bipolar	20,000,000	3	120
TL072	FET	200,000	20	86
OP27	FET	1,800,000	3	120
OPA2134	FET	1,000,000	3	100
OPA627	FET	1,000,000	20	116

www.raguvaran.puzl.com

perfect-resistor CMRR is about −90 dB, degrading from 100 Hz upwards. This sort of performance is not attainable in practice. We will shortly see why not.

Op-Amp CMRR Effects

Op-amps have their own CMRR, and we need to know how much this will affect the final CMRR of the balanced interconnection. The answer is that if all resistors are accurate, the overall CMRR is equal to the CMRR of the op-amp [5]. Since op-amp CMRR is typically very high (see the examples in Table 14.6), it is very unlikely to be the limiting factor.

The CMRR of an op-amp begins to degrade above a certain frequency, typically at 6 dB/octave. This is (fortunately) at a higher frequency than the open-loop pole, and is frequently around 1 kHz. For example, the OP27 has a pole frequency at about 3 Hz, but the CMRR remains flat at 120 dB until 2 kHz, and it is still greater than 100 dB at 20 kHz.

Amplifier Component Mismatch Effects

We saw earlier in this chapter that when the output and input impedances on a balanced line have a high ratio between them and are accurately matched we got a very good CMRR; this was compromised by the imperfections of op-amps, but the overall results were still very good – and much higher than the CMRRs measured in practice. There remains one place where we are still away in theory-land: we have so far assumed the resistances around the op-amp were all exactly accurate. We must now face reality, admit that these resistors will not be perfect, and see how much damage to the CMRR they will do.

SPICE simulation gives us Table 14.7. The situation with LF op-amp gains of both 100,000 and 1,000,000 is examined, but the effects of finite op-amp bandwidth or op-amp CMRR are not included. R1 in Figure 14.5 is varied while R2, R3, and R4 are all kept at precisely 10 kΩ, and the balanced output source impedances are set to exactly 100 Ω.

Table 14.7 shows with glaring clarity that our previous investigations, which took only output and input impedances into account, and determined that 68 Ω output resistors and 20 kΩ input impedances gave a CMRR of −89.5 dB for a 1% deviation in either, were actually quite unrealistic, and even adding in op-amp imperfections left us with unduly optimistic results. If a 1% tolerance resistor is used for R1 (and nowadays there is no financial incentive to use anything less accurate), the CMRR is dragged down at once to −46 dB; the same figure results from varying any other one of the four resistances by itself. If you are prepared to shell out for 0.1% tolerance resistors, the CMRR is a rather better −66 dB.

This shows that there really is no point in worrying about the gain of the op-amp you use in balanced inputs; the effect of mismatches in the resistors around that op-amp are far greater.

www.raguvaran.puzl.com

TABLE 14.7 How resistor tolerances affect the CMRR with some realistic op-amp open-loop gains

R1 (Ω)	R1 deviation (%)	Gain (×)	CMRR (dB)
10k	0	100,000	−94.0
10.001k	0.01	100,000	−90.6
10.01k	0.1	100,000	−66.5
10.1k	1	100,000	−46.2
11k	10	100,000	−26.6
10k	0	1,000,000	−114.1
10.001k	0.01	1,000,000	−86.5
10.01k	0.1	1,000,000	−66.2
10.1k	1	1,000,000	−46.2
11k	10	1,000,000	−26.6

The results in the table give an illustration of how resistor accuracy affects CMRR, but it is only an illustration, because in real life – a phrase that seems to keep cropping up, showing how many factors affect a practical balanced interconnection – all four resistors will of course be subject to a tolerance, and a more realistic calculation would produce a statistical distribution of CMRR rather than a single figure. One method is to use the Monte Carlo function in SPICE, which runs multiple simulations with random component variations and collates the results. However you do it you must know (or assume) how the resistor values are distributed with their tolerance window. Usually you don't know, and finding out by measuring hundreds of resistors is not a task that appeals to all of us.

It is straightforward to assess the worst-case CMRR, which occurs when all resistors are at the limit of the tolerance in the most unfavorable direction. The CMRR in dB is then:

$$\mathrm{CMRR} = 20\log\left(\frac{1+\mathrm{R2}/\mathrm{R1}}{4T/100}\right) \qquad \text{(Equation 14.1)}$$

where R1 and R2 are as in Figure 14.5, and T is the tolerance (%).

This deeply pessimistic equation tells us that 1% resistors give a worst-case CMRR of only 34.0 dB, that 0.5% parts give only 40.0 dB and expensive 0.1% parts yield but 54.0 dB. Things are not, however, quite that bad in actuality, as the chance of everything being as wrong as possible is actually very small. I have measured the CMRR of more of these balanced inputs, built with 1% resistors, than I care to contemplate, but I do not recall that I ever saw one with an LF CMRR worse than 40 dB.

There are eight-pin SIL packages that offer four resistors that ought to have good matching, if not accurate absolute values; be very, very wary of these as they usually contain thick-film

www.raguvaran.puzl.com

resistive elements that are not perfectly linear. In a test I did a 10 kΩ SIL resistor with 10 Vrms across it generated 0.0010% distortion. Not a huge amount perhaps, but in the quest for perfect audio, resistors that do not stick to Ohm's Law are not a good start.

To conclude this section, it is clear that in practical use it is the errors in the balanced amplifier resistors that determine the CMRR, though both unbalanced capacitances (C1, C2 in Figure 14.7) and the finite op-amp bandwidth are likely to cause further degradation at high audio frequencies. If you are designing both ends of a balanced interconnection and you are spending money on precision resistors, you should put them in the input amplifier, not the balanced output. The LF gain of the op-amp, and op-amp CMRR, have virtually no effect.

At this point it has to be said that balanced input amplifiers with four 1% resistors around them are used extensively in the audio business, and almost always prove to have adequate CMRR. When more CMRR is thought desirable, for example in high-end mixing consoles, one of the resistances is made trimmable with a preset. This means a lot of tweaking in manufacture, but the upside is that this is a quick set-and-forget adjustment that will not need to be touched again unless one of the four resistors needs replacing, and that is unlikely. CMRRs at LF of more than 60 dB can easily be obtained by this method, but the CMRR at HF will degrade due to the op-amp gain roll-off.

A Practical Balanced Input

The simple balanced input circuit shown in Figure 14.4 is not fit to face the outside world without additional components. Figure 14.7 shows a fully equipped version. Firstly, and most important,

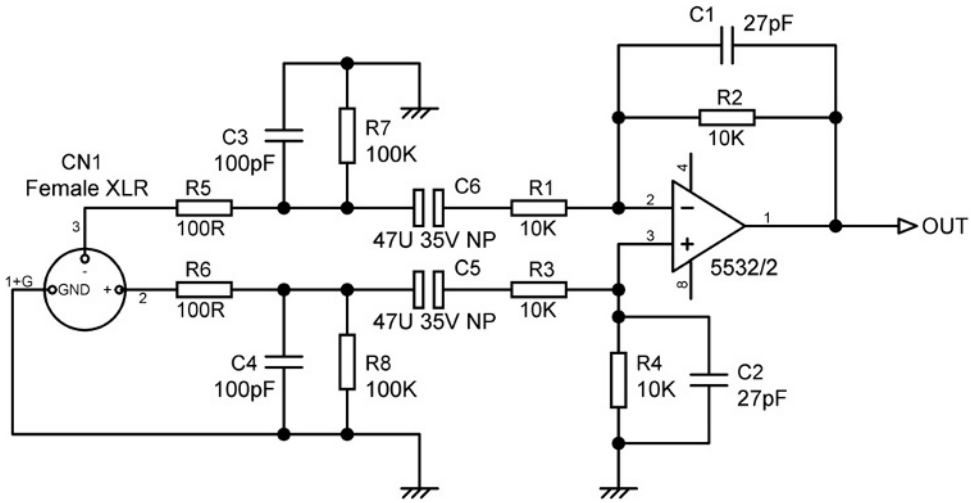


Figure 14.7: Balanced input amplifier with the extra components required for DC blocking and EMC immunity

www.raguvaran.puzl.com

C1 has been added across the feedback resistor R2; this prevents stray capacitances from pin 2 to ground causing extra phase-shifts that lead to HF instability. The value required for stability is small, much less than that which would cause an HF roll-off anywhere near the top of the audio band. The values here of 10k and 27 pF give −3 dB at 589 kHz, and such a roll-off is only down by 0.005 dB at 20 kHz. C2, of equal value, must be added across R4 to maintain the balance of the amplifier, and hence its CMRR, at high frequencies.

C1 and C2 must not be relied upon for EMC immunity as C1 is not connected to ground, and there is every chance that RF will demodulate at the op-amp inputs. A passive RF filter is therefore added to each input, in the shape of R5, C3 and R6, C4, so the capacitors will shunt incoming RF to ground before it reaches the op-amp. Put these as close to the input socket as possible to minimize radiation inside the enclosure.

I explained earlier in this chapter when looking at unbalanced inputs that it is not easy to guess what the maximum source impedance will be, given the existence of 'passive preamplifiers' and valve equipment. Neither are likely to have a balanced output, unless implemented by transformer, but either might be used to feed a balanced input, and so the matter needs thought.

In the unbalanced input, circuit resistances had to be kept as low as practicable to minimize the generation of Johnson noise that would compromise the inherently low noise of the stage. The situation with a standard balanced input is, however, different from the unbalanced case as there have to be resistances around the op-amp, and they must be kept up to a certain value to give acceptably high input impedances; this is why a balanced input like this one is much noisier. We could therefore make R5 and R6 much larger without a measurable noise penalty if we reduce R1 and R3 accordingly to keep unity gain. In Figure 14.7 R5 and R6 are kept at 100 Ω, so if we assume 50 Ω output resistances in both legs of the source equipment, then we have a total of 150 Ω, and 150 Ω and 100 pF give −3 dB at 10.6 MHz. Returning to a possible passive preamplifier with a 10 kΩ potentiometer, its maximum output impedance of 2.5k plus 100 Ω with 100 pF gives −3 dB at 612 kHz, which remains well clear of the top of the audio band.

As with the unbalanced input, replacing R5 and R6 with small inductors will give much better RF filtering but at increased cost. Ideally a common-mode choke (two bifilar windings on a small toroidal core) should be used as this improves performance. Check the frequency response to make sure the LC circuits are well damped and not peaking at the turnover frequency.

C5 and C6 are DC-blocking capacitors. They must be rated at no less than 35 V to protect the input circuitry, and are the non-polarized type as external voltages are of unpredictable polarity. The lowest input impedance that can occur with this circuit when using 10 kΩ resistors is, as described above, 6.66 kΩ when it is being driven in the balanced mode. The low-frequency roll-off is therefore −3 dB at 0.51 Hz. This may appear to be undesirably low, but the important point is not the LF roll-off but the possible loss of CMRR at low frequencies due to imbalance in the values of C5 and C6; they are electrolytics with a significant

www.raguvaran.puzl.com

tolerance. Therefore they should be made large so their impedance is a small part of the total input impedance; 47 μF is shown here but 100 or 220 μF can be used to advantage if there is the space to fit them in. The low-end frequency response must be defined somewhere in the input system, and the earlier the better, to prevent headroom or linearity being affected by subsonic disturbances, but this is not a good place to do it. A suitable time-constant immediately after the input amplifier is the way to go, but remember that capacitors used as time-constants may distort unless they are NP0 ceramic, polystyrene, or polypropylene (see Chapter 2 for more on this).

R7, R8 are DC drain resistors to prevent charges lingering on C5 and C6. These can be made lower than for the unbalanced input as the input impedances are lower, so a value of, say, 100 kΩ rather than 220 kΩ makes relatively little difference to the total input impedance.

A useful property of this kind of balanced amplifier is that it does not go mad when the inputs are left open-circuit – in fact it is actually *less* noisy than with its inputs shorted to ground. This is the opposite of the 'normal' behavior of a high-impedance unterminated input. This is because two things happen: open-circuiting the hot input doubles the resistance seen by the non-inverting input of the op-amp, raising its noise contribution by 3 dB. However, opening the cold input makes the noise gain drop by 6 dB, giving a net drop in noise output of approximately 3 dB. This of course refers only to the internal noise of the amplifier stage, and pickup of external interference is always possible on an unterminated input. The input impedances here are modest, however, and the problem is less serious than you might think. Having said that, deliberately leaving inputs unterminated is always bad practice.

If this circuit is built with four 10 kΩ resistors and a 5532 op-amp section, the noise output is −104.8 dBu with the inputs terminated to ground via 50 Ω resistors. As noted above, the input impedance of the cold input is actually lower than the resistor connected to it when working balanced, and if it is desirable to raise this input impedance to 10 kΩ, it could be done by raising the four resistors to 16 kΩ; this slightly degrades the noise output to −103.5 dBu. Table 14.8 gives some examples of how the noise output depends on the resistor value; the third column gives the noise with the input unterminated, and shows that in each case the amplifier is about 3 dB quieter when open-circuited. It also shows that a useful improvement in noise performance is obtained by dropping the resistor values to the lowest that a 5532 can

TABLE 14.8 Noise output (dBu) measured from simple balanced amps using a 5532 section

R value (Ω)	50 Ω terminated inputs	Open-circuit inputs	Terminated/open difference
100k	−95.3	−97.8	2.5
10k	−104.8	−107.6	2.8
2k0	−109.2	−112.0	2.8
820	−111.7	−114.5	2.8

www.raguvaran.puzl.com

easily drive (the op-amp has to drive the feedback resistor), though this usually gives unacceptably low input impedances. More on that at the end of the chapter.

Variations on the Balanced Input Stage

I now give a collection of balanced input circuits that offer advantages or extra features over the standard balanced input configuration. The circuit diagrams often omit stabilizing capacitors, input filters, and DC-blocking capacitors to improve the clarity of the basic principle. They can easily be added; in particular, bear in mind that a stabilizing capacitor like C1 in Figure 14.7 is often needed between the op-amp output and the negative input to guarantee freedom from high-frequency oscillation.

Combined Unbalanced and Balanced Inputs

If both unbalanced and balanced inputs are required, it is extremely convenient if it can be arranged so that no switching between them is required. Switches cost money, mean more holes in the metalwork, and add to assembly time. Figure 14.8 shows an effective way to implement this. In balanced mode, the source is connected to the balanced input and the unbalanced input left unterminated. In unbalanced mode, the source is connected to the unbalanced input and the balanced input left unterminated, and no switching is required. It might appear that these unterminated inputs would pick up extra noise, but in practice this is not the case. It works very well and I have used it successfully in high-end equipment for two prestigious manufacturers.

As described above, in the world of hi-fi, balanced signals are at twice the level of the equivalent unbalanced signals, and so the balanced input must have a gain of 0.5 or −6 dB

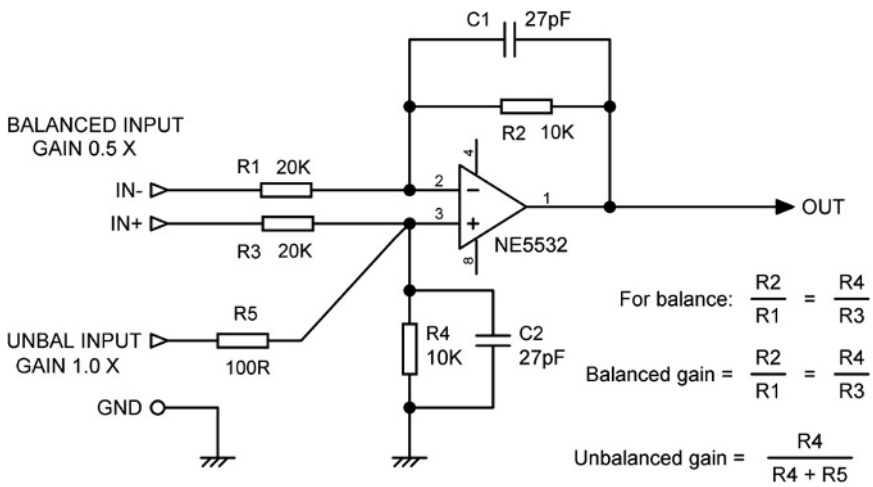


Figure 14.8: Combined balanced and unbalanced input amplifier with no switching required

www.raguvaran.puzl.com

relative to the unbalanced input to get the same gain by either path. This is done here by increasing R1 and R3 to 20 kΩ. The balanced gain can be greater or less than unity, but the gain via the unbalanced input is always 1. The differential gain of the amplifier and the constraints on the component values for balanced operation are shown in Figure 14.4, and are not repeated in the text to save space. This applies to the rest of the balanced inputs in this chapter.

There are two minor compromises in this circuit that need to be noted. Firstly, the noise performance in unbalanced mode is worse than for the dedicated unbalanced input described earlier in this chapter, because R2 is effectively in the signal path and adds Johnson noise. Secondly, the input impedance of the unbalanced input cannot be very high because it is set by R4, and if this is increased in value all the resistances must be increased proportionally and the noise performance will be markedly worse. It is important that only one input cable should be connected at a time, because if an unterminated cable is left connected to an unused input, the cable capacitance to ground can cause frequency response anomalies and might in adverse circumstances cause HF oscillation. A prominent warning on the back panel and in the manual is a very good idea.

The Superbal Input

This version of the balanced input amplifier, shown in Figure 14.9, has been referred to as the 'Superbal' circuit because it gives equal impedances into the two inputs for differential signals. It was originated by David Birt of the BBC [6]. With the circuit values shown the

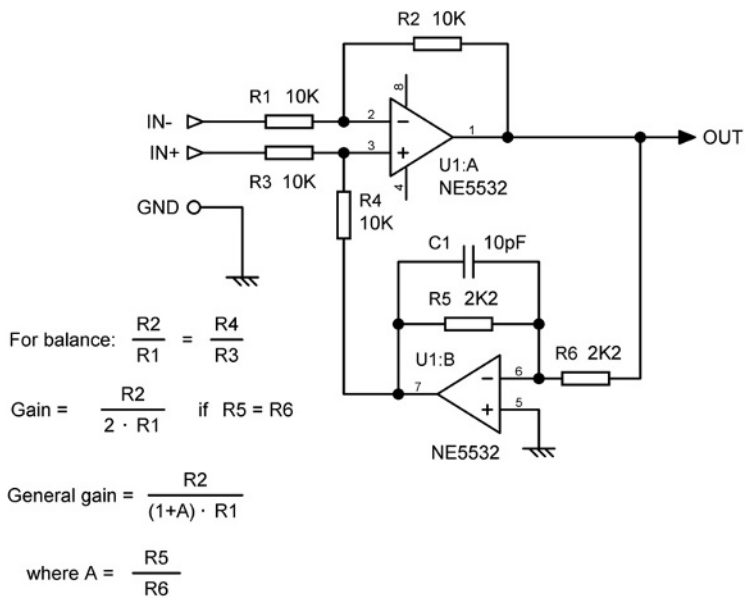


Figure 14.9: The Superbal balanced input amplifier

www.raguvaran.puzl.com

differential input impedance is exactly 10 kΩ via both hot and cold inputs. The common-mode input impedance is 20 kΩ as before.

In the standard balanced input R4 is connected to ground, but here its lower end is actively driven with an inverted version of the output signal, giving symmetry. The increased amount of negative feedback reduces the gain with four equal resistors to −6 dB instead of unity. The gain can be reduced below −6 dB by giving the inverter a gain of more than 1; if R1, R2, R3, and R4 are all equal, the gain is $1/(A + 1)$, where A is the gain of the inverter stage. This is of limited use as the inverter U1:B will now clip before the forward amplifier U1:A, reducing headroom. If the gain of the inverter stage is gradually reduced from unity to zero, the stage slowly turns back into a standard balanced amplifier with the gain increasing from −6 dB to unity and the input impedances becoming less and less equal. If a gain of less than unity is required it should be obtained by increasing R1 and R3.

R5 and R6 should be kept as low in value as possible to minimize Johnson noise; there is no reason why they have to be equal in value to R1, etc. The only restriction is the ability of U1:A to drive R6 and U1:B to drive R5, both resistors being effectively grounded at one end. The capacitor C1 will almost certainly be needed to ensure HF stability; the value in the figure is only a suggestion. It should be kept as small as possible because reducing the bandwidth of the inverter stage impairs CMRR at high frequencies.

Switched-Gain Balanced Inputs

A balanced input stage that can be switched to two different gains while maintaining CMRR is very useful. Equipment often has to give optimal performance with both semi-pro (−7.8 dBu) and professional (+4 dBu) input levels. If the nominal internal level of the system is in the normal range of −2 to −6 dBu, the input stage must be able to switch between amplifying and attenuating, while maintaining good CMRR in both modes.

The brute-force way to change gain in a balanced input stage is to switch the values of either R1 and R3, or R2 and R4, in Figure 14.4, keeping the pairs equal in value to maintain the CMRR; this needs a double-pole switch for each input channel. A much more elegant technique is shown in Figure 14.11. Perhaps surprisingly, the gain of a differential amplifier can be manipulated by changing the drive to the feedback arm (R2 etc.) only, and leaving the other arm R4 unchanged, without affecting the CMRR. The essential point is to keep the source resistance of the feedback arm the same, but drive it from a scaled version of the op-amp output. Figure 14.10 does this with the network R5, R6, which has a source resistance made up of 6k8 in parallel with 2k2, which is 1.662 kΩ. This is true whether R6 is switched to the op-amp output (low gain setting) or to ground (high gain setting), for both have effectively zero impedance. For low gain the negative feedback is not attenuated, but fed through to R2 and R7 via R5, R6 in parallel. For high gain R5 and R6 become a potential divider, so the

www.raguvaran.puzl.com

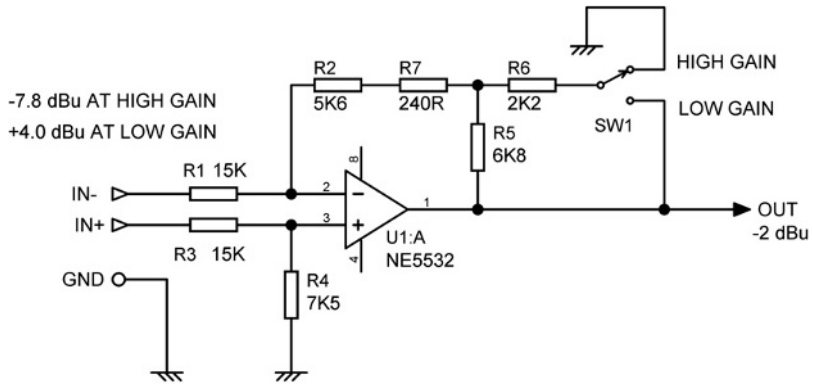


Figure 14.10: A balanced input amplifier with gain switching that maintains good CMRR

amount of feedback is decreased and the gain increased. The value of R2 + R7 is reduced from 7k5 by 1.662 kΩ to allow for the source impedance of the R5, R6 network; this requires the distinctly non-standard value of 5.838 kΩ that is here approximated by R2 and R7, which give 5.6 kΩ + 240 Ω = 5.840 kΩ. This value is the best that can be done with E24 resistors; it is obviously out by 2 Ω, but that is much less than a 1% tolerance on R2, and so will have only a vanishingly small effect on the CMRR.

Note that this stage can attenuate as well as amplify if R1, R3 are set to be greater than R2, R4, as shown here. The nominal output level of the stage is assumed to be −2 dBu; with the values shown the two gains are −6.0 and +6.2 dB, so +4 and −7.8 dBu respectively will give −2 dBu at the output. Other pairs of gains can of course be obtained by changing the resistor values; the important thing is to stick to the principle that the value of R2 + R7 is reduced from the value of R4 by the source impedance of the R5, R6 network. With the values shown the differential input impedance is 11.25 kΩ via the cold and 22.5 kΩ via the hot input. The common-mode input impedance is 22.5 kΩ.

This neat little circuit has the added advantage that nothing bad happens when the switch is moved with the circuit operating. When the wiper is between contacts you simply get a gain intermediate between the high and low settings, which is pretty much the ideal situation.

Variable-Gain Balanced Inputs

The beauty of a variable-gain balanced input is that it allows you to get the incoming signal up or down to the nominal internal level as soon as possible, minimizing both the risk of clipping and contamination with circuit noise. The obvious method of making a variable-gain differential stage is to use dual-gang pots to vary either R1, R3 or R2, R4 together, to maintain CMRR. This is clumsy, and gives a CMRR that is both bad and highly variable due to the

www.raguvaran.puzl.com

inevitable mismatches between pot sections. For a stereo input the required four-gang pot is an unappealing proposition.

There is, however, a way to get a variable gain with good CMRR, using a single pot section. The principle is essentially the same as for the switched-gain amplifier above: keep constant the source impedance driving the feedback arm, but vary the voltage applied. The principle is shown in Figure 14.11. To the best of my knowledge I invented this circuit in 1982; any comments on this point are welcome. The feedback arm R2 is driven by voltage-follower U1:B. This eliminates the variations in source impedance at the pot wiper, which would badly degrade the CMRR. R6 limits the gain range and R5 modifies the gain law to give it a more usable shape. When the pot is fully up (minimum gain) R5 is directly across the output of U1:A so do not make it too low in value. If a center-detent pot is used to give a default gain setting, this may not be very accurate as it partly depends on the ratio of pot track (no better than ±10% tolerance, and sometimes worse) to 1% fixed resistors.

This configuration is very useful as a general line input with an input sensitivity range of −20 to +10 dBu. For a nominal output of 0 dBu, the gain of Figure 14.11 is +20 to −10 dB, with R5 chosen for 0 dB gain at the central wiper position. An op-amp in a feedback path may appear a dubious proposition for HF stability, because of the extra phase-shift it introduces, but here it is working as a voltage-follower, so its bandwidth is maximized and in practice the circuit is dependably stable.

High-Input-Impedance Balanced Inputs

We saw earlier that high input impedances are required to maximize CMRR of a balanced interconnection, but the input impedances offered by the standard balanced circuit are limited by the need to keep the resistor values down to control Johnson noise. High-impedance balanced inputs are also useful for interfacing to valve equipment in the strange world of retro

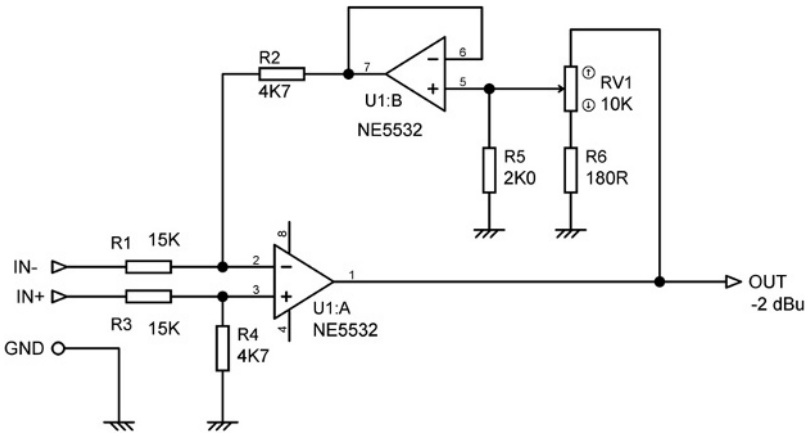


Figure 14.11: Variable-gain balanced input amplifier

www.raguvaran.puzl.com

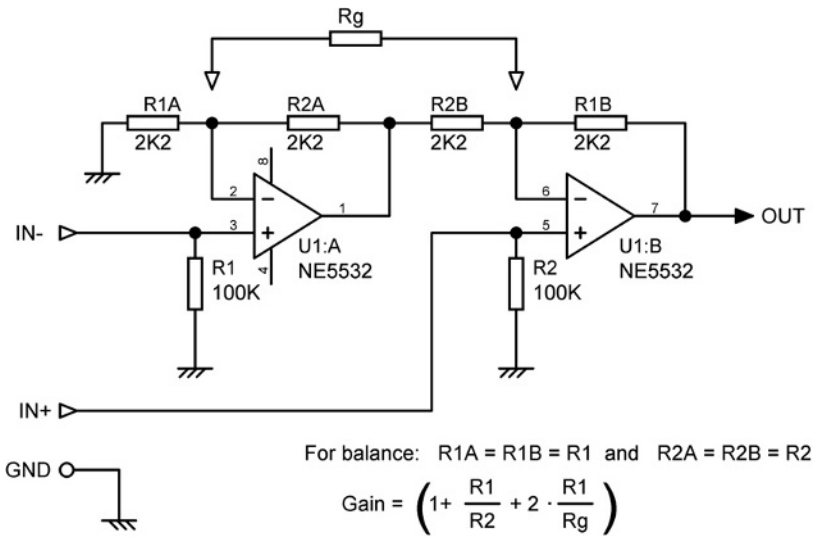


Figure 14.12: High-input-impedance balanced input

hi-fi. Adding output cathode-followers to valve circuitry is expensive and consumes a lot of extra power, and so the output is often taken directly from the anode of a gain stage, and even a so-called bridging load of 10 kΩ may seriously compromise the distortion performance and output capability of the source equipment.

Figure 14.12 shows a configuration where the input impedances are determined only by the bias resistances R1 and R2. They are shown here as 100 kΩ, but may be considerably higher if op-amp bias currents permit. A useful property of this circuit is that adding a single resistor R_g increases the gain, but preserves the circuit balance and CMRR. This configuration cannot be set to attenuate because the gain of an op-amp with series feedback cannot be reduced below unity.

It is of course always possible to give a basic balanced input a high input impedance by putting unity-gain buffers in front of each input, but that uses three op-amp sections rather than two. Sometimes, however, it is appropriate. Much more on that later.

We noted earlier that the simple balanced input is surprisingly quiet and well behaved when its inputs are unterminated. This is not the case with this configuration, which because of its high input impedances will be both noisy and susceptible to picking up external interference.

The Inverting Two-Op-Amp Input

The configuration depicted in Figure 14.13 has its uses because the hot and cold inputs have the same impedances for differential signals, as well as for common-mode voltages [7]. It is not suited to high input impedances at normal gains because high resistor values would have to be

www.raguvaran.puzl.com

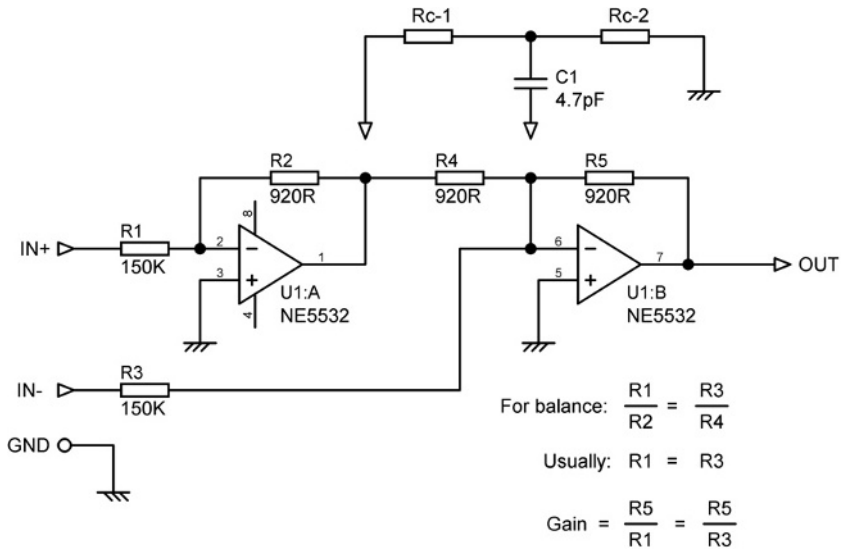


Figure 14.13: Inverting two-op-amp balanced input. Values shown interface a 100 V line to −2 dBu internal level

used throughout and they would generate excess Johnson noise, but if it is interfacing with a high voltage source so the gain must be well below unity, R1 and R3 can be made high in value, and R2, R4, R5 set low, the latter components keeping the noise down. The CMRR may degrade at HF because the hot signal has gone through an extra op-amp and suffered phase-shift, interfering with the subtraction; this can be compensated for by the network Rc-1, Rc-2, and C1; the values needed depend on op-amp type and must be checked by CMRR measurements.

I have used this configuration for balanced input feeds from a 100 V loudspeaker line of the sort still in use for distributing audio over wide areas. The circuit values shown reduce the 100 V input to a nominal −2 dBu (615 mV) internal level. If this circuit is used for that purpose then effective input overvoltage protection on the inputs is essential as large voltage transients are possible on a loudspeaker line if parts of it are unplugged or plugged with signal present, due to the inductance of the line-matching transformers. This protection can be conveniently provided by the usual diode clamps to the supply rails; this is dealt with at the end of this chapter. The input resistors will be of high value in this application so there is very little possibility of excessive inputs overheating the resistors or 'pumping up' the op-amp supply rails.

The Instrumentation Amplifier

Almost every book on balanced or differential inputs includes the three-op-amp circuit of Figure 14.14 and praises it as the highest expression of the differential amplifier. It is usually called the instrumentation amplifier configuration because of its undoubted superiority for data acquisition. (Specialized ICs exist that are sometimes also called instrumentation

www.raguvaran.puzl.com

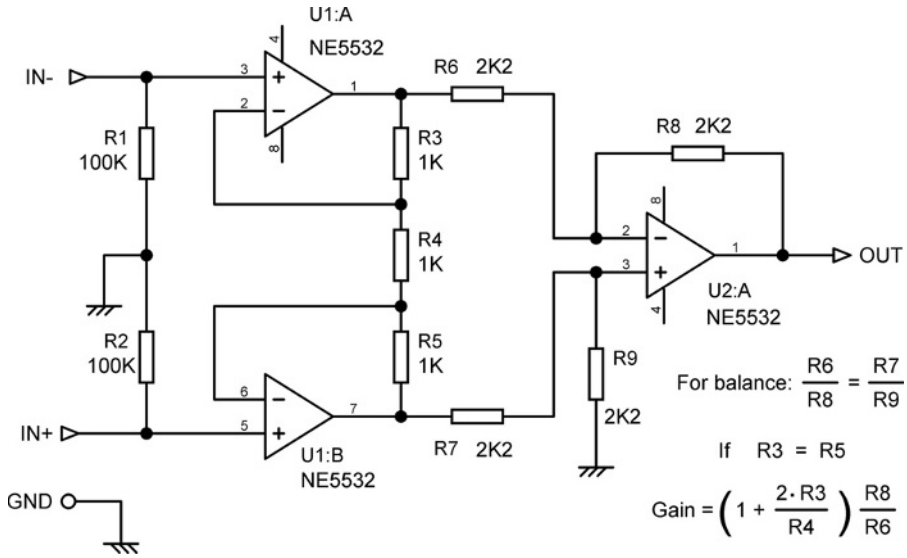


Figure 14.14: The instrumentation amplifier configuration. Gain here is 3 times

amplifiers or in-amps; these are designed for very high CMRR data acquisition. They are expensive and in general not optimized for audio work.)

The beauty of the three-op-amp configuration is that the dual input stage buffers the balanced line from the input impedances of the final differential amplifier; the four resistances around it can now be made much lower in value, reducing their Johnson noise by a significant amount, while keeping the CMRR benefits of presenting high input impedances to the balanced line. The other feature, which is usually much more emphasized because of its unquestionable elegance, is that the dual input stage, with its shared feedback network R3, R4, R5, can be set to have a high differential gain by giving R4 a low value, but its common-mode gain is always unity; this property is not affected by mismatches in R3 and R5. The final amplifier then does its usual job of common-mode rejection, and the combined CMRR can be very good indeed if the first stage gain is high.

This is all well and good, but as it stands this stage is not very useful for audio balanced line inputs. A data-acquisition application like ECG monitoring may need a gain of thousands of times, which will allow a stunning CMRR to be achieved without using precision resistors, but the cruel fact is that, in audio usage, gain at this point is very often simply not wanted. In a typical signal path comprised of op-amps, the nominal internal level is usually between −6 and 0 dBu, and the level coming in is at the professional level of +4 dBu; what is needed is 6 dB of attenuation rather than any gain. Gain at this point and attenuation later would introduce what can only be described as a headroom bottleneck. If the incoming level was the semi-pro −7.8 dBu a small amount of gain could be introduced, but then the CMRR advantage would be equally small, and certainly not worth the cost of two op-amp sections.

www.raguvaran.puzl.com

Transformer Balanced Inputs

When it is essential that there is no galvanic connection (i.e. no electrical conductor) between two pieces of equipment, transformer inputs are indispensable. They are also useful if EMC conditions are severe. Figure 14.15 shows a typical transformer input. The transformer usually has a 1:1 ratio, and is enclosed in a metal shielding can that must be grounded. Good line transformers have an interwinding shield that must also be grounded or the high-frequency CMRR will be severely compromised. The transformer secondary must see a high impedance as this is reflected to the primary and represents the input impedance; here it is set by R2, and a buffer drives the circuitry downstream. In addition, if the secondary loading is too heavy there will be increased transformer distortion at low frequencies. Line-input transformers are built with small cores and are only intended to deliver very small amounts of power; they are *not* interchangeable with line-output transformers. A most ingenious approach to dealing with this distortion problem by operating the input transformer core at near-zero flux was published by Paul Zwicky in 1986 [8].

There is a bit more to correctly loading the transformer secondary. If it is simply loaded with a high-value resistor there will be peaking of the frequency response due to resonance between the transformer leakage inductance and the winding capacitance. This is shown in Figure 14.16, where a Sowter 3276 line input transformer (a high-quality component) was given a basic resistive loading of 100 kΩ. The result was trace A, which has a 10 dB peak around 60 kHz. This is bad not only because it accentuates the effect of out-of-band noise, but because it intrudes on the audio frequency response, giving a lift of 1 dB at 20 kHz. Reducing the resistive load R2 would damp the resonance, but it would also reduce the input impedance. The answer is to add a Zobel network, i.e. a resistor and capacitor in series, across the secondary; this has no effect except at high frequencies. The first attempt used R1 = 2k7 and C1= 1 nF, giving trace B, where the peaking has been reduced to 4 dB around 40 kHz, but the 20 kHz lift is actually slightly greater. R1 = 2k7 and C1= 2 nF gave trace C, which is a bit

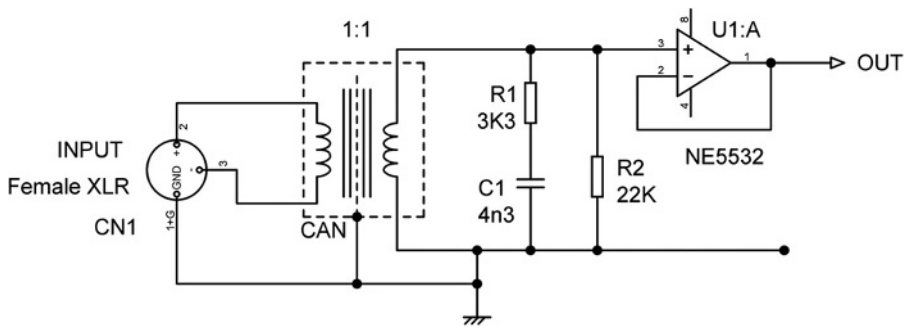


Figure 14.15: A transformer balanced input. R1 and C1 are the Zobel network that damps the secondary resonance

www.raguvaran.puzl.com

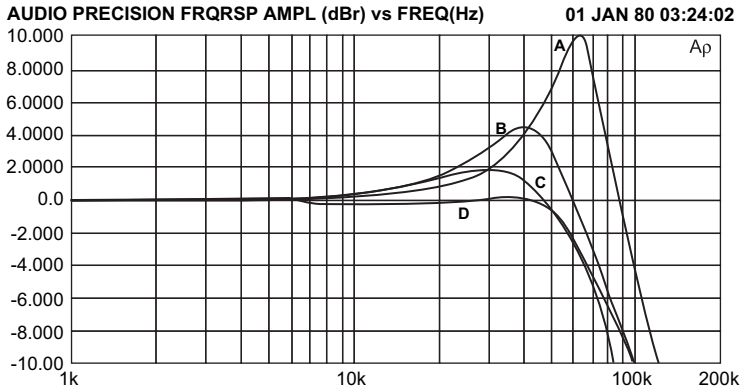


Figure 14.16: Optimizing the frequency response of a transformer balanced input with a Zobel network

better in that it only has a 2 dB peak. A bit more experimentation ended up with R1 = 3k3 and C1= 4.3 nF (3n3 + 1 nF) and yielded trace D, which is pretty flat, though there is a small droop around 10 kHz. The Zobel values are fairly critical for the flattest possible response, and must certainly be adjusted if the transformer type is changed.

Input Overvoltage Protection

Input overvoltage protection is not common in hi-fi applications, but is regarded as essential in most professional equipment. The normal method is to use clamping diodes, as shown in Figure 14.18, that prevent certain points in the input circuitry from moving outside the supply rails.

This is straightforward, but there are two points to watch. Firstly, the ability of this circuit to withstand excessive input levels is not without limit. Sustained overvoltages may burn out R5 and R6, or pump unwanted amounts of current into the supply rails; this sort of protection is mainly aimed at transients. Secondly, diodes have a non-linear junction capacitance when they are reverse biased, so if the source impedance is significant the diodes will cause distortion at high frequencies. To quantify this problem here are the results of a few tests. If Figure 14.17 is fed from the low impedance of the usual kind of line-output stage, the impedance at the diodes will be about 1 kΩ and the distortion induced into an 11 Vrms, 20 kHz input will be below the noise floor. However, if the source impedance is high so the impedance at the diodes is increased to 10 kΩ, with the same input level, the THD at 20 kHz is degraded from 0.0030% to 0.0044% by adding the diodes. I have thought up a rather elegant way to eliminate this effect completely, but this is not the place to disclose it.

www.raguvaran.puzl.com

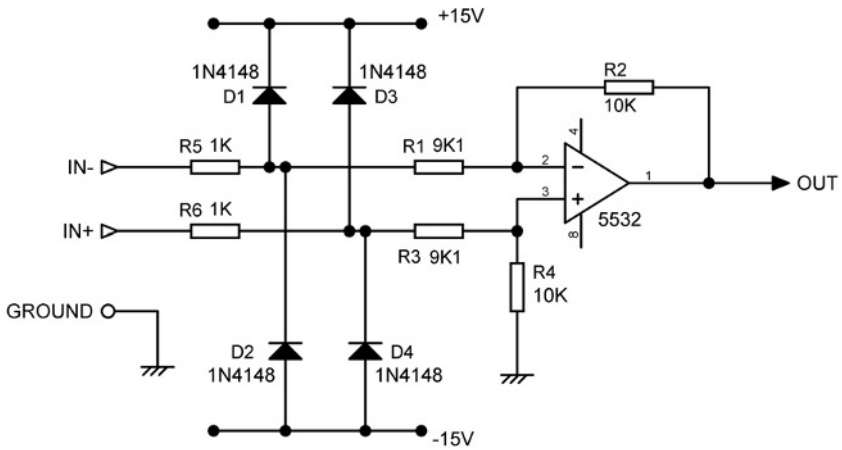


Figure 14.17: Input overvoltage protection for a balanced input amplifier

Low-Noise Balanced Inputs

So far we have not said much about the noise performance of balanced inputs, though on our way through the chapter we have remarked several times that the standard balanced input amplifier with four 10 kΩ resistors shown in Figure 14.18(a) is markedly noisier than an unbalanced input like that in Figure 14.1. The unbalanced input stage, with its input terminated by 50 Ω to ground, has a noise output of −119.0 dBu over the usual 22 Hz – 22 kHz bandwidth. If the balanced circuit is built with 10 kΩ resistors and a 5532 section, the noise output is −104.8 dBu with the inputs similarly terminated. This is a big difference of 14.2 dB.

In the hi-fi world in particular, where an amplifier may have both unbalanced and balanced inputs, many people would feel that this situation is the wrong way round. Surely the balanced input, with its professional XLR connector and its much-vaunted rejection of ground noise, should show a better performance in all departments? Well, it does – except as regards internal noise, and a 14 dB discrepancy is both clearly audible and hard to explain away. This section explains how to design it away instead.

We know that the source of the extra noise is the relatively high resistor values around the op-amp (see Table 14.8 earlier in the chapter), but these cannot be reduced in the simple balanced input without reducing the input impedances below what is acceptable. The answer is to lower the resistor values but buffer them from the input with a pair of voltage-followers. This arrangement is shown in Figure 14.18(b); 5532s are a good choice for this as they combine low voltage noise with low distortion and good load-driving capability. Since the input buffers are voltage-followers with 100% feedback, their gain is very accurately defined at unity and CMRR is not degraded; CMRR is still defined by the resistor tolerances, and by the bandwidth of the differential op-amp.

www.raguvaran.puzl.com

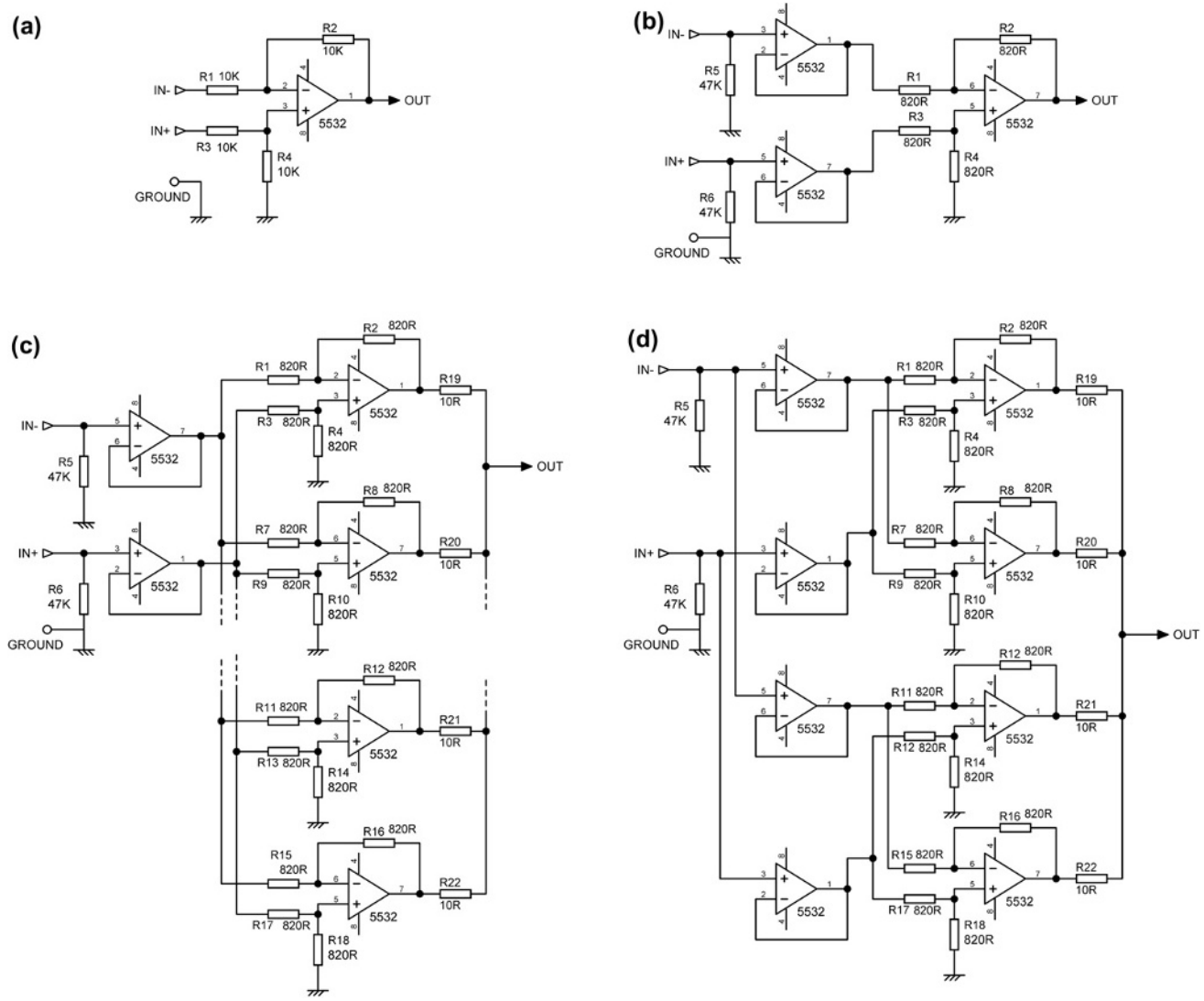


Figure 14.18: Low-noise unity-gain balanced inputs using multiple 5532 buffers and differential amplifiers

www.raguvaran.puzl.com

There is a limit to how far the four resistors can be reduced, as the differential stage has to be driven by the input buffers, and it also has to drive its own feedback arm. If 5532s are used a safe value that gives no measurable deterioration of the distortion performance is about 820 Ω, and a 5532 differential stage alone (without the buffers) and 4 × 820 Ω resistors gives a noise output of −111.7 dBu, which is 6.6 dB lower than the standard 4 × 10 kΩ version. Adding the two input buffers degrades this only slightly to −110.2 dB, because we are adding only the voltage noise component of the two new op-amps, and we are still 5.1 dB quieter than the original 4 × 10 kΩ version. It is interesting that we now have three op-amps in the signal path instead of one, but we still get a significantly lower noise level.

This might appear to be all we can do; it is not possible to reduce the value of the four resistors around the differential amplifier any further without compromising linearity. However, there is almost always some way to go further in the great game that is electronics, and here are three possibilities. A step-up transformer could be used to exploit the low source impedance (remember we are still assuming the source impedances are 50 Ω) and it might well work superbly in terms of noise alone, but transformers are always heavy, expensive, susceptible to magnetic fields, and of doubtful low-frequency linearity. We would also very quickly run into headroom problems; balanced line input amplifiers are normally required to attenuate rather than amplify.

We could design a discrete op-amp hybrid stage with discrete input transistors, which are quieter than those integrated into IC op-amps, coupled to an op-amp to provide raw loop gain; this can be quite effective but you need to be very careful about high-frequency stability, and it is difficult to get an improvement of more than 6 dB. Thirdly, we could design our own op-amp using all discrete parts; this approach tends to have fewer stability problems as all circuit parameters are accessible, but it definitely requires rather specialized skills, and the result takes up a lot of PCB area.

Since none of those three approaches is appealing, now what? One of the most useful techniques in low-noise electronics is to use two identical amplifiers so that the gains add arithmetically, but the noise from the two separate amplifiers, being uncorrelated, partially cancels. Thus, we get a 3 dB noise advantage each time the number of amplifiers used is doubled. This technique works very well with multiple op-amps, as expounded in Chapter 1; let us apply it and see how far it may be usefully taken.

Since the noise of a single 5532-section unity-gain buffer is only −119.0 dBu, and the noise from the 4 × 820 Ω differential stage (without buffers) is a much higher −111.7 dBu, the differential stage is clearly the place to start work. We will begin by using two identical 4 × 820 Ω differential amplifiers as shown in the top section of Figure 14.18(c), both driven from the existing pair of input buffers. This will give partial cancellation of both resistor and op-amp noise from the two differential stages if their outputs are summed. The main question is how to

www.raguvaran.puzl.com

sum the two amplifier outputs; any active solution would introduce another op-amp, and hence more noise, and we would almost certainly wind up worse off than when we started. The answer is, however, beautifully simple. We just connect the two amplifier outputs together with 10 Ω resistors; the gain does not change but the noise output drops. The signal output of both amplifiers is nominally the same, so no current should flow from one op-amp output to the other. In practice there will be slight gain differences due to resistor tolerances, but with 1% resistors I have never experienced any hint of a problem. The combining resistor values are so low at 10 Ω that their Johnson noise contribution is negligible.

The use of multiple differential amplifiers has another advantage – the CMRR errors are also reduced in the same way that the noise is reduced. This is also similar to the use of multiple parallel capacitors to improve RIAA accuracy, as explained in Chapter 7.

We therefore have the arrangement of Figure 14.18(c), with single input buffers (i.e. one per input) and two differential amplifiers, and this reduces the noise output by 2.3 dB to −112.5 dBu, which is quieter than the original 4 × 10 kΩ version by an encouraging 7.4 dB. We do not get a full 3 dB noise improvement because both differential amplifiers are handling the noise from the input buffers, which is correlated and so is not reduced by partial cancellation. The contribution of the input buffer noise is further brought out if we take the next step of using four differential amplifiers. There is of course nothing special about using amplifiers in powers of two. It is perfectly possible to use three or five differential amplifiers in the array, which will give intermediate amounts of noise reduction. If you have a spare op-amp section, then put it to work!

So, leaving the input buffers unchanged, we use them to drive an array of four differential amplifiers. These are added on at the dotted lines in the lower half of Figure 14.18(c). We get a further improvement, but only by 1.5 dB this time. The output noise is down to −114.0 dBu, quieter than the original 4 × 10 kΩ version by 8.9 dB. You can see that at this point we are proceeding by decreasing steps, as the input buffer noise is starting to dominate, and there seems little point in doubling up the differential amplifiers again; the amount of hardware could be regarded as a bit excessive, and so would the PCB area occupied. The increased loading on the input buffers is also a bit of a worry.

A more fruitful approach is to tackle the noise from the input buffers, by doubling them up as in Figure 14.18(d), so that each buffer drives only two of the four differential amplifiers. This means that the buffer noise will also undergo partial cancellation, and will be reduced by 3 dB. There is, however, still the contribution from the differential amplifier noise, and so the actual improvement on the previous version is 2.2 dB, bringing the output noise down to −116.2 dBu, which is quieter than the original 4 × 10 kΩ version by a thumping 11.1 dB. Remember that there are two inputs, and 'double buffers' means two buffers per hot and cold input, giving a total of four in the complete circuit.

www.raguvaran.puzl.com

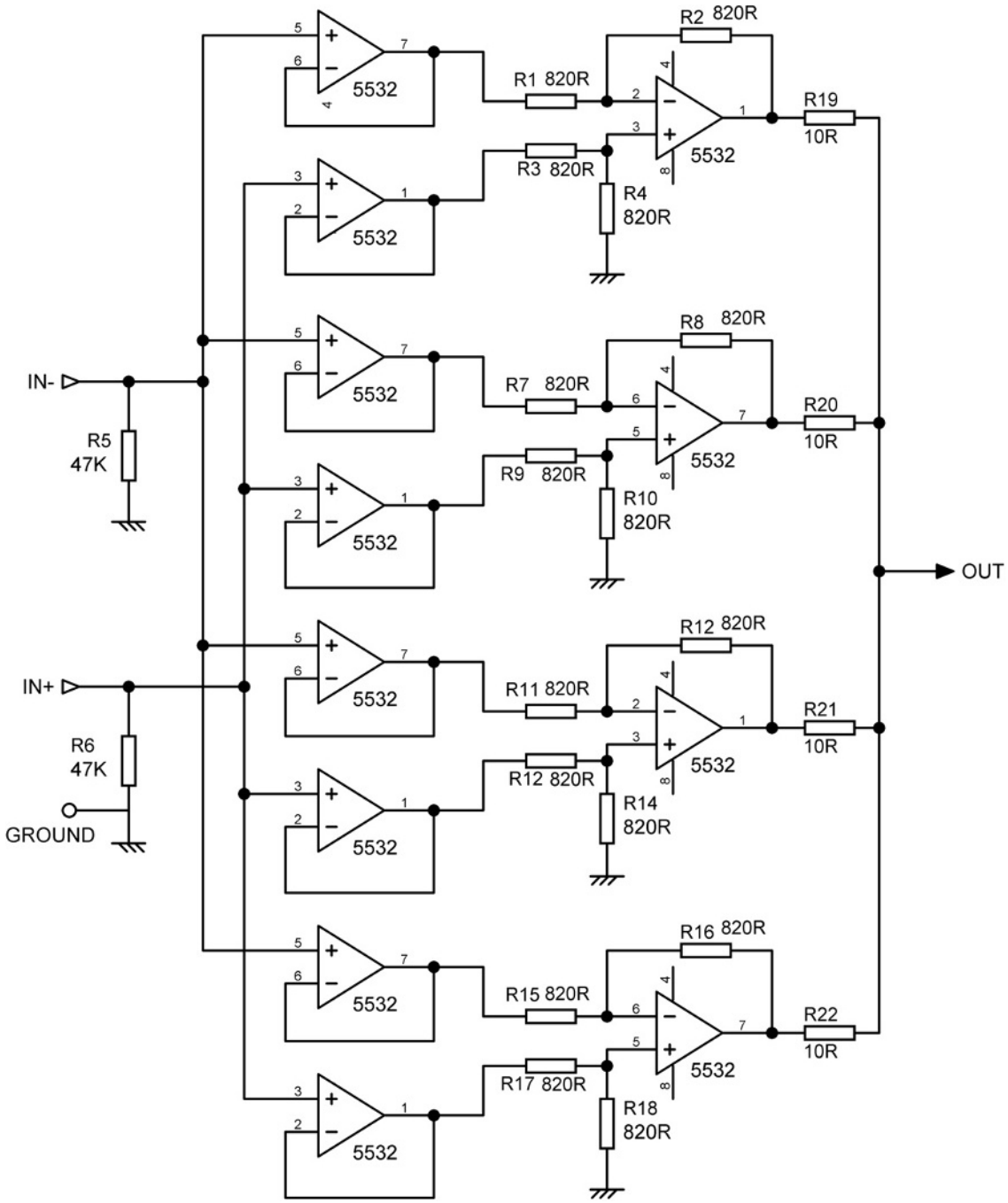


Figure 14.19: The final 5532 low-noise unity-gain balanced input stage, with quad input buffers and four differential amplifiers. The noise output is only −117.0 dBu

www.raguvaran.puzl.com

Since doubling up the input buffers gave us a useful improvement, it's worth trying again, so we have a structure with quad buffers and four differential amplifiers, as shown in Figure 14.19, where each differential amplifier now has its very own buffer. This improves on the previous version by a rather less satisfying 0.8 dB, giving an output noise level of −117.0 dBu, quieter than the original 4 × 10 kΩ version by 11.9 dB. The small improvement we have gained indicates that the focus of noise reduction needs to be returned to the differential amplifier array, but the next step would seem to be using eight amplifiers, which is not very appealing. Thoughts about ears of corn on chessboards tend to intrude at this point.

This is a good moment to pause and see what we have achieved. We have built a balanced input stage that is quieter than the standard circuit by 11.9 dB, using standard components of low cost. We have used increasing numbers of them, but the total cost is still small compared with enclosures, power supplies, etc. On the other hand, the power consumption is clearly several times greater. The technology is highly predictable and the noise reduction reliable, in fact bulletproof. The linearity is as good as that of a single op-amp of the same type, and in the same way there are no HF stability problems.

What we have not done is build a balanced input that is quieter than the unbalanced one – we are still 2.0 dB short of that target, but at least we have reached a point where the balanced input is not obviously noisier.

These noise results are summarized in Table 14.10 at the end of the chapter.

Low-Noise Balanced Inputs in Real Life

Please don't think that this examination of low-noise input options is merely a voyage off into pure theory. It has real uses and has been applied in practice. The Cambridge Audio 840W power amplifier is a design of mine which, I might modestly mention in passing, won a CES Innovation Award in January 2008. This unit has both unbalanced and balanced inputs, and conventional technology would have meant that the balanced inputs would have been significantly the noisier of the two. Since the balanced input is the 'premium' input, many people would think there was something amiss with this state of affairs. We therefore decided the balanced input had to be quieter than the unbalanced input. Using 5532s in an architecture similar to those outlined above, this requirement proved straightforwardly attainable, and the final balanced input design was both economical and quieter than its unbalanced neighbor by a dependable 0.9 dB. Two other versions were evaluated that made the balanced input quieter than the unbalanced one by 2.8 dB, and by 4.7 dB, at somewhat greater cost and complexity. These were put away for possible future upgrades.

The Signal Transfer Company [9] manufactures a low-noise balanced input card based on these principles that has 47 kΩ input impedances, unity gain, and a noise output of only −115 dBu.

www.raguvaran.puzl.com

Ultra-Low-Noise Balanced Inputs

In the section on low-noise balanced inputs above, we reclined briefly on our laurels, having achieved an economical balanced input stage with output noise at the extremely low level of −117.0 dBu. Regrettably this is still 2 dB noisier than a simple unbalanced input. It would be wrong to conclude from this that the resources of electronic design are exhausted. At the end of the noise-reduction sequence we were aware that the dominant noise source was currently the differential amplifier array, and we shrank from doubling up again to use eight amplifiers because of issues of cost and the PCB area occupied. We will take things another step by taking a much more relaxed view of cost (as one does at the 'high end') and see how that changes the game. We will, however, retain some concern about PCB area.

An alternative way to make the differential amplifier array quieter is simply to use op-amps that are quieter. These will inevitably be more expensive – much more expensive – than the ubiquitous 5532. Because of the low resistor values around the op-amps, we need to focus on low voltage noise rather than low current noise, and there are several that are significantly better than the 5532, as shown by the typical noise density figures in Table 14.9.

Clearly, moving to the 5534A will give a significant noise reduction, but since there is only a single op-amp per package, and external compensation is needed, the board area used will be much greater. The new chip on the block is the LM4562, a bipolar op-amp that has finally surpassed the 5532 in performance. The input voltage noise density is typically 2.7 nV/$\sqrt{\text{Hz}}$, substantially lower than the 5 nV/$\sqrt{\text{Hz}}$ of the 5532. For applications with low source impedances, this implies a handy noise advantage of 5 dB or more.The LM4562 is a dual op-amp and will not take up more space. At the time of writing it is something like 10 times more expensive than the 5532.

Step 1 – we replace all four op-amps in the differential amplifiers with LM4562s. They are a drop-in replacement with no circuit adjustments required at all. We leave the quad 5532 input buffers in place. The noise output drops by an impressive 1.9 dB, giving an output noise level of −118.9 dBu, quieter than the original 4 × 10 kΩ version by 14.1 dB, and only 0.1 dB noisier than the unbalanced stage.

TABLE 14.9 Voltage and noise densities for low-noise balanced-input op-amp candidates

Op-amp	Voltage noise density (nV/$\sqrt{\text{Hz}}$)	Current noise density (pA/$\sqrt{\text{Hz}}$)
5532	5	0.7
5534A	3.5	0.4
LM4562	2.7	1.6
AD797	0.9	2
LT1028	0.85	1

www.raguvaran.puzl.com

Step 2 – replace the quad 5532 buffers with quad LM4562 buffers. Noise falls by only 0.6 dB, the output being −119.5 dBu, but at last we have a balanced stage that is quieter than the unbalanced stage, by a small but solid 0.5 dB.

One of the pre-eminent low-noise-with-low-source-resistance op-amps is the AD797 from Analog Devices, which has a remarkably low voltage noise of 0.9 nV/$\sqrt{\text{Hz}}$ (typical at 1 kHz) but it is a very expensive part, costing between 20 and 25 times more than a 5532 at the time of writing. The AD797 is a single op-amp, while the 5532 is a dual, so the cost per op-amp is actually 40–50 times greater, and more PCB area is required, but the potential improvement is so great we will overlook that.

Step 3 – we replace all four op-amps in the differential amplifiers with AD797s, putting the 5532s back into the input buffers in the hope that we might be able to save money somewhere. The noise output drops by a rather disappointing 0.4 dB, giving an output noise level of −119.9 dBu, quieter than the original 4 × 10 kΩ version by 15.1 dB.

Perhaps putting those 5532s back in the buffers was a mistake? Our fourth and final move in this game of electronic chess is to replace all the quad 5532 input buffers with dual (not quad) AD797 buffers. This requires another four AD797s (two per input) and is once more not a cheap strategy. We retain the four AD797s in the differential amplifiers. The noise drops by another 0.7 dB, yielding an output noise level of −120.6 dBu, quieter than the original 4 × 10 kΩ version by 15.8 dB, and quieter than the unbalanced stage by a satisfying 1.6 dB. You can do pretty much anything in electronics with a bit of thought and a bit of money.

You are probably wondering what happened to the LT1028 lurking at the bottom of Table 14.9. It is true that its voltage noise density is slightly better than that of the AD797, but there is a subtle snag. As described in Chapter 7, the LT1028 has bias-current cancellation circuitry that injects correlated noise currents into the two inputs. These will cancel if the impedances seen by the two inputs are the same, but in moving-magnet amplifier use the impedances differ radically and the LT1028 is not useful in this application. The input conditions here are more benign, but the extra complication is unwelcome and I have never used the LT1028 in audio work. In addition, it is a single op-amp with no dual version.

This is not of course the end of the road. The small noise improvement in the last step we made tells us that the differential amplifier array is still the dominant noise source, and further development would have to focus on this. A first step would be to see if the relatively high current noise of the AD797s is significant with respect to the surrounding resistor values. If so, we need to see if the resistor values can be reduced without degrading linearity at full output. We should also check the Johnson noise contribution of all those 820 Ω resistors; they are generating −123.5 dBu each at room temperature, but of course the partial cancellation effect applies to them as well.

All these noise results are also summarized in Table 14.10.

www.raguvaran.puzl.com

TABLE 14.10 A summary of the noise improvements made to the balanced input stage

Buffer type	Amplifier	Noise output (dBu)	Improvement on previous version (dB)	Improvement over 4 × 10 kΩ diff. amp (dB)	Noisier than unbalanced input by: (dB)
	5532 voltage-follower	−119.0			0.0 dB ref.
None	Standard diff. amp 10K 5532	−104.8	0	0.0 dB ref.	14.2
None	Single diff. amp 820R 5532	−111.7	6.9	6.9	7.3
Single 5532	Single diff. amp 820R 5532	−110.2	5.4	5.4	8.8
Single 5532	Dual diff. amp 820R 5532	−112.50	2.3	7.4	6.5
Single 5532	Quad diff. amp 820R 5532	−114.00	1.5	8.9	5.0
Dual 5532	Quad diff. amp 820R 5532	−116.20	2.2	11.1	2.8
Quad 5532	Quad diff. amp 820R 5532	−117.00	0.8	11.9	2.0
Quad 5532	Quad diff. amp 820R LM4562	−118.9	1.9	14.1	0.1
Quad LM4562	Quad diff. amp 820R LM4562	−119.5	0.6	14.7	−0.5
Quad 5532	Quad diff. amp 820R AD797	−119.9	0.4	15.1	−0.9
Dual AD797	Quad diff. amp 820R AD797	−120.6	0.7	15.8	−1.6

References

[1] N. Muncy, Noise susceptibility in analog and digital signal processing systems, JAES 3 (6) (June 1995) p. 447.

[2] T. Williams, EMC for Product Designers, Newnes (Butterworth-Heinemann) (1992) p. 176.

[3] T. Williams, EMC for Product Designers, Newnes (Butterworth-Heinemann) (1992) p. 173.

[4] S. Winder, What's the difference? Electronics World (November 1995) p. 989.

[5] J. Graeme, Amplifier Applications of Opamps, McGraw-Hill, 1999, p. 16.

[6] D. Birt, Electronically balanced analogue-line interfaces, Proceedings of the Institute of Acoustics Conference, Windermere, UK, November 1990.

[7] W. Jung (Ed.), Op Amp Applications Handbook, Newnes, Elsevier, 2006.

[8] P. Zwicky, Low-distortion audio amplifier circuit arrangement, US patent no. 4,567,443 (1986).

[9] http://www.signaltransfer.freeuk.com/

www.raguvaran.puzl.com

CHAPTER 15

Line Outputs

Unbalanced Outputs

There are only two electrical output terminals for an unbalanced output – signal and ground. However, the unbalanced output stage in Figure 15.1(a) is fitted with a three-pin XLR connector, to emphasize that it is always possible to connect the cold wire in a balanced cable to the ground at the output end and still get all the benefits of common-mode rejection if you have a balanced input. If a two-terminal connector is fitted, the link between the cold wire and ground has to be made inside the connector, as shown in Figure 14.2 in the chapter on line inputs.

The output amplifier in Figure 15.1(a) is configured as a unity-gain buffer, though in some cases it will be connected as a series-feedback amplifier to give gain. A non-polarized DC-blocking capacitor C1 is included: 100 μF gives a −3 dB point of 2.6 Hz with one of those notional 600 Ω loads. The op-amp is isolated from the line shunt capacitance by a resistor R2, in the range 47–100 Ω, to ensure HF stability, and this unbalances the hot and cold line impedances. A drain resistor R1 ensures that no charge can be left on the output side of C1; it is placed *before* R2, so it causes no attenuation. In this case the loss would only be 0.03 dB, but such errors can build up to an irritating level in a large system and it costs nothing to avoid them.

If the cold line is simply grounded as in Figure 15.1(a), then the presence of R2 degrades the CMRR of the interconnection to an uninspiring −43 dB even if the balanced input at the other end of the cable has infinite CMRR in itself and perfectly matched 10 kΩ input impedances.

To fix this problem, Figure 15.1(b) shows what is called an impedance-balanced output. There are now three physical terminals: hot, cold, and ground. The cold terminal is neither an input nor an output, but a resistive termination R3 with the same resistance as the hot terminal output impedance R2. If an unbalanced input is being driven, this cold terminal can be either shorted to ground locally or left open-circuit. The use of the word 'balanced' is perhaps unfortunate as when taken together with an XLR output connector it implies a true balanced output with anti-phase outputs, which is *not* what you are getting. The impedance-balanced approach is not particularly cost-effective, as it requires significant extra money to be spent on an XLR

Small Signal Audio Design; ISBN: 9780240521770
Copyright © 2010 Elsevier Ltd; All rights of reproduction, in any form, reserved.

www.raguvaran.puzl.com

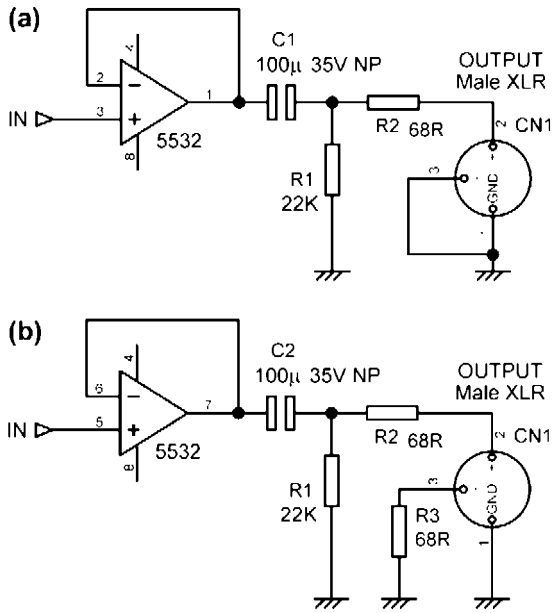


Figure 15.1: Unbalanced outputs. (a) Simple output. (b) Impedance-balanced output for improved CMRR when driving balanced inputs

connector. Adding an op-amp inverter to make it a proper balanced output costs little more, especially if there happens to be a spare op-amp half available, and it sounds much better in the specification.

Zero-Impedance Outputs

Both the unbalanced outputs shown in Figure 15.1 have series output resistors to ensure stability when driving cable capacitance. This increases the output impedance and can lead to increased crosstalk in some situations. One such scenario that was fixed by the use of a so-called 'zero-impedance' output is described in Chapter 17 in the section on mixer insert points.

Figure 15.2(a) shows how the same technique is applied to an unbalanced output stage with 10 dB of gain. Feedback at audio frequencies is taken from outside isolating resistor R3 via R2, while the HF feedback is taken from inside R3 via C2 so it is not affected by load capacitance and stability is unimpaired. Using a 5532 op-amp, the output impedance is reduced from 68 to 0.24 Ω at 1 kHz – a dramatic reduction that would reduce capacitative crosstalk by 49 dB. Output impedance increases to 2.4 Ω at 10 kHz and 4.8 Ω at 20 kHz as op-amp open-loop gain falls with frequency. The impedance-balancing resistor on the cold pin has been replaced by a link to match the near-zero output impedance at the hot pin.

www.raguvaran.puzl.com

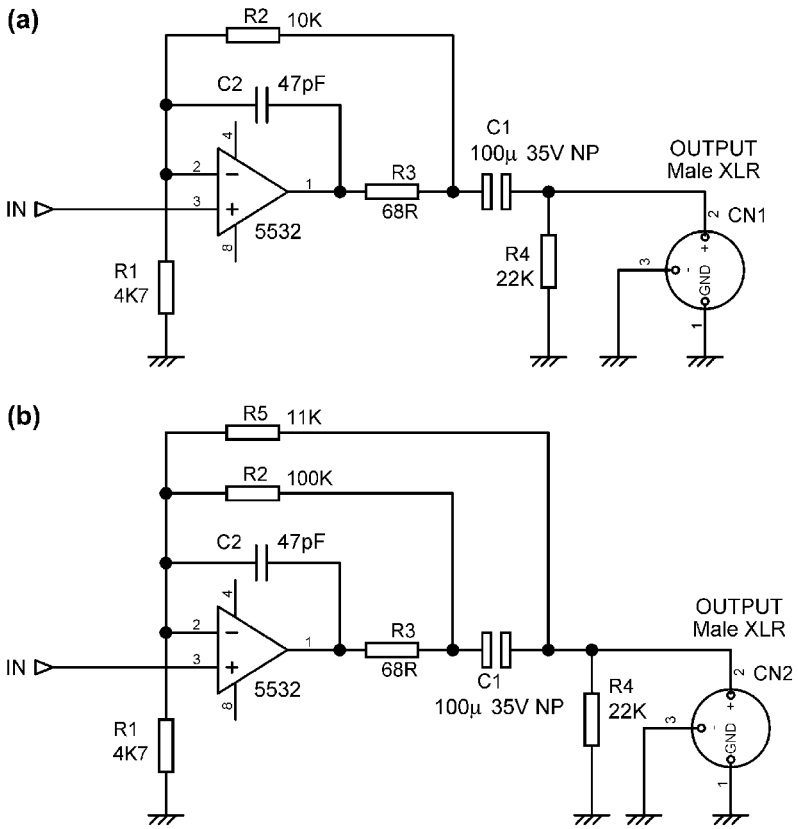


Figure 15.2: (a) Zero-impedance output. (b) Zero-impedance output with NFB around output capacitor

Figure 15.2(b) shows a refinement of this scheme with three feedback paths. In Chapter 2 we saw how electrolytic coupling capacitors can introduce distortion even if the time-constant is long enough to give a flat LF response. In Figure 15.2(b) most of the feedback is now taken from outside C1, via R5, so it can correct capacitor distortion. The DC feedback goes via R2, now much higher in value, and the HF feedback goes through C2 as before to maintain stability with capacitative loads. R2 and R5 in parallel come to 10 kΩ so the gain is the same. Any circuit with separate DC and AC feedback paths must be checked carefully for frequency response irregularities, which may happen well below 10 Hz.

Ground-Canceling Outputs

This technique, also called a ground-compensated output, appeared in the early 1980s in mixing consoles. It allows ground voltages to be canceled out even if the receiving equipment has an unbalanced input; it prevents any possibility of creating a phase error by

www.raguvaran.puzl.com

mis-wiring, and it costs virtually nothing except for the provision of a three-pin output connector.

Ground canceling (GC) separates the wanted signal from the unwanted ground voltage by addition at the output end of the link, rather than by subtraction at the input end. If the receiving equipment ground differs in voltage from the sending ground, then this difference is added to the output signal so that the signal reaching the receiving equipment has the same ground voltage superimposed upon it. Input and ground therefore move together and the ground voltage has no effect, subject to the usual effects of component tolerances. The connecting lead is differently wired from the more common unbalanced-out balanced-in situation, as now the cold line is to be joined to ground at the *input* or receiving end.

An inverting unity-gain ground-cancel output stage is shown in Figure 15.3(a). The cold pin of the output socket is now an input, and has a unity-gain path summing into the main signal going to the hot output pin to add the ground voltage. This path R3, R4 has a very low input impedance equal to the hot terminal output impedance so if it *is* used with a balanced input, the line impedances will be balanced, and the combination will still work effectively. The 6 dB of attenuation in the R3–R4 divider is undone by the gain of 2 set by R5, R6. It is unfamiliar to most people to have the cold pin of an output socket as a low impedance input, and its very low input impedance minimizes the problems caused by mis-wiring. Shorting it locally to ground merely converts the output to a standard unbalanced type. On the other hand, if the cold input is left unconnected then there will be a negligible increase in noise due to the very low input resistance of R3.

This is the most economical GC output, and is very useful to follow an inverting summing amplifier as it corrects the phase. Obviously the phase inversion is not always convenient, and Figure 15.3(b) shows a non-inverting GC output stage with a gain of 6.6 dB. R5 and R6 set up a gain of 9.9 dB for the amplifier, but the overall gain is reduced by 3.3 dB by attenuator R3, R4. The cold line is now terminated by R7, and any signal coming in via the cold pin is attenuated by R3, R4 and summed at unity gain with the input signal. The stage must be fed from a very low impedance such as an op-amp output to work properly. There is a slight compromise on noise performance here because attenuation is followed by amplification.

Ground-canceling outputs are an economical way of making ground loops innocuous when there is no balanced input, and it is rather surprising they are not more popular; perhaps people find the notion of an input pin on an output connector unsettling. In particular, GC outputs would appear to offer the possibility of a quieter interconnection than the standard balanced interconnection because a relatively noisy balanced input is not required (see Chapter 14). Ground-canceling outputs can also be made zero-impedance using the techniques described earlier.

www.raguvaran.puzl.com

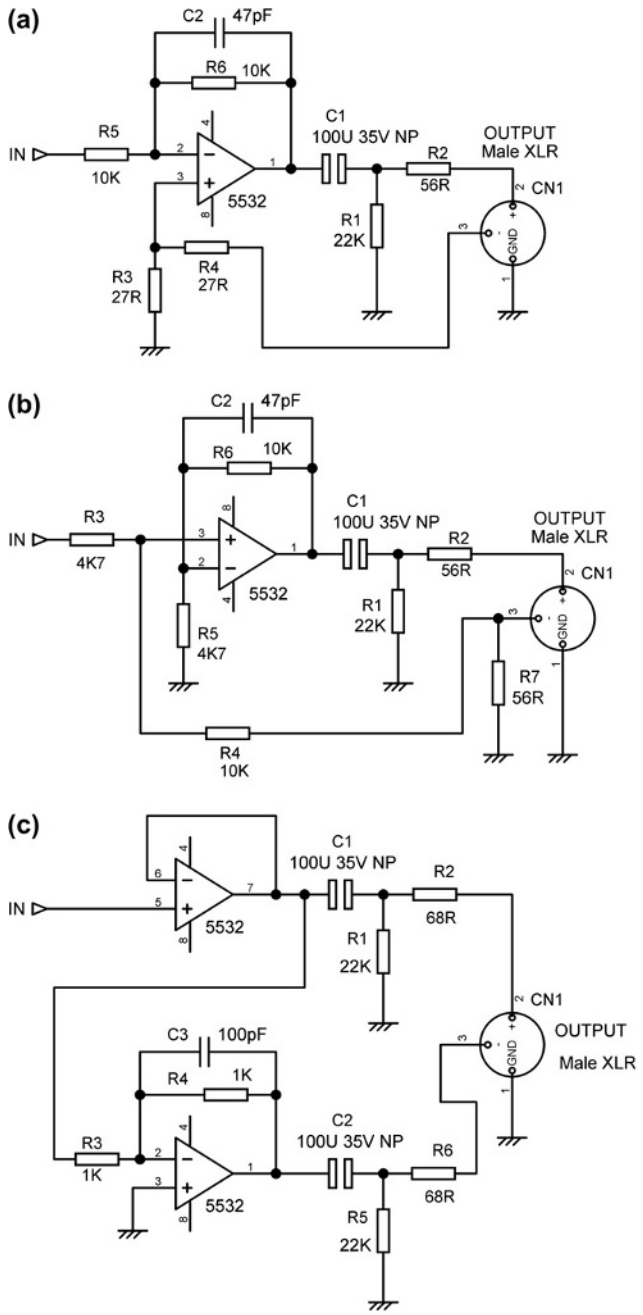


Figure 15.3: (a) Inverting ground-canceling output. (b) Non-inverting ground-canceling output. (c) A true balanced output

www.raguvaran.puzl.com

Balanced Outputs

Figure 15.3(c) shows a balanced output, where the cold terminal carries the same signal as the hot terminal but phase-inverted. This can be arranged simply by using an op-amp stage to invert the normal in-phase output. The resistors R3, R4 around the inverter should be as low in value as possible to minimize Johnson noise, because this stage is working at a noise gain of 2, but bear in mind that R3 is effectively grounded at one end and its loading, as well as the external load, must be driven by the first op-amp. A unity-gain follower is shown for the first amplifier, but this can be any other shunt or series-feedback stage as convenient. The inverting output if not required can be ignored; it must not be grounded, because the inverting op-amp will then spend most of its time clipping in current limiting, probably injecting unpleasant distortion into the grounding system. Both hot and cold outputs must have the same output impedances (R2, R6) to keep the line impedances balanced and the interconnection CMRR maximized.

A balanced output has the advantage that the total signal level on the line is increased by 6 dB, which if correctly handled can improve the signal-to-noise ratio. It is also less likely to crosstalk to other lines even if they are unbalanced, as the currents injected via the stray capacitance from each line will tend to cancel; how well this works depends on the physical layout of the conductors. All balanced outputs give the facility of correcting phase errors by swapping hot and cold outputs. This is, however, a two-edged sword, because it is probably how the phase got wrong in the first place.

There is no need to worry about the exact symmetry of level for the two output signals; ordinary tolerance resistors are fine. Such gain errors only affect the signal-handling capacity of the interconnection by a small amount. This simple form of balanced output is the norm in hi-fi balanced interconnection, but is less common in professional audio, where the quasi-floating output described below gives more flexibility.

Quasi-Floating Outputs

The purely electronic output stage in Figure 15.4(a) emulates a floating transformer winding; if both hot and cold outputs are driving signal lines, then the outputs are balanced, as if a center-tapped output transformer were being used, though clearly the output is not galvanically isolated from ground. If, however, the cold output is grounded, the hot output doubles in amplitude so the total level hot-to-cold is unchanged. This condition is detected by the current-sensing feedback taken from the outside of the 75 Ω resistor R10, and the current driven into the shorted cold output is automatically reduced to a low level that will not cause problems.

Similarly, if the hot output is grounded, the cold output doubles in amplitude and remains out of phase; the total hot–cold signal level is once more unchanged. This system has the

www.raguvaran.puzl.com

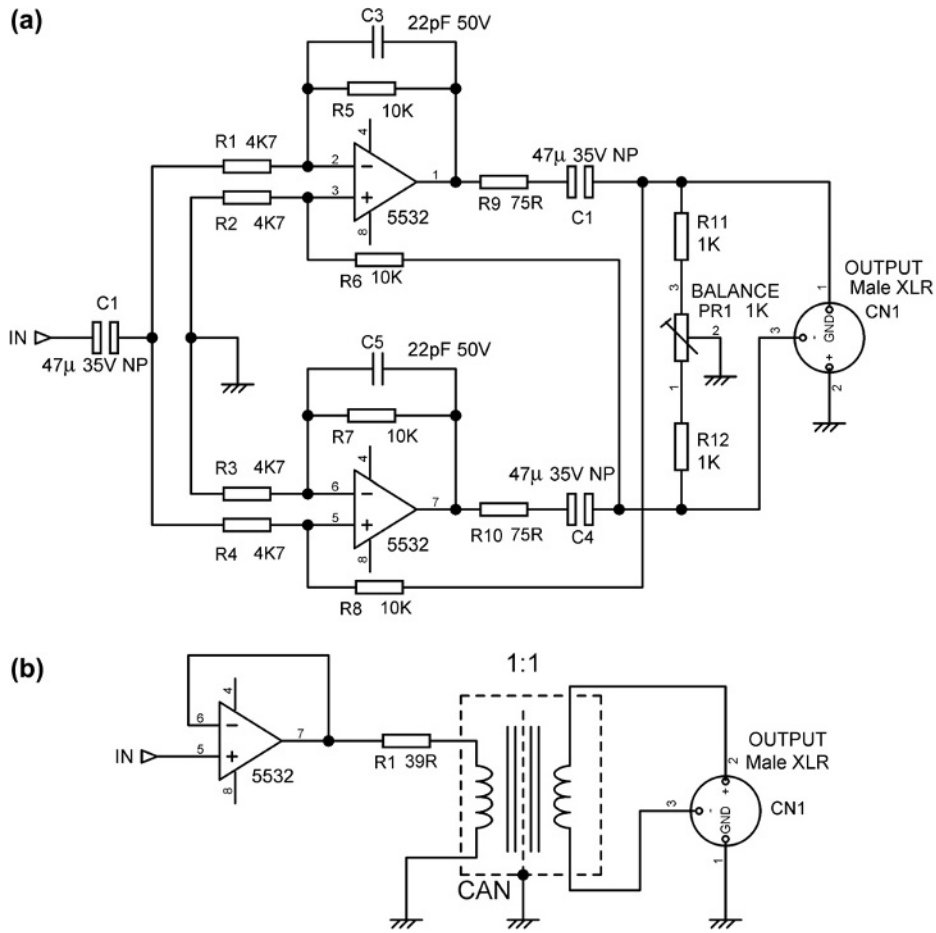


Figure 15.4: (a) Quasi-floating balanced output. (b) Transformer balanced output

advantage that it can give the same level into either a balanced or unbalanced input, given an appropriate connector at the input end; 6 dB of headroom are, however, lost when the output is used in unbalanced mode. It is most useful in recording studios where various bits of equipment may be temporarily connected; it is of less value in a PA system with a fixed equipment line-up.

When an unbalanced output is being driven, the quasi-floating output can be wired to work as a ground-canceling connection, with rejection of ground noise no less effective than the true balanced mode. This requires the cold output to be grounded at the remote (input) end of the cable. Under adverse conditions this might cause HF instability, but in general the approach is sound. If you are using exceptionally long cable, then it is wise to check that all is well.

If the cold output is grounded locally, i.e. at the sending end of the cable, then it works as a simple unbalanced output, with no noise rejection. When a quasi-floating output stage is used

www.raguvaran.puzl.com

unbalanced, the cold leg *must* be grounded, or common-mode noise will degrade the noise floor by at least 10 dB, and there may be other problems with increased distortion.

Quasi-floating outputs use a rather subtle mixture of positive and negative feedback of current and voltage. This performs the required function quite well, but a serious drawback is that it accentuates the effect of resistor tolerances, and so a preset resistor is normally required to set the outputs for equal amplitude; the usual arrangement is shown in Figure 15.4(a). If it is not correctly adjusted one side of the output will clip before the other and reduce the total output headroom. This is a set-and-forget adjustment unless it becomes necessary to change any of the resistors in the circuit.

Transformer Balanced Outputs

If true galvanic isolation between equipment grounds is required, this can only be achieved with a transformer. The technique is not that common, because of the cost, weight, and performance problems of transformers, but is sometimes found in touring PA systems (for example, in the mic-splitter box on the stage) and broadcast environments where huge RF field strengths are encountered. It is also used for test-gear outputs to prevent ground-loop formation. A basic transformer balanced output is shown in Figure 15.4(b).

Transformers have well-known problems with linearity at low frequencies. This is because the voltage induced into the secondary winding depends on the rate of change of the magnetic field in the core, and so the lower the frequency, the greater the change in magnitude must be for transformer action. The current drawn by the primary winding to establish this field is non-linear, because of the well-known non-linearity of iron cores. If the primary had zero resistance, and was fed from a zero source impedance, as much distorted current as was needed would be drawn and no one would ever know there was a problem. But… there is always some primary resistance, and this alters the primary current drawn so that distortion is introduced into the magnetic field established, and so into the secondary output voltage. Sometimes there is also series resistance such as R1 deliberately inserted into the primary circuit, presumably with the intention of avoiding HF in stability provoked by transformer winding and line capacitances; this makes the LF distortion problem worse, and a better means of isolation is a low-value inductor of say 4 μH in parallel with a low-value damping resistor of around 47 Ω.

The LF distortion can be reduced by applying negative feedback via a tertiary transformer winding. Another cure is to cancel out the transformer primary resistance by an electronically generated negative resistance; comprehensive details on this approach can be found in Bruce Hofer's patent, which covers the transformer output of the Audio Precision System 1 [1]. DC flowing through the primary winding is also bad for linearity, and if necessary should be stopped by a DC-blocking capacitor.

www.raguvaran.puzl.com

Driving Heavy Loads

The output stages described above assume that the load to be driven will not be less than 600 Ω and can be handled by a single 5532 op-amp section. If you want to keep optimal linearity then the lower limit is perhaps 700–800 Ω. There are many reasons to drive lower load resistance, such as headphone outputs, or driving special low-resistance circuitry designed to minimize noise.

Op-amps do exist with a greater load-driving capability – for example, the NJM4556A from JRC is capable of driving 150 Ω, but this device seems to have achieved little market penetration, possibly because its linearity even into light loads is distinctly inferior to that of the 5532. In 2009 none of the usual distributors is carrying it, and it is not clear if it is still in production.

In some circumstances paralleled 5532 stages are the simplest answer. In Figure 15.5 the upper trace shows a single 5532 section attempting to drive 5 Vrms into 220 Ω. The distortion is very high for a 5532, and it is clearly running out of current capability. Adding a second section in parallel with the outputs coupled by 10 Ω resistors, as shown in Figure 1.10 in Chapter 1, drops the THD back to almost nothing. The gains of the paralleled voltage-follower stages are very closely equal, so the small series output resistors can allow for minor gain mismatches and prevent currents passing from op-amp to op-amp.

This can be a very simple and cost-effective way to solve the problem of driving medium loads, and it can be extended by connecting more 5532 sections in parallel, apparently without limit. Figure 15.6 shows four parallel 5532 sections driving 100 Ω to 5 Vrms with very little distortion. An array of 12 parallel 5532 sections can drive a 25 Ω load with 5 Vrms very

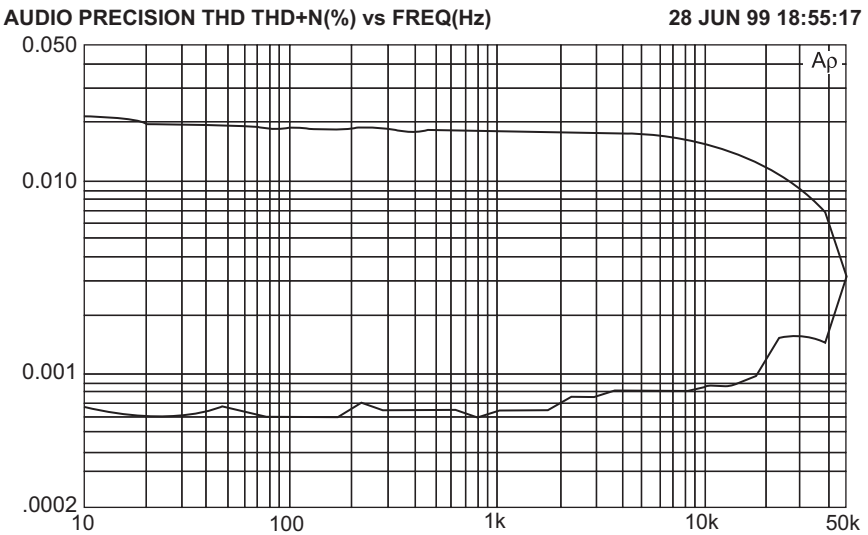


Figure 15.5 One and two 5532 sections attempting to drive 5 Vrms into 220 Ω

www.raguveran.puzl.com

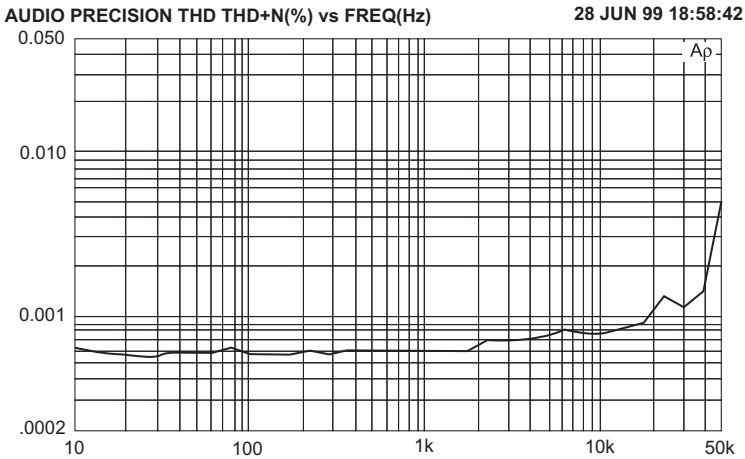


Figure 15.6: Four parallel 5532 sections driving 100 Ω with 5 Vrms. The THD plot is almost identical to that of two 5532 sections driving 220 Ω

effectively, and I have experimentally driven 8 Ω speakers in the same way, at extremely low distortion. That does, however, require rather a lot of 5532s.

As described in Chapter 1, paralleled op-amps not only increase drive capability but also reduce noise and allow a lower effective output impedance when driving capacitative loads.

For loads beyond the capabilities of the two halves of a 5532 in parallel, it may be more economical to adopt a hybrid circuit. (These are 'hybrid' circuits in that they combine IC op-amps with discrete transistors; the name does not refer to thick-film construction or anything like that.) The usual procedure is to add a Class-B output stage after the op-amp. A Class-A output stage is perfectly feasible if ultimate quality is required, though the power consumption is naturally much higher.

The arrangement in Figure 15.7(a) is economical and dependable, and I have used it to drive headphones in low-cost mixing consoles. The output stage is the simplest possible version of Class-B, but it really works quite well, though its distortion characteristics, as seen in Figure 15.8, are not comparable with a 5532's. The output devices are TO-92 small-signal types with high beta, and this is crucial to an acceptable distortion performance; the bias is fixed by D1, D2. Gain is 3.2 times. Inserting extra stages into op-amp feedback loops must be done with care, to avoid adding extra phase-shifts that may cause HF instability; here there are no problems, and no extra stabilizing components are required. The output is AC coupled by C2 so that DC-offset protection is not required, and output resistor R7 takes care of short-circuit protection, and also does something to cater for the different levels of drive required by 600 and 30 Ω headphones; it seems to be generally accepted that headphones do not require a very low output impedance.

www.raguvaran.puzl.com

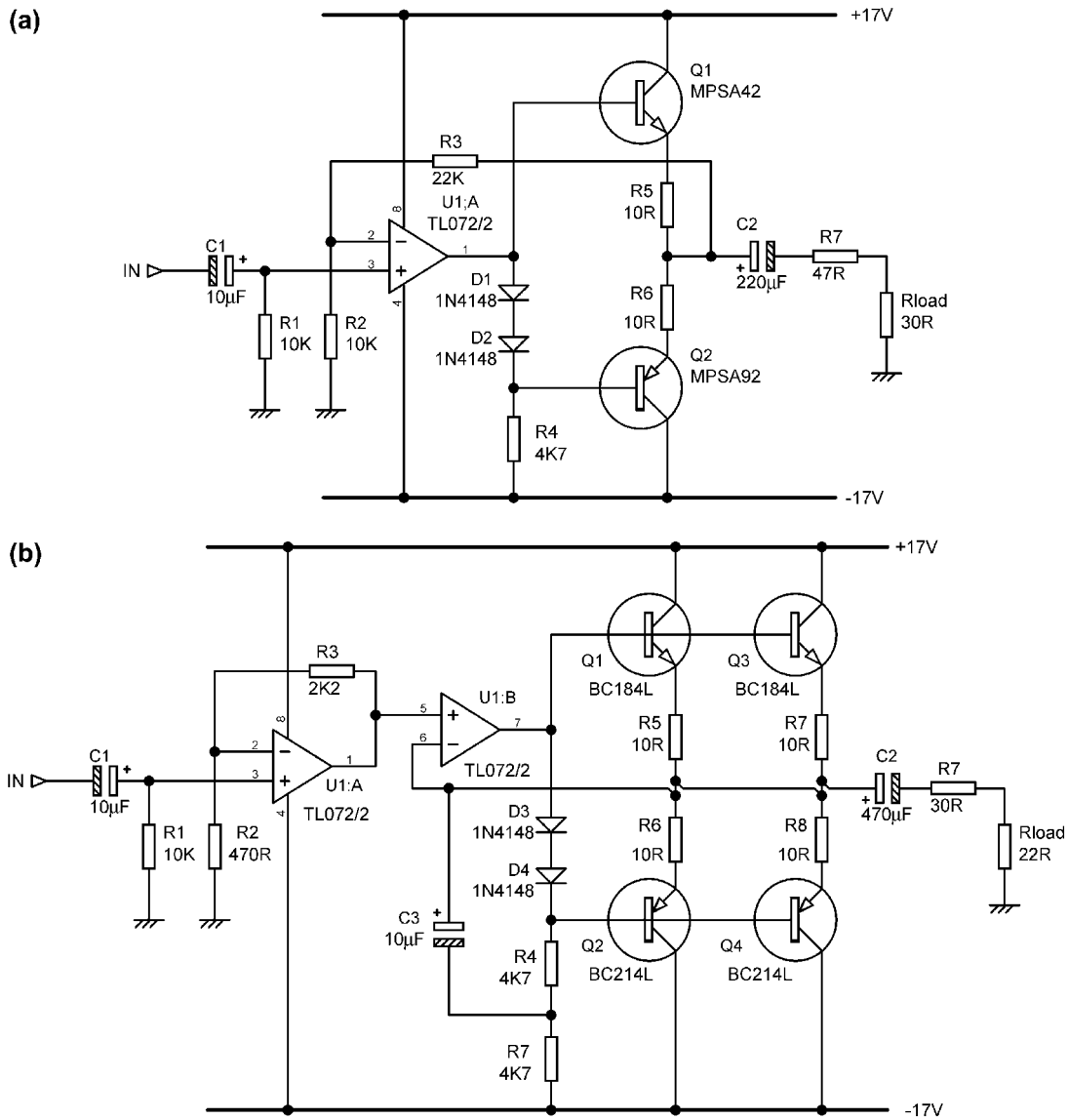


Figure 15.7: Two headphone amplifiers that combine a TL072 with a discrete Class-B output stage. (a) The basic version. (b) A superior variant giving greater output and lower distortion

The arrangement in Figure 15.7(b) is a deluxe version with lower distortion and greater output into lower impedances. I have deployed it in high-end mixing consoles and in preamplifiers with a headphone output. The gain is higher at 5.7 times, and is provided by U1:A, allowing U1:B to work at unity gain so it has maximal negative feedback for

www.raguvaran.puzl.com

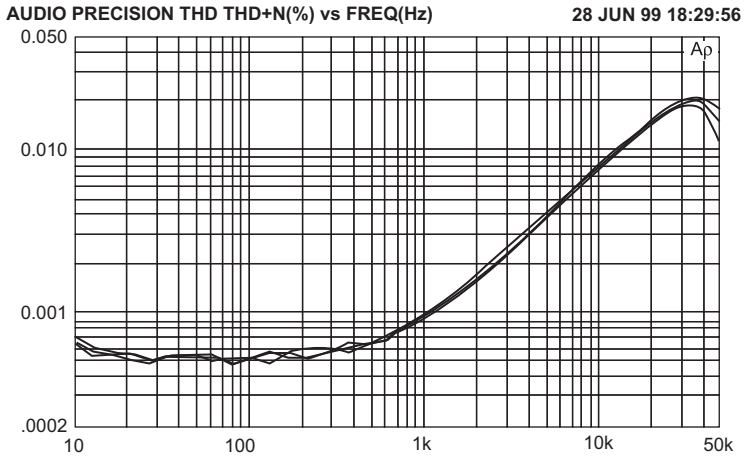


Figure 15.8: THD against load for TL072 with discrete Class-B output stage. No-load, 470R, and 220R loading. Gain 3×, output level 7.75 Vrms (before output series resistor)

correcting errors in the output stage. Bootstrapping is added to R4, R7 to maintain a more constant current through bias diodes D1, D2. The distortion is reduced to 0.002% at 10 kHz and 0.004% at 20 kHz with 6 Vrms at the amplifier output when driving a 30 Ω load.

Reference

[1] B. Hofer, Low-distortion transformer-coupled circuit, US patent no. 4,614,914 (1986).

www.raguvaran.puzl.com

CHAPTER 16

Signal Switching

The switching and routing of analog signals is a fundamental part of signal processing, but not one that is easily implemented if accuracy and precision are required. This chapter focuses on audio applications, but the basic parameters such as isolation and linearity are equally relevant in many fields.

Mechanical Switches

A mechanical switch normally makes a solid unequivocal connection when it is closed, and it is as 'on' as the resistance of its contacts and connections allows; these are small fractions of an ohm and are unlikely to cause trouble in small-signal audio design. Switches are, however, in general terms a good deal less 'off'. The insulation resistance may be measured in peta-ohms, but what does the damage is the inevitable capacitance between contacts. This is usually small in pF, but quite large enough to dominate the degree of offness obtainable at high audio frequencies. Its effects naturally depend on the impedance at the 'receiving' side of the switch. For all the tests discussed here this was 10 kΩ.

Using an ALPS SPUN type push-switch, at 10 kHz the offness is only −66 dB, and grounding the unused side of the switch only improves offness by about 2 dB. A graph of the result can be seen in Chapter 17 (Figure 17.18); the offness naturally degrades by 6 dB/octave. Switch capacitance is an important issue in designing mixer routing systems.

In another test a miniature three-way slide switch gave −70 dB at 10 kHz. Once again, grounding the unused contact at the end of the switch only gave a 4 dB improvement, and it is wise to assume that, in general, grounding unused switch sections will not help much.

Switch intercontact capacitance is quite easy to determine: measure the offness, i.e. the loss of the RC circuit, and since R is known C can be calculated easily. Once it is known for a given switch construction, it is easy to calculate the offness for different loading resistances. Interestingly, the intercontact capacitance of switches seems to be relatively constant, even though they vary widely in size and construction. This seems to be because the smaller switches have smaller contacts, with a smaller area, but on the other hand they are closer together.

Small Signal Audio Design; ISBN: 9780240521770
Copyright © 2010 Elsevier Ltd; All rights of reproduction, in any form, reserved.

www.raguvaran.puzl.com

Input-Select Switching

Some time ago, Morgan Jones [1] raised the excellent point of crosstalk in the input-select switching of preamplifiers. If the source impedance is significant then this may be a serious problem. While I agree that his use of a rotary switch with twice the required number of positions and grounding alternate contacts is slightly superior to the conventional use of rotary switches, measuring a popular Lorlin switch type showed the improvement to be only 5 dB. I am also unhappy with all those redundant 'Mute' positions between input selections, so when I design a preamp I normally choose interlocked push-switches rather than a rotary switch. A four-changeover format can then be used to reduce crosstalk.

The problem with conventional input-select systems like that shown in Figure 16.1(a) is that the various input tracks necessarily come into close proximity, with significant crosstalk through capacitance C_{stray} to the common side of the switch, i.e. from A to B. Using two changeovers per input side (i.e. four for stereo) allows the intermediate connection B–C to be

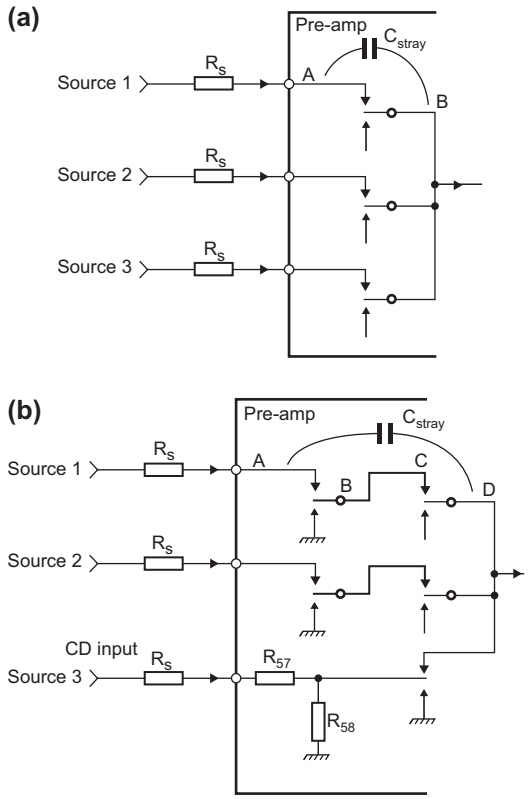


Figure 16.1: (a) Two-changeover select switches give poor signal rejection due to switch capacitance C_{stray}. (b) Using four-changeover switches improves offness by 21 dB at 10 kHz. Note CD input attenuator

www.raguvaran.puzl.com

grounded by the NC contact of the first switch section, and keeps the 'hot' input A much further away from the common input line D, as shown in Figure 16.1(b). C_{STRAY} is now much smaller.

The crosstalk data in Table 16.1 was gathered at 10 kHz, with 10 kΩ loading resistances.

TABLE 16.1 Offness of various switch configurations

	10 kHz
Simple rotary switch	−71 dB
Rotary with alternate contacts grounded	−76 dB
Two-changeover switch	−74 dB
Four-changeover switch	−95 dB

The emphasis here is on minimizing crosstalk between different sources carrying different signals, as interchannel (L–R) crosstalk is benign by comparison. Interchannel isolation is limited by the placement of Left and Right channels on the same switch, with the contact rows parallel, and limits L–R isolation to −66 dB at 10 kHz with a high 10 kΩ source impedance. Actual source impedances are likely to be lower, with both intersource and interchannel crosstalk proportionally reduced, so a more probable 1 kΩ source gives 115 dB of intersource rejection at 10 kHz for the four-changeover configuration.

The third input of Figure 16.1(b) has a resistive attenuator intended to bring CD outputs down to the same level as other sources. In this case intersource crosstalk can be improved simply by back-grounding the attenuator output when it is not in use, so only a two-pole switch is required for good isolation of this input.

The Virtual Contact

What do you do if you need a changeover switch – to select one of two signal sources – but only have a make contact? Here is a technique that can be a lifesaver when you have screwed up on ordering a switch with a long lead-time, or if you live in a Dilbertian world of last-minute spec changes.

Figure 16.2 demonstrates the principle. With the switch S open, source A goes through voltage-follower A and R_{feed} to the output voltage-follower. With the switch closed, the much lower impedance output of voltage-follower B takes over and the contribution from A is now negligible. To give good rejection of A, the output impedance of follower B must be much, much lower than the value of R_{feed}, so an op-amp output must be used directly. Figure 16.3 shows how good the rejection of A can be using 5532 op-amps as the voltage-followers.

At first this technique looks a bit op-amp-intensive. However, there is often no need to use dedicated voltage-followers if a similar low-impedance feed is available from a previous stage

www.raguvaran.puzl.com

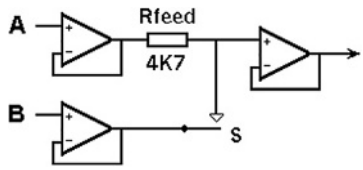


Figure 16.2: The virtual contact concept. When the switch S is closed the signal B overrides signal A

that uses an op-amp with a large amount of negative feedback. Likewise, the output voltage-follower may often be dispensed with if the following load is reasonably high.

There is also the rejection of B when the switch is open, to consider. The impedance of R_{feed} means there is the potential for capacitative crosstalk across the open switch contacts. The amount depends on the value of R_{feed} and on switch construction.

If the offness of B is more important than the offness of A, then R_{feed} should be a lower value, to minimize the effects of the capacitance. Do not make R_{feed} too low as A drives through it into effectively a short-circuit when B is selected.

If the offness of A is more important, R_{feed} should be higher to increase its ratio to the output impedance of B; be aware that making it too high may introduce excessive Johnson noise.

The rejection of A shown in Figure 16.3 worsens at high frequencies, as the dominant pole of op-amp B reduces its open-loop gain and the output impedance rises. The slopes are 6 dB/octave as usual.

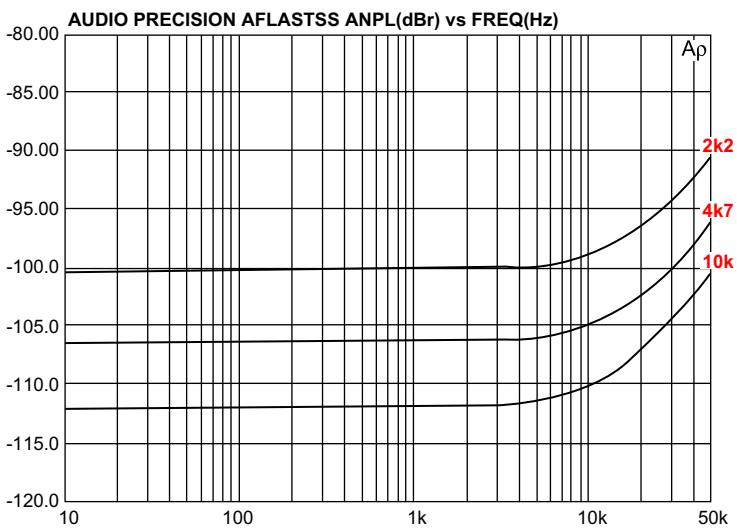


Figure 16.3: Rejection of signal A for 2k2, 4k7, and 10 kΩ R_{feed} resistors, using 5532 op-amps

www.raguvaran.puzl.com

This technique is particularly useful for switching between three sources with a center-off toggle switch.

Relay Switching

Any electronic switching technique must face comparison with relays, which are still very much with us. Relays give total galvanic isolation between control and signal, zero contact distortion, and in audio terms have virtually unlimited signal-handling capability. They introduce negligible series resistance and shunt leakage to ground is usually also negligible. Signal offness can be very good, but as with other kinds of switching this depends on intelligent usage. There will always be capacitance between open contacts, and if signal is allowed to crosstalk through this to nominally off circuitry, the 'offness' will be no better than other kinds of switching.

Obviously relays have their disadvantages. They are relatively big, expensive, and not always as reliable as more than a hundred years of development should have made them. Their operating power is significant, though it can be reduced by circuitry that applies full voltage to pull in the relay and then a lower voltage to keep it closed. Some kinds of power relay can introduce disastrous distortion if used for switching audio because the signal passes through the magnetic soft-iron frame; however, such problems are likely to be confined to the output circuits of large power amplifiers. For small-signal switching the linearity of relay contacts can normally be regarded as perfect.

Electronic Switching

Electronic switching is usually implemented with complementary metal oxide semiconductor (CMOS) analog gates, of which the well-known 4016 is the most common example, and these are examined first. However, there are many special applications where discrete JFETs provide a better solution, so these are dealt with in the second part.

Switching with CMOS Analog Gates

CMOS analog gates, also known as transmission gates, are quite different from CMOS logic gates, though the underlying process technology is the same. Analog gates are bilateral, which means that either of the in/out leads can be the input or output; this is emphatically not true for logic gates. The 'analog' part of the name emphasizes that they are not restricted to fixed logic levels, but pass whatever signal they are given with low distortion. The 'low' there requires a bit of qualification, as will be seen later.

When the gate is switched on, the connection between the two pins is a resistance that passes current in each direction as usual, depending on the voltage between the two terminals. Analog gates have been around for a long time, and are in some ways the obvious method of electronic switching. They do, however, have significant drawbacks.

www.raguvaran.puzl.com

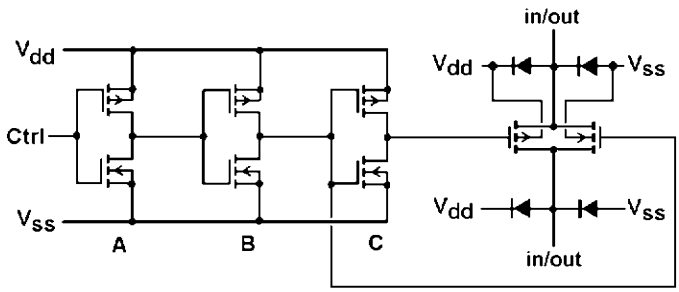


Figure 16.4: The internal circuitry of a 4000 series analog gate

Analog gates such as the 4016 are made up of two metal oxide semiconductor field-effect transistors (MOSFETs) of opposite polarity connected back to back. The internal structure of a 4016 analog gate is shown in Figure 16.4. The two transmission FETs with their protective diodes are shown on the right; on the left is the control circuitry. A and B are standard CMOS inverters whose only function is to sharpen up the rather soggy voltage levels that 4000 series CMOS logic sometimes provides. The output of B directly controls one FET, and inverter C develops the anti-phase control voltage for the FET of opposite polarity, which requires an inverted gate voltage to turn it on or off.

MOSFETS are of the enhancement type, requiring a voltage to be applied to the gate to turn them on (in contrast, JFETs work in depletion mode and require a gate voltage to turn them *off*), so as the channel approaches the gate voltage, the device turns off more. An analog gate with only one polarity of FET would be of little use because R_{on} would become very high at one extreme of the voltage range. This is why complementary FETs are used; as one polarity finds its gate voltage decreasing, turning it off, the other polarity has its gate voltage increasing, turning it more on. It would be nice if this process canceled out so the R_{on} was constant, but sadly it just doesn't work that way. Figure 16.5 shows how R_{on} varies with input voltage, and the peaky curve gives a strong hint that something is turning on as something else turns off.

Figure 16.5 also shows that R_{on} is lower and varies less when the higher supply voltage is used; since these are enhancement FETs the on-resistance decreases as the available control voltage increases. If you want the best linearity then always use the maximum rated supply voltage.

Since R_{on} is not very linear, the smaller its value the better. The 4016 R_{on} is specified as 115 Ω typical, 350 Ω max, over the range of input voltages and with a 15 V supply. The 4066 is a version of the 4016 with lower R_{on}, 60 Ω typical, 175 Ω max under the same conditions. This option can be very useful both in reducing distortion and improving offness, and in most cases there is no point in using the 4016. The performance figures given below assume the use of the 4066 except where stated.

www.raguvaran.puzl.com

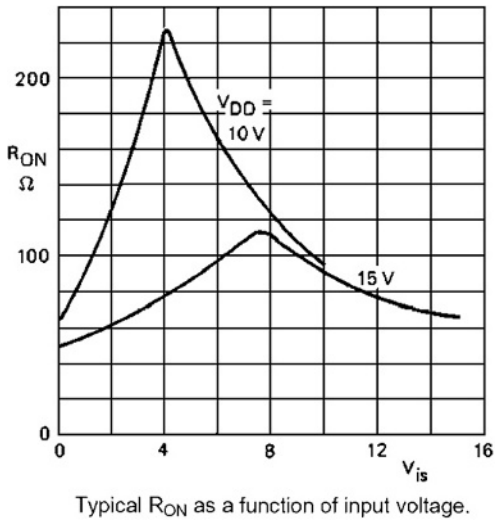


Figure 16.5: Typical variation of the gate series resistance R_{on}

CMOS Gates in Voltage Mode

Figure 16.6 shows the simplest and most obvious way of switching audio on and off with CMOS analog gates. This series configuration is in a sense the 'official' way of using them, the only snag being that it doesn't work very well.

Figure 16.7 shows the measured distortion performance of the simple series gate using the 4016 type. The distortion performance is a long way from brilliant, exceeding 0.1% just above 2 Vrms. These tests, like most in this section, display the results for a single sample of the semiconductor in question. Care has been taken to make these representative, but there will inevitably be some small variation in parameters like R_{on}. This may be greater when comparing the theoretically identical products of different manufacturers.

Replacing the 4016 gate with a 4066 gives a reliable improvement due to the lower R_{on}. THD at 2 Vrms (10 kΩ load) has dropped to a third of its previous level. There seems to be no downside to using 4066 gates instead of the more common and better-known 4016, and they are used exclusively from this point on, unless otherwise stated.

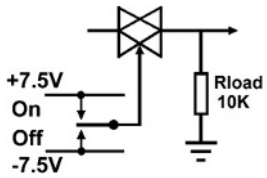


Figure 16.6: Voltage-mode series switching circuit using analog gate

www.raguvaran.puzl.com

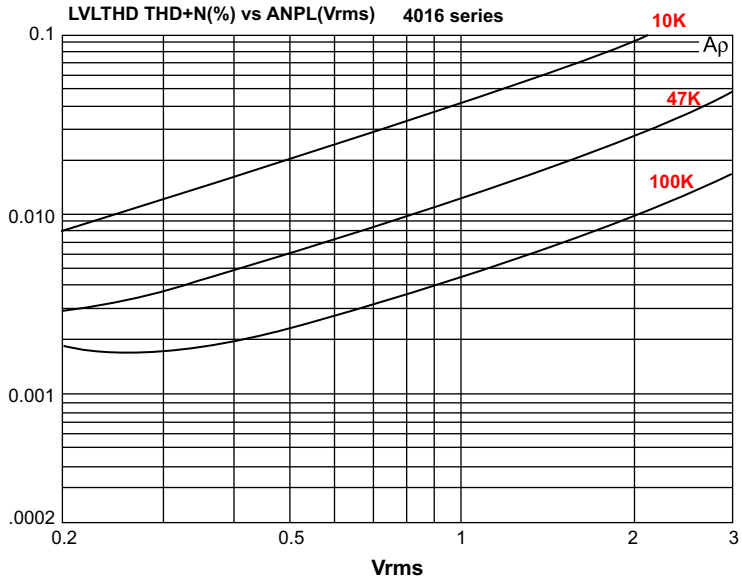


Figure 16.7: 4016 series gate THD versus level, with different load resistances

The distortion is fairly pure second harmonic, except at the highest signal levels where higher-order harmonics begin to intrude. This is shown in Figures 16.8 and 16.9 by the straight line plots beginning to bend upwards above 2 Vrms.

Analog gate distortion is flat with frequency as far as audio is concerned, so no plots of THD versus frequency are shown; they would just be a rather uninteresting set of horizontal lines.

This circuit gives poor offness when off, as shown by Figure 16.10. The offness is limited by the stray capacitance in the package feeding through into the relatively high load impedance. If this is 10 kΩ the offness is only −48 dB at 20 kHz, which would be quite inadequate for most applications. The load impedance could be reduced below 10 kΩ to improve offness – for example, 4k7 offers about a 7 dB improvement – but this degrades the distortion, which is already poor at 0.055% for 3 Vrms, to 0.10%.

Using 4066 gates instead of 4016s does not improve offness in this configuration. The internal capacitance that allows signals to leak past the gate seems to be the same for both types. The maximum signal level that can be passed through (or stopped) is limited by the CMOS supply rails and conduction of the protection diodes. While it would in some cases be possible to contrive a bootstrapped supply to remove this limitation, it is probably not a good route to head down.

Figure 16.11 shows a CMOS three-way switch. When analog gates are used as a multi-way switch, the offness problem is much reduced, because capacitative feedthrough of the

www.raguvaran.puzl.com

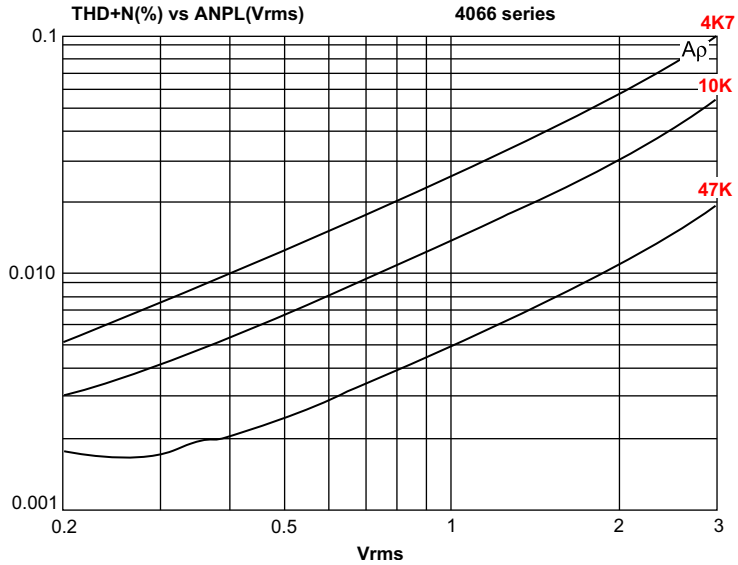


Figure 16.8: 4066 THD versus level, with different load resistances

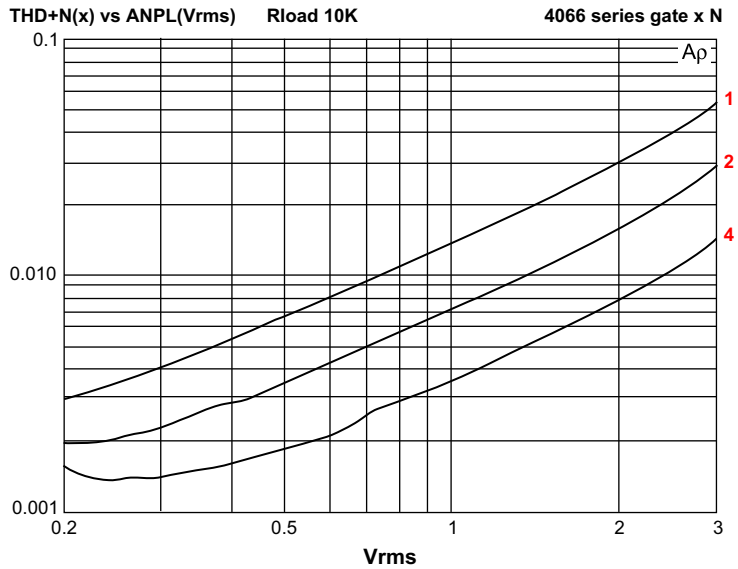


Figure 16.9: THD versus level, for different numbers of series 4066 gates connected in parallel

unwanted inputs is attenuated by the low R_{on} looking back into the (hopefully) low impedance of the active input, such as an op-amp output. If this is not the case then the crosstalk from nominally off inputs can be serious.

In this circuit the basic poor linearity is unchanged, but since the crosstalk problem is much less, there is often scope for increasing the load impedance to improve linearity. This makes R_{on}

www.raguvaran.puzl.com

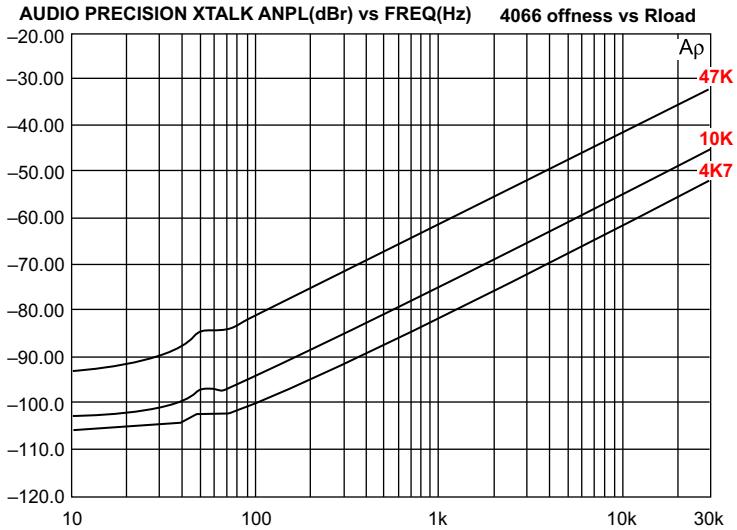


Figure 16.10: 4016 and 4066 offness versus load resistance; −48 dB at 20 kHz with a 10 kΩ load

a smaller proportion of the total resistance. The control voltages must be managed so that only one gate is on at a time, so there is no possibility of connecting two op-amp outputs together.

It may appear that if you are implementing a true changeover switch, which always has one input on, the resistor to ground is redundant, and just a cause of distortion. Omitting it is, however, very risky, because if all CMOS gates are off together even for an instant, there is no DC path to the op-amp input and it will register its displeasure by snapping its output to one of the rails. This does not sound nice.

Figure 16.12 shows the offness of a changeover system, for two types of FET-input op-amps. The offness is much improved to −87 dB at 20 kHz, an improvement of 40 dB over the simple series switch; at the high-frequency end, however, it still degrades at the same rate of

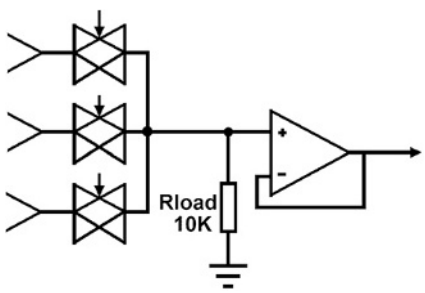


Figure 16.11: A one-pole, three-way switch made from analog gates

www.raguvaran.puzl.com

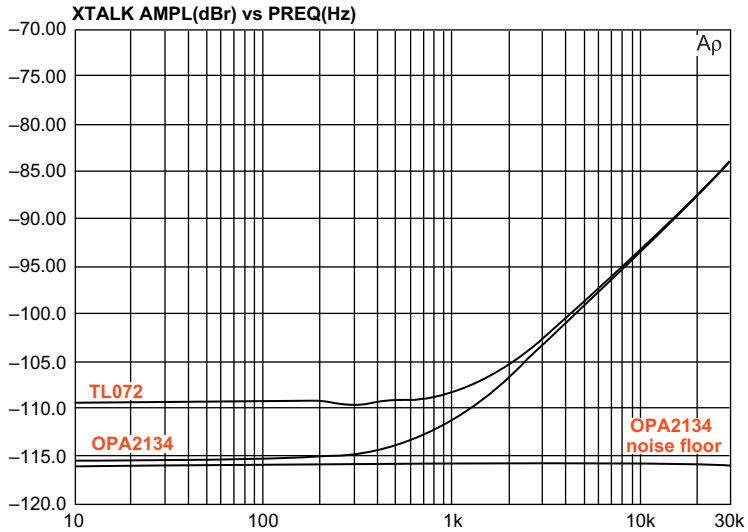


Figure 16.12: Voltage-mode changeover circuit offness for TL072 and OPA2134. R_{load} = 10k

6 dB/octave. It is well known that the output impedance of an op-amp with negative feedback increases with frequency at this rate, as the amount of internal gain falls, and this effect is an immediate suspect. However, there is actually no detectable signal on the op-amp output (as shown by the lowest trace) and it is also not very likely that two completely different op-amps would have exactly the same output impedance. I was prepared for a subtle effect, but the true explanation is that the falling offness is simply due to feedthrough via the internal capacitance of the analog gate.

It now remains to explain why the OPA2134 apparently gives better offness in the flat low-frequency region. In fact it does not; the flat parts of the trace represent the noise floor for that particular op-amp. The OPA2134 is a more sophisticated and quieter device than the TL072, and this is reflected in the lower noise floor.

There are two linearity problems. Firstly, the on-resistance itself is not totally linear. Secondly, and more serious, the on-resistance is modulated when the gates move up and down with respect to their fixed control voltages.

It will by now probably have occurred to most readers that an on/off switch with good offness can be made by making a changeover switch with one input grounded. This is quite true, but since much better distortion performance can be obtained by using the same approach in current mode, as explained below, I am not considering it further here.

Figure 16.13 shows a shunt muting circuit. This gives no distortion in the 'On' state because the signal is no longer going through the R_{on} of a gate. However, the offness is limited by the R_{on}, forming a potential divider with the series resistor R; the latter cannot be very high in

www.raguvaran.puzl.com

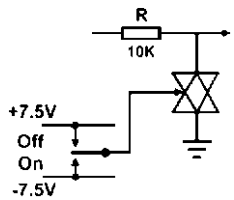


Figure 16.13: Voltage-mode shunt CMOS muting circuit

value or the circuit noise will be degraded. There is, however, the advantage that the offness plot is completely flat with frequency. Note that the On and Off states of the control voltage are now inverted.

Table 16.2 gives the measured results for the circuit, using the 4066. The offness can be improved by putting two or more of these gates in parallel, but since doubling the number *N* only gives 6 dB improvement, it is rarely useful to press this approach beyond four gates.

TABLE 16.2 Offness versus number of shunt 4066 analog gates used, with R = 10 KΩ

No. of gates	Offness (dB)
1	−37
2	−43
4	−49

CMOS Gates in Current Mode

Using these gates in current mode – usually by defining the current through the gate with an input resistor and dropping it into the virtual-earth input of a shunt-feedback amplifier – gives much superior linearity. It removes the modulation of channel resistance as the gate goes up and down with respect to its supply rails and, in its more sophisticated forms, can also remove the signal voltage limit and improve offness.

Figure 16.14 shows the simplest version of a current-mode on/off switch. An important design decision is the value of R_{in} and R_{nfb}, which are often equal to give unity gain. Too low

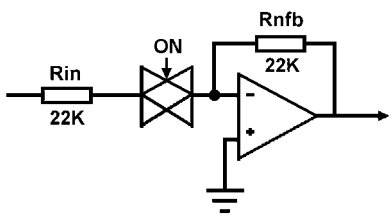


Figure 16.14: The simplest version of a current-mode on/off switch

www.raguvaran.puzl.com

TABLE 16.3 Distortion produced by a current-mode switch using 4016 gates, showing the gate contribution is small

	1 kHz	10 kHz	20 kHz
THD via 4016, +20 dBu	0.0025%	0.0039%	0.0048%
THD: 4016 shorted, +20 dBu	0.0020%	0.0036%	0.0047%
Offness (dB)	−68	−48	−42

a value increases the effect of the non-linear R_{on}, while too high a value degrades offness, as it makes the gate stray capacitance more significant, and also increases Johnson noise. In most cases 22 kΩ is a good compromise.

Table 16.3 gives the distortion for +20 dBu (7.75 Vrms) in/out, and shows that it is now very low compared with voltage-mode switchers working at much lower signal levels (compare the table data with Figures 16.8 and 16.9). The increase in THD at high frequencies is due to a contribution from the op-amp. However, the offness is pretty poor, and would not be acceptable for most applications. The problem is that with the gate off, the full signal voltage appears at the gate input and crosstalks to the summing node through the package's internal capacitance. In practical double-sided PCB layouts the inter-track capacitance can usually be kept very low by suitable layout, but the internal capacitance of the gate is inescapable.

In Figures 16.14 and 16.15, the CMOS gate is powered from a maximum of ±7.5 V. This means that in Figure 16.11, signal breakthrough begins at an input of 5.1 Vrms. This is much too low for op-amps running off their normal rail voltages, and several decibels of headroom are lost.

Figure 16.15 shows a partial cure for this. Resistor R_{in2} is added to attenuate the input signal when the CMOS gate is off, preventing breakthrough. There is no effect on signal gain when the gate is on, but the presence of R_{in2} does increase the noise gain of the stage.

As with all shunt-feedback stages, this circuit introduces a phase inversion, which is sometimes convenient but usually not.

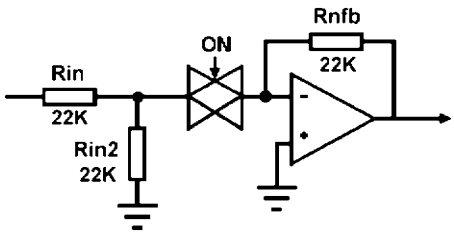


Figure 16.15: Current-mode switch circuit with breakthrough prevention resistor R_{in2}

www.raguvaran.puzl.com

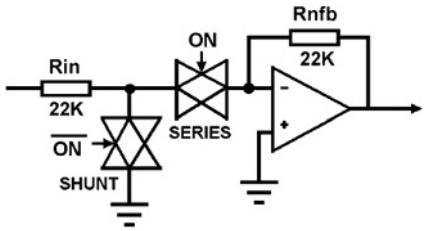


Figure 16.16: A series-shunt current-mode switch

Series-Shunt Current Mode

We now extravagantly use two 4016 CMOS gates, as shown in Figure 16.16. When the switch is on, the series gate passes the signal through as before; the shunt gate is off and has no effect. When the switch is off the series gate is off and the shunt gate is on, sending almost all the signal at A to ground so that the remaining voltage is very small. The exact value depends on the 4016 specimen and its R_{on} value, but is about 42 dB below the input voltage. This deals with the offness (by greatly reducing the signal that can crosstalk through the gate internal capacitance) and also increases the headroom by several decibels, as there is now effectively no voltage signal to break through when it exceeds the rails of the series gate.

Two anti-phase control signals are now required. If you have a spare analog gate it can generate the inverted control signal, as shown in Figure 16.17.

The distortion generated by this circuit can be usefully reduced by using two gates in parallel for the series switching, as in Table 16.4; this gate doubling reduces the ratio of the variable R_{on} to the fixed series resistor and so improves the linearity. Using two in parallel is sufficient to render the distortion negligible. (The higher distortion figures at 10 and 20 kHz are due to distortion generated by the TL072 op-amp used in the measurements.)

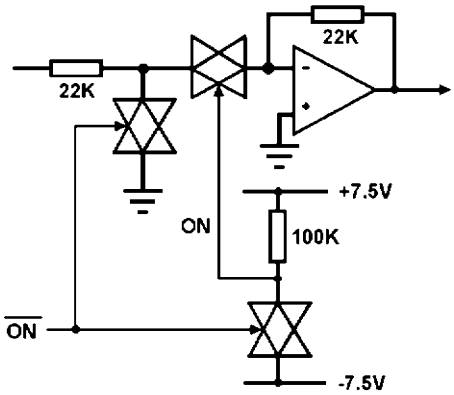


Figure 16.17: Generating the control signals with a spare analog gate

www.raguvaran.puzl.com

TABLE 16.4 Distortion levels with series-shunt switching

	1 kHz	10 kHz	20 kHz
THD via 4016 × 1, +20 dBu	0.0016%	0.0026%	0.0035%
THD via 4016 × 2, +20 dBu	0.0013%	0.0021%	0.0034%
THD 4016 shorted, +20 dBu	0.0013%	0.0021%	0.0034%
Offness 4016 × 1 (dB)	−109	−91	−86
Offness 4016 × 1, J111 (dB)	<−116	−108	−102

As before the input and output levels are +20 dBu, well above the nominal signal levels expected in op-amp circuitry; measurements taken at more realistic levels would show only noise.

Discrete FETs have lower R_{on} than analog gates. If a J111 JFET is used as the shunt switching element the residual signal at A is further reduced, to about 60 dB below the input level, with a consequent improvement in offness, demonstrated by the bottom row in Table 16.4. This could also be accomplished by using two or more CMOS gates for the shunt switching.

Control Voltage Feedthrough in CMOS Gates

When an analog gate changes state, some energy from the control voltage passes into the audio path via the gate-channel capacitance of the switching FETs, through internal package capacitances, and through any stray capacitance designed into the PCB. Since the control voltages of analog gates move snappily, due to the internal inverters, this typically puts a click rather than a thump into the audio. Attempts to slow down the control voltage going into the chip with RC networks are not likely to be successful for this reason. In any case, slowing down the control voltage change simply converts a click to a thump; the FET gates are moving through the same voltage range, and the feedthrough capacitance has not altered, so the same amount of electric charge has been transferred to the audio path – it just gets there more slowly.

The only certain way to reduce the effect of transient feedthrough is to soak it up in a lower value of load resistor. The same electric charge is applied to a lower resistor value (the feedthrough capacitance is tiny, and controls the circuit impedance) so a lower voltage appears. Unfortunately reducing the load tends to increase the distortion, as we have already seen; the question is if this is acceptable in the intended application.

Discrete JFET Switching

Having looked in detail at analog switching using CMOS gates, and having seen how well they can be made to work, you might be puzzled as to why anyone should wish to perform the same function with discrete JFETs. There are at least two advantages in particular applications.

www.raguvaran.puzl.com

Firstly, JFETs can handle the full output range of op-amps working from maximum supply rails, so higher signal levels can often be switched directly without requiring op-amps to convert between current and voltage mode.

Secondly, the direct access to the device gate allows relatively slow changes in attenuation (though still measured in milliseconds, for reasons that will emerge) rather than the rapid on–off action that CMOS gates give as a result of their internal control-voltage circuitry. This is vital in creating mute circuits that essentially implement a fast fade rather than a sharp cut, and so do not generate clicks and thumps by abruptly interrupting the signal. The downside is that they require carefully tailored voltages to drive the gates, and these cannot always be conveniently derived from the usual op-amp supply rails.

JFETs in Voltage Mode: The Series JFET Switch

The basic JFET series switching circuit is shown in Figure 16.18. With the switch open there is no other connection to the gate other than the bootstrap resistor R_{boot}, V_{gs} is zero, and so the FET is on. When the switch is closed, the gate is pulled down to a sufficiently negative voltage to ensure that the FET is biased off even when the input signal is at its negative limit.

The JFET types J111 and J112 are specially designed for analog switching and pre-eminent for this application. The channel on-resistances are low and relatively linear. This is a depletion-mode FET, which requires a negative gate voltage to actively turn it off. The J111 requires a more negative V_{gs} to ensure it is off, but in return gives a lower $R_{ds(on)}$, which means lower distortion.

J111, J112 (and J113) are members of the same family – in fact they are the same device, selected for gate/channel characteristics, unless I am much mistaken. Table 16.5 shows how the J111 may need 10 V to turn it off, but gives a 30 Ω on-resistance or $R_{ds(on)}$ with zero gate voltage. In contrast the J112 needs only 5.0 V at most to turn it off, but has a higher $R_{ds(on)}$ of 50 Ω. The trade-off is between ease of generating the gate-control voltages, and linearity. The higher the $R_{ds(on)}$, the higher the distortion, as this is a non-linear resistance.

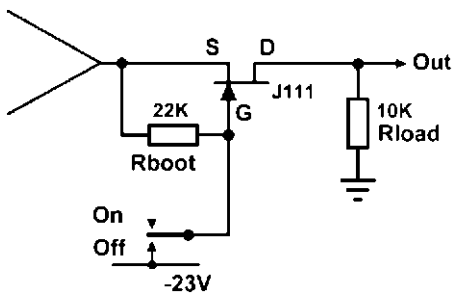


Figure 16.18: The basic JFET switching circuit, with gate bootstrap resistor

www.raguvaran.puzl.com

TABLE 16.5 Characteristics of the J111 JFET series

	J111	J112	J113
$V_{gs(off)}$ min	−3.0	−1.0	−0.5
$V_{gs(off)}$ max	−10	−5.0	−3.0
$R_{ds(on)}$	30	50	100

FET tolerances are notoriously wide, and nothing varies more than the V_{gs} characteristic. It is essential to take the full range into account when designing the control circuitry.

Both the J111 and J112 are widely used for audio switching. The J111 has the advantage of the lowest distortion, but the J112 can be driven directly from 4000 series logic running from ±7.5 V rails, which is often convenient. The J113 appears to have no advantage to set against its high $R_{ds(on)}$ and is rarely used – I have never even seen one.

The circuits below use either J111 or J112, as appropriate. The typical version used is shown, along with typical values for associated components.

Figure 16.18 has S (source) and D (drain) marked on the JFET. In fact the J11x devices appear to be perfectly symmetrical, and it seems to make no difference which way round they are connected, so further diagrams omit this. As JFETs, in practical use they are not particularly static sensitive.

The off voltage must be sufficiently negative to ensure that V_{gs} never becomes low enough to turn the JFET on. Since a J111 may require a V_{gs} of −10 V to turn it off, the off voltage must be 10 V below the negative saturation point of the driving op-amp – hence the −23 V rail. This is not exactly a convenient voltage, but the rail does not need to supply much current and the extra cost in something like a mixing console is relatively small.

To turn a JFET on, the V_{gs} must be held at 0 V. That sounds simple enough, but it is actually the more difficult of the two states. Since the source is moving up and down with the signal, the gate must move up and down in exactly the same way to keep V_{gs} at zero. This is done by bootstrap resistor R_{boot} in Figure 16.18. When the JFET is off, DC flows through this resistor from the source; it is therefore essential that this path be DC coupled and fed from a low impedance such as an op-amp output, as shown in these diagrams. The relatively small DC current drawn from the op-amp causes no problems.

Figure 16.19 is a more practical circuit using a driver transistor to control the JFET. (If you had a switch contact handy, you would presumably use it to control the audio directly.) The pull-up resistor R_c keeps diode D reverse-biased when the JFET is on; this is its sole function, so the value is not critical. It is usually high to reduce power consumption. I have used anything between 47 and 680 kΩ with success.

www.raguvaran.puzl.com

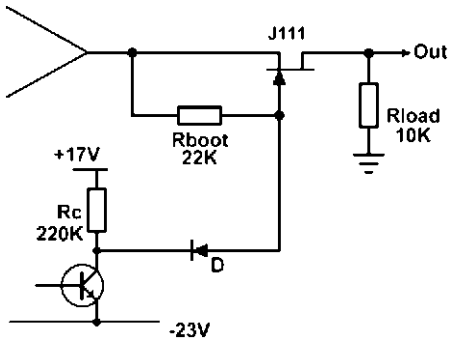


Figure 16.19: Using a transistor and diode for gate control

Sometimes DC blocking is necessary if the op-amp output is not at a DC level of 0 V. In this case the circuit of Figure 16.20 is very useful; the audio path is DC blocked but not the bootstrap resistor, which must always have a DC path to the op-amp output. R_{drain} keeps the capacitor voltage at zero when the JFET is held off.

Figure 16.21 shows the distortion performance with a load of 10 kΩ. The lower curve is the distortion from the 5532 op-amp alone. The signal level was 7.75 Vrms (+20 dBu).

Figure 16.22 shows the distortion performance with various heavier loadings, from 10 kΩ down to 1 kΩ. As is usual in the world of electronics, heavier loading makes things worse. In this case, it is because the non-linear R_{on} becomes a more significant part of the total circuit resistance. The signal level was 7.75 Vrms (+20 dBu).

Figure 16.23 shows the distortion performance with different values of bootstrap resistor. The lower the value, the more accurately the drain follows the source at high audio frequencies, and so the lower the distortion. The signal level was 7.75 Vrms (+20 dBu) once again. There appears to be no disadvantage to using bootstrap resistors of 22 kΩ or so, except in special circumstances, as explained below.

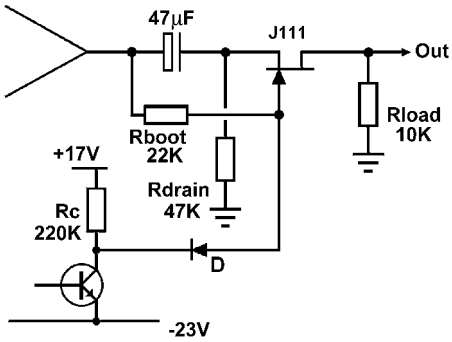


Figure 16.20: The JFET switching circuit with a DC-blocking capacitor

www.raguvaran.puzl.com

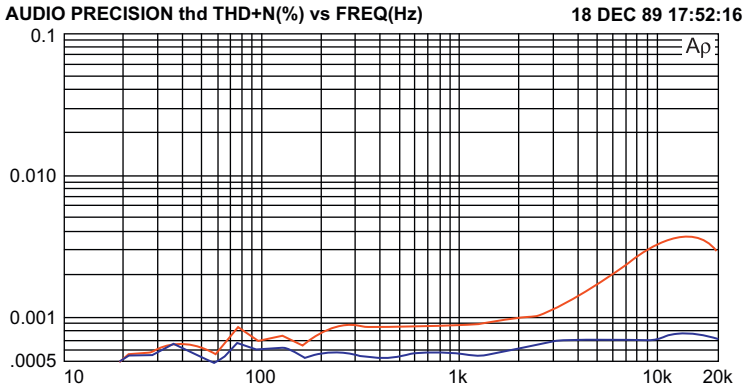


Figure 16.21: The JFET distortion performance with a load of 10k

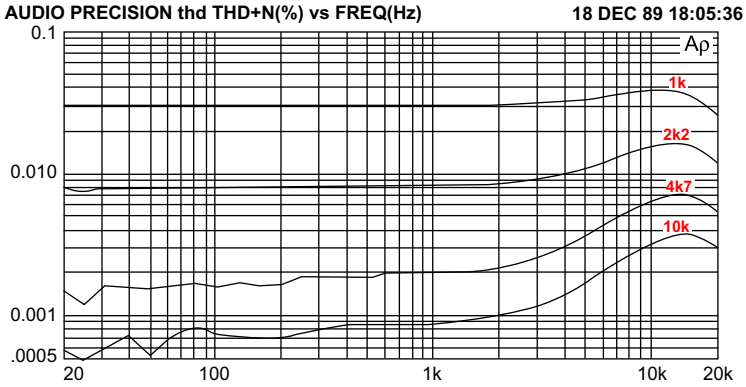


Figure 16.22: The JFET distortion performance versus loading

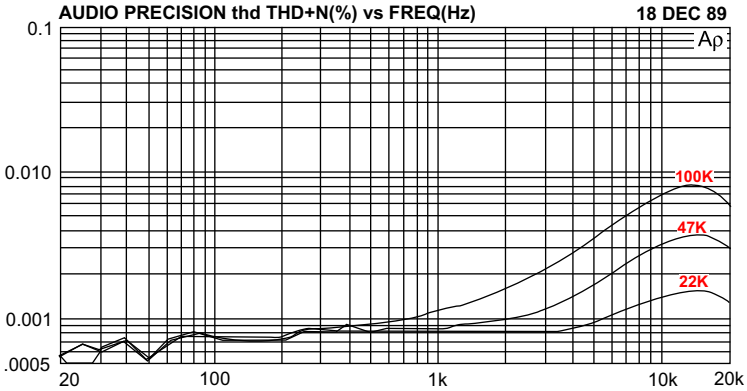


Figure 16.23: The distortion performance with different values of bootstrap resistor

www.raguvaran.puzl.com

Two series JFET switches can be simply combined to make a changeover switch, as shown in Figure 16.24. The valid states are A on, B on, or both off. Both on is not a good option because the two op-amps will then be driving each other's outputs through the JFETs.

It is possible to cascade FET switches, as in Figure 16.25, which is taken from a real mixing console application. Here the main output is switched between A and B as before, but a second

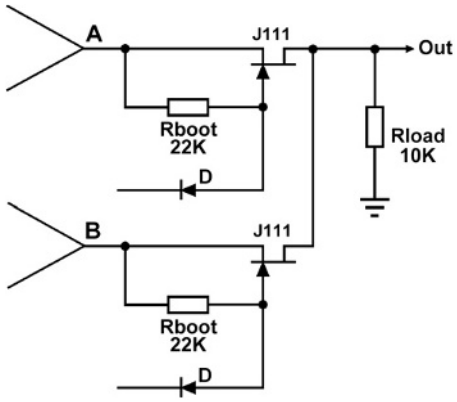


Figure 16.24: A JFET changeover switch

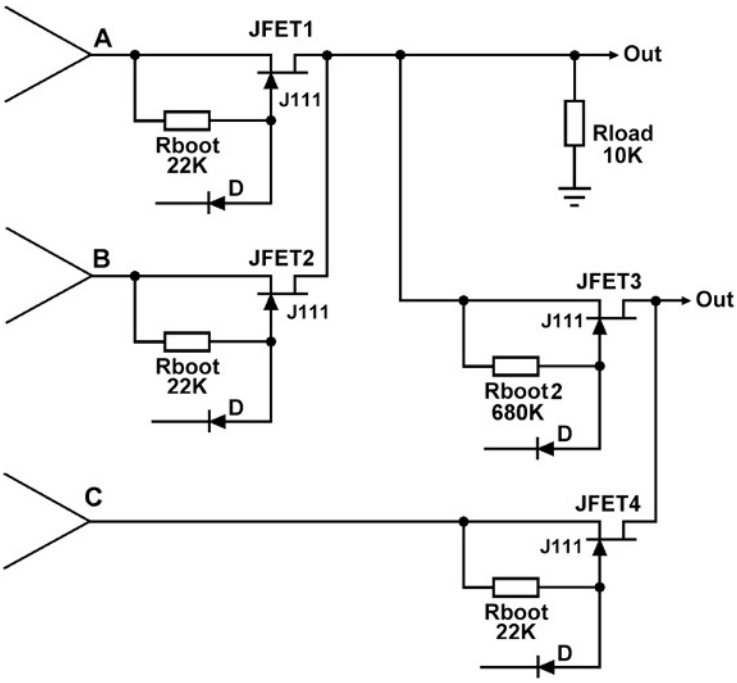


Figure 16.25: Cascaded FET switches

www.raguvaran.puzl.com

auxiliary output is switched between this selection and another input C by JFET3 and JFET4. The current drawn by the second bootstrap resistor R_{boot2} must flow through the $R_{ds(on)}$ of the first FET, and will thus generate a small click. R_{boot2} is therefore made as high as possible to minimize this effect, accepting that the distortion performance of the JFET3 switch will be compromised at HF; this was acceptable in the application as the second output was not a major signal path. The bootstrap resistor of JFET4 can be the desirable lower value as this path is driven direct from an op-amp.

The Shunt JFET Switch in Voltage Mode

The basic JFET shunt switching circuit is shown in Figure 16.26. Like the shunt analog gate mute, it gives poor offness but good linearity in the On state, so long as its gate voltage is controlled so it never allows the JFET to begin conducting. Its great advantage is that the depletion JFET will be in its low-resistance state before and during circuit power-up, and can be used to mute switch-on transients. Switch-off transients can also be effectively muted if the drive circuitry is configured to turn on the shunt FETs as soon as the mains disappears, and keep them on until the various supply rails have completely collapsed.

I have used the circuit of Figure 16.26 to mute the turn-on and turn-off transients of a hi-fi preamplifier. Since this is an output that is likely to drive a reasonable length of cable, with its attendant capacitance, it is important to keep R1 as low as possible, to minimize the possibility of a drooping treble response. This means that the $R_{ds(on)}$ of the JFET puts a limit on the offness possible. The output series resistor R1 is normally in the range 47–100 Ω, when it has as its only job the isolation of the output op-amp from cable capacitance. Here it has a value of 1 kΩ, which is a distinct compromise – it is not suited for use with very long cables. Even with this value the muting obtained was not quite adequate at −27 dB so two J111s were used in parallel, giving a further −6 dB of attenuation. The resulting −33 dB across the audio band

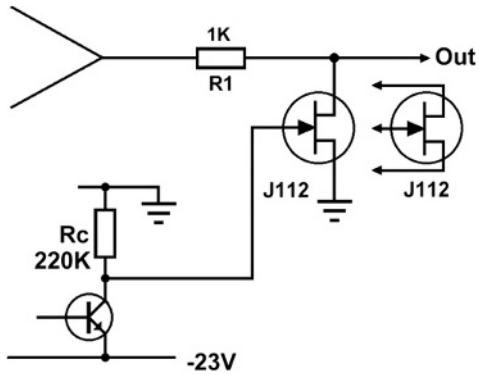


Figure 16.26: The basic JFET shunt switching circuit. Adding more JFETs in parallel increases the offness, but each −6 dB requires doubling the number

www.raguvaran.puzl.com

was sufficient to render the transients inaudible. The offness is not frequency dependent as the impedances are low and so stray capacitance is irrelevant.

JFETs in Current Mode

JFETs can be used in the current mode, just as for analog gates. Figure 16.27 shows the basic muting circuit, with series FET switching only. R_{in2} attenuates the signal seen by the FET when it is off, to prevent breakthrough; its presence means that the gain of the circuit is somewhat less than unity with the values shown, but the gain may be readily adjusted by altering the value of R_{nfb}. Figure 16.28 illustrates the distortion performance and Figure 16.29 the offness of this circuit. Neither is startlingly good.

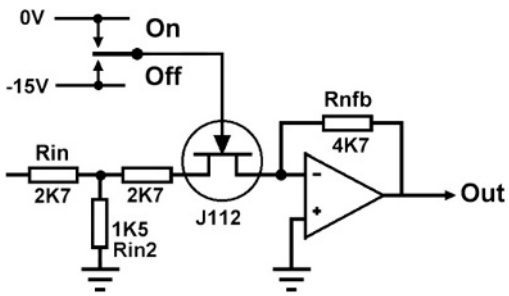


Figure 16.27: The simplest FET mute circuit: single-FET muting forces a crosstalk/linearity trade-off

In designing a mute block, we want low distortion *and* good offness at the same time, so the series-shunt configuration, which proved highly effective with CMOS analog gates, is the obvious choice. The basic circuit is shown in Figure 16.30, the distortion performance is illustrated in Figure 16.31, and the offness in Figures 16.32 and 16.33. A small capacitor C across the feedback resistor is usually required to ensure HF stability, due to the FET capacitances hanging on the summing node at D.

Due to the shunt-feedback configuration, this circuit introduces a phase inversion. I have often been forced to follow this circuit with another inverting stage that does nothing except get the phase right again. In this situation, it is sometimes advantageous to put the inverting stage *before* the mute block, so that any crosstalk to its sensitive summing node is muted with the signal by the following mute block.

The control voltages to the series and shunt JFETs are complementary as with the CMOS version, but now they can be slowed down by RC networks to make the operation gradual, as shown in Figure 16.34. The exact way in which the control voltages overlap is easy to control, but the V_{gs}/resistance law of the FET is not (and it is about the most variable FET parameter there is) and so the overlap of FET conduction is rather variable. However, I should say at once

www.raguvaran.puzl.com

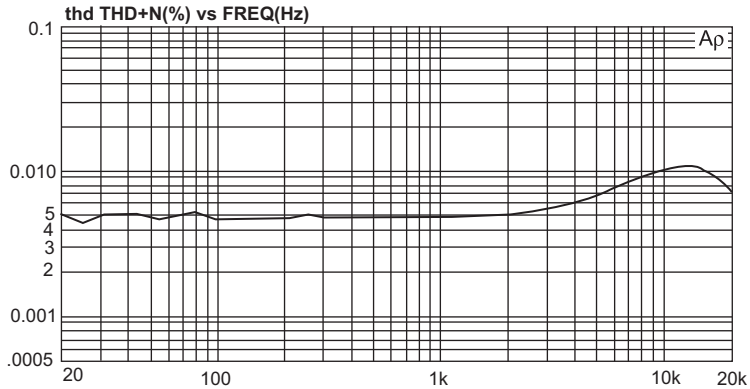


Figure 16.28: THD of the single-FET circuit in Figure 16.27

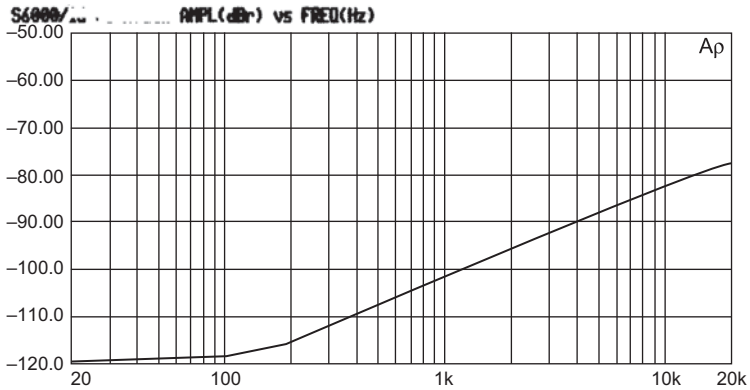


Figure 16.29: Offness of the single-FET circuit in Figure 16.27. It only manages −982 dB at 10 kHz

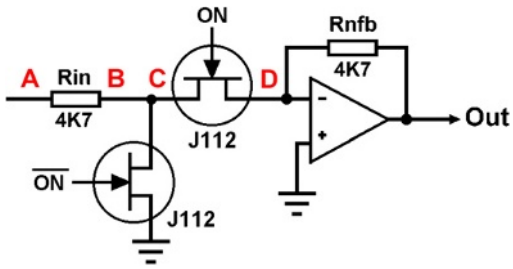


Figure 16.30: Series-shunt mode mute bloc circuit

www.raguvaran.puzl.com

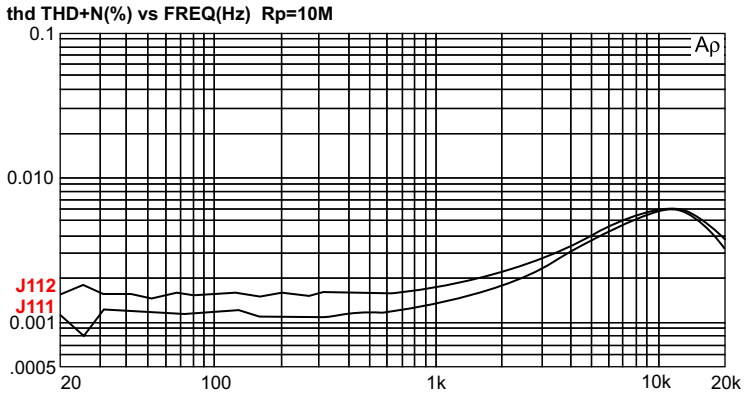


Figure 16.31: The THD of the mute bloc in Figure 16.30, with R = 4K7 and R_p = 10 MΩ. The increase in JFET distortion caused by using a J112 rather than a J111 is shown. The rising distortion above 1 kHz comes from the op-amp

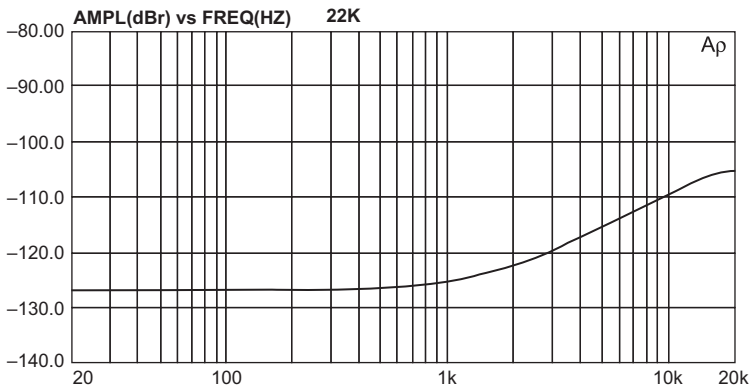


Figure 16.32: Offness of mute bloc in Figure 16.27 with R_{in} = R_{nfb} = 22k

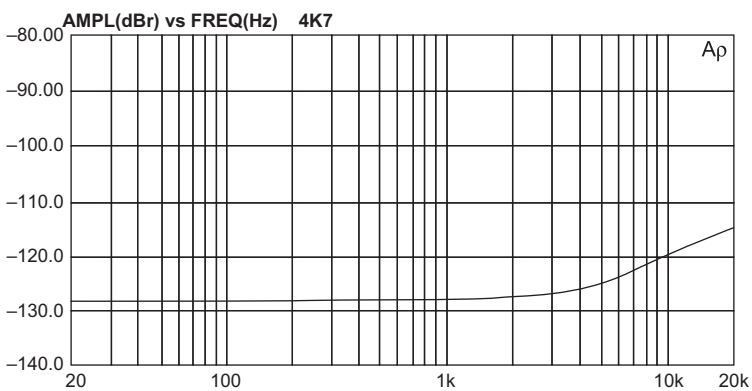


Figure 16.33: Offness of mute bloc in Figure 16.27 with R_{in} = R_{nfb} = 4K7. Offness is better and the noise floor (the flat section below 2 kHz) has been lowered by about 2 dB

www.raguvaran.puzl.com

that this system does work, and works well enough to go in top-notch mixing consoles. As you go into the muted condition the series JFET turns off and the shunt JFET turns on, and if the overlap gets to be too much in error, the following bad things can happen:

1. If the shunt FET turns on too early, while the series JFET is still mostly on, a low-resistance path is established from the op-amp VE point to ground, causing a large but brief rise in stage noise gain. This produces a 'chuff' of noise at the output as muting occurs.

2. If the shunt FET turns on too late, so the series JFET is mostly off, and the large signal voltage presented to the series FET causes visibly serious distortion. I say 'visibly' because it is well known that even quite severe distortion is not obtrusive if it occurs only briefly. The transition here is usually fast enough for this to be the case; it would not, however, be a practical way to generate a slow fade. The conclusion is that we should err on the side of distortion rather than noise.

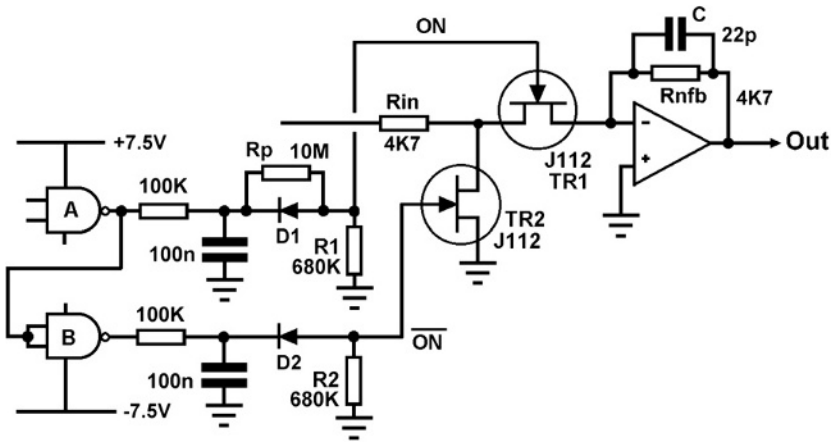


Figure 16.34: Circuitry to generate drive voltages for series-shunt JFET mute bloc

Reducing Distortion by Biasing

The distortion generated by this circuit block is of considerable importance, because if the rest of the audio path is made up of 5532 op-amps – which is likely in professional equipment – then this stage can generate more distortion than the rest of the signal path combined, and dominate this aspect of the performance. It is therefore worth examining any way of increasing the linearity that we can think of.

We have already noted that to minimize distortion, the series JFET should be turned on as fully as possible to minimize the value of the non-linear $R_{ds(on)}$. When a JFET has a zero gate-source voltage, it is normally considered fully on. It is, however, possible to turn it even more on than this.

www.raguvaran.puzl.com

The technique is to put a small positive voltage on the gate, say about 200–300 mV. This further reduces the $R_{ds(on)}$ in a smoothly continuous manner, without forward biasing the JFET gate junction and injecting DC into the signal path. This is accomplished in Figure 16.34 by the simple addition of R_p, which allows a small positive voltage to be set up across the 680k resistor R1. The value of R_p is usually in the 10–22 MΩ range, for the circuit values shown here.

Care is needed with this technique, because if temperatures rise the JFET gate diode may begin to conduct after all, and DC will leak into the signal path, causing thumps and bangs. In my experience 300 mV is about the upper safe limit for equipment that gets reasonably warm internally, i.e. about 50°C. Caution is the watchword here, for unwanted transients are *much* less tolerable than slightly increased distortion.

As with analog CMOS gates, an important consideration with this circuit is the impedance at which it works, i.e. the values of R_{in} and R_{nfb}. These are usually of equal resistance for unity gain so we will call their value R:

1. Raising R reduces distortion because it minimizes the effect of $R_{ds(on)}$ variation in the series JFET.
2. Lowering R reduces the noise generated by the circuit, and improves offness as it reduces the effect of stray capacitances. It also reduces the effect of control-voltage feedthrough via the gate-channel capacitances.

Figures 16.33 and 16.34 examine how the offness of the circuit is affected by using values of 4k7 and 22 kΩ. The latter gives −110 dB rather than −120 dB at 10 kHz. In my not inconsiderable experience with this circuit, R = 4k7 is the best choice when J112s are used. Values below 4k7 are not usual as distortion will increase as the JFET $R_{ds(on)}$ becomes a larger part of the total resistance in the circuit. The loading effect of R_{in} on the previous stage must also be considered.

JFET Drive Circuitry

The series-shunt mute bloc requires two complementary drive voltages, and these are most easily generated by 4000 series CMOS running from ±7.5 V rails. NAND gates are shown in Figure 16.34 as they are convenient for interfacing with other bits of control logic, but any standard CMOS output can be used. It is vital that the JFET gates get as close to 0 V as possible, ensuring that the series gate can be fully on and give minimum distortion, so the best technique is to run the logic from these ± rails and use diodes to clamp the gates to 0 V.

Thus, in Figure 16.34, when the mute bloc is passing signal, the signal from gate A is high, so D1 is reverse-biased and the series JFET TR1 gate is held at 0 V by R1, keeping it on (the role of R_p was explained in the previous section). Meanwhile, D2 is conducting as the NAND-gate output driving it is low, so the shunt JFET TR2 gate is at about −7 V and it is

www.raguvaran.puzl.com

firmly switched off. This voltage is more than enough to turn off a J112, but cannot be guaranteed to turn off a J111, which may require −10 V (see Table 16.5). This is one reason why the J112 is more often used in this application – it is simpler to generate the control voltages. When the mute bloc is off, the conditions are reversed, with the output of A low, turning off TR1, and the output of B high, turning on TR2.

When switching audio signals, an instantaneous cut of the signal is sometimes not what is required. When a non-zero audio signal is abruptly interrupted there is bound to be a click. Perhaps surprisingly, clever schemes for making the instant of switching coincide with a zero crossing give little improvement. There may no longer be a step-change in level, but there is still a step-change in slope and the ear once more interprets this discontinuity as a click.

What is really needed is a fast-fade over about 10 ms. This is long enough to prevent clicks, without being so slow that the timing of the event becomes sloppy. This is normally only an issue in mixing consoles, where it is necessary for things to happen in real time. Such fast-fade circuits are often called 'mute blocks' to emphasize that they are more than just simple on–off switches. Analog gates cannot be slowly turned on and off due to their internal circuitry for control-voltage generation, so discrete JFETs must be used. Custom chips to perform the muting function have been produced, but the ones I have evaluated have been expensive, single source, and give less than startling results for linearity and offness; this situation is of course subject to change.

In Figure 16.34 the rate of control voltage is determined by the RC networks at the NAND-gate outputs. Ingenious schemes involving diodes to make the up/down rates different have been tried many times but my general conclusion is that they give little if any benefit.

Physical Layout and Offness

The offness of this circuit is extremely good, provided certain precautions are taken in the physical layout. In Figure 16.35 there are two possible crosstalk paths that can damage the offness. The path C–D, through the internal capacitances of the series JFET, is rendered innocuous as C is connected firmly to ground by the shunt JFET. However, point A is still alive

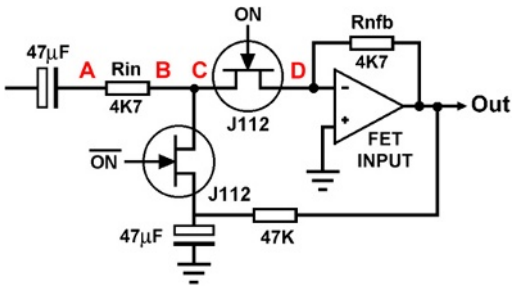


Figure 16.35: Circuit of JFET mute showing stray capacitances and DC handling

www.raguvaran.puzl.com

with full amplitude signal, and it is the stray capacitance from A to D that defines the offness at high frequencies.

Given the finite size of R_{in}, it is often necessary to extend the PCB track B–C to get A far enough from D. This is no problem if done with caution. Remember that the track B–C is at virtual earth when the mute bloc is on, and so vulnerable to capacitative crosstalk from other signals straying into the area.

Dealing with the DC Conditions

The circuits shown so far have been stripped down to their bare essentials to get the basic principles across. In reality, things are (surprise) a little more complicated. Op-amps have non-zero offset and bias voltages and currents, and if not handled properly these will lead to thumps and bangs. There are several issues:

1. If there is any DC voltage at all passed on from the previous stage, this will be interrupted along with the signal, causing a click or thump. The foolproof answer is of course a DC-blocking capacitor, but if you are aiming to remove all capacitors from the signal path, you may have a problem. DC servos can partly make up the lack, but since they are based on op-amp integrators they are no more accurate than the op-amp, while DC blocking is foolproof.

2. The offset voltage of the mute block op-amp. If the noise gain is changed when the mute operates (which it is), the changing amplification of this offset will change the DC level at the output. The answer is shown in Figure 16.35. The shunt FET is connected to ground via a blocking capacitor to prevent gain changes. This capacitor does not count as 'being in the signal path' as audio only goes through it when the circuit is muted. Feedback of the op-amp offset voltage to this capacitor via the 47 kΩ resistor renders it innocuous.

3. The input bias and offset currents of the op-amp. These are much more of a problem and are best dealt with by using JFET op-amps such as TL072 or OPA2134, where the bias and offset currents are negligible at normal equipment temperatures. All of the distortion measurements in this chapter were made with TL072 op-amps in place.

A Soft Changeover Circuit

This circuit (Figure 16.36) is designed to give a soft changeover between two inputs – in effect a fast crossfade. It is the same mute block but with two separate inputs, either or both of which can be switched on. The performance at +20 dBu in/out is summarized in Table 16.6.

The THD increase at 20 kHz is due to the use of a TL072 as the op-amp. J112 JFETs are used in all positions.

www.raguvaran.puzl.com

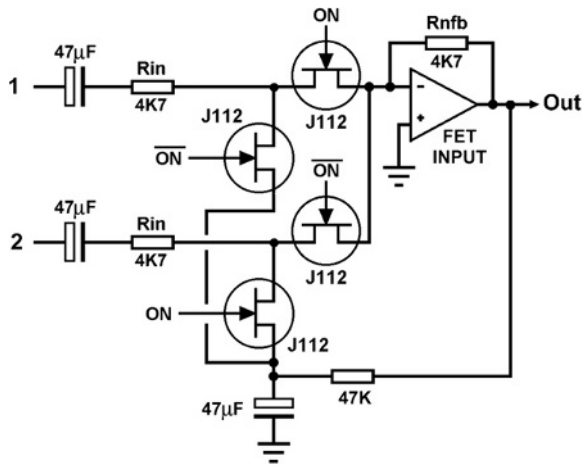


Figure 16.36: Circuit of JFET soft changeover switching

TABLE 16.6 Distortion produced by JFET soft changeover switch

	1 kHz	10 kHz	20 kHz
THD +20 dBu	0.0023%	0.0027%	0.0039%
Offness (dB)	−114	−109	−105

This circuit is intended for soft-switching applications where the transition between states is fast enough for a burst of increased distortion to go unnoticed. It is not suitable for generating slow crossfades in applications like disco mixers, as the exact crossfade law is not very predictable.

Control-voltage feedthrough in JFETs

All discrete FETs have a small capacitance between the gate and the device channel, so changes in the gate voltage will therefore cause a charge to be transferred to the audio path, just as for CMOS analog gates. As before, slowing down the control voltage change tends to give a thump rather than a click; the same amount of electric charge has been transferred to the audio path, but more slowly. Lowering the circuit impedances is effective in reducing feedthrough, but it is of limited effectiveness. Halving the impedance only reduces the amplitude by 6 dB, and such a reduction is likely to increase distortion.

Reference

[1] M. Jones, Designing valve preamps, Electronics World (March 1996) p. 193.

www.raguvaran.puzl.com

This page intentionally left blank

www.raguvaran.puzl.com

CHAPTER 17

Mixer Subsystems

This chapter deals with the specialized circuit blocks that make up a mixing consoles. Some functions, such as microphone amplification, line input and output, and equalization, have already been explored in previous chapters. The other useful blocks are presented here in the order that a signal encounters them as it goes through the console, but many of them are applicable to all the various kinds of module – input channels, groups, and master sections. For example, fader postamplifiers and insert points will be found in all three types of module.

Mixer Bus Systems

In all but the smallest mixers there is a need to connect together all the modules so they have access to the mixing buses, power-supply rails, and logic and control lines.

There are three basic ways of connecting together the modules. The smallest mixers are usually constructed on a single large PCB lying parallel to a one-piece front panel, and here 'modules' means repeated circuitry rather than physically-separate modules. The all-embracing PCB minimizes the money spent on connectors, and the time plugging them in during assembly, but there are obvious limitations to the size of mixer you can build in this way. A definite problem is the need to run summing buses laterally, as this results in them winding their way between controls and circuit blocks, threatening a mediocre crosstalk performance. The use of double-sided PCBs helps greatly with this, but very often there are still awkward points such as the need for the feed to and from the faders to cross over the mix bus area. This can easily wreck the crosstalk performance; one solution is to use what I call a 'three-layer board'. The mix buses are on the bottom of the PCB, the top layer above it carries a section of ground plane, and the fader connections are made by wire links above that. Given a tough solder-resist, no further insulation of the links is necessary; if you have doubts then laying a rectangle of component-ident screen print under the links will give another layer of insulation without adding labor cost.

Medium-sized mixers are commonly made with separate modules, connected together with ribbon cable bearing insulation-displacement connectors (IDCs). The advent of IDC ribbon cables (a long time ago now) had a major effect on the affordability of mixing consoles. These cables naturally join Pin 1 to Pin 1 on every module, and so on, leading to a certain inflexibility in design.

Small Signal Audio Design; ISBN: 9780240521770
Copyright © 2010 Elsevier Ltd; All rights of reproduction, in any form, reserved.

www.raguvaran.puzl.com

Large mixers use a motherboard system, where each module plugs into a PCB at the bottom of the frame, which is typically divided into 'bins' holding eight or 12 modules. This provides (at considerably increased expense) total flexibility in the running of buses and the interconnection of modules.

To me, it is a 'mixing bus' or a 'summing bus'. I realize that some of the world spells it 'buss' and I am probably wasting my time pointing out that the latter is wrong but I am still going to do it. The term dates from the dawn of electrical distribution, when circuits were connected together by copper bars called 'omnibus bars'. This inevitably got shortened to 'bus-bars' and in mixing consoles it was further abbreviated to 'bus', which somehow turned into 'buss'. The Oxford English Dictionary says buss means 'to kiss' (archaic), which seems somehow not quite appropriate. I have grave doubts if my protest here will make any difference to common usage, but sometimes you've just got to make a stand.

Input Arrangements

Most mixer channels have both microphone and line inputs. On the lower-cost consoles these are usually switched to a single amplifier with a wide gain range, the line input being attenuated to a suitable level first. This approach is covered in Chapter 13 on microphone amplifiers. High-end consoles have separate line-input amplifiers, removing some compromises on CMRR and noise performance. Dedicated line-input amplifiers are dealt with in Chapter 14.

Equalization

Mixer input channels have more or less sophisticated tone controls to modify the frequency response, either to correct imperfections or produce specific effects. This subject is fully dealt with in Chapter 10.

Insert Points

The addition of effects for general use, such as reverberation, is normally handled by an effects send system. However, if a specific effect (say, flanging) is going to be used on one channel only then it is far more efficient and convenient to connect the external effects unit in series with the signal path of the channel itself. This is done by means of an insert point (usually just called an 'insert'), which is a jack with normalling contacts arranged so that the signal flows to it and back to the channel again when nothing is plugged in. When a jack is inserted the normalling connection is broken and the signal flows through the external unit.

Inserts are also often fitted to groups.

www.raguvaran.puzl.com

Inserts come in two versions, illustrated in Figure 17.1. The single-jack version is economical in panel space but is restricted to unbalanced operation. The two-jack version is superior because it allows the use of balanced send and returns, and in addition the OUT jack socket can be used as a direct output because inserting a jack in it does not break the signal path; only inserting a jack into the IN jack socket does that.

When the mixer has a patchbay, the insert sends (and, indeed, other console outputs) are likely to find their way there through a quite considerable length of ribbon cable, which has significant capacitance between its conductors. It is easy to get into a situation where the crosstalk performance of the console is limited by capacitative crosstalk between outputs, despite their low impedance. Output amplifiers commonly have a series resistor to isolate the amplifier from the capacitance of the cabling and prevent HF instability, and the minimum safe value of this resistor defines the output impedance, which is usually in the region of 47–100 Ω.

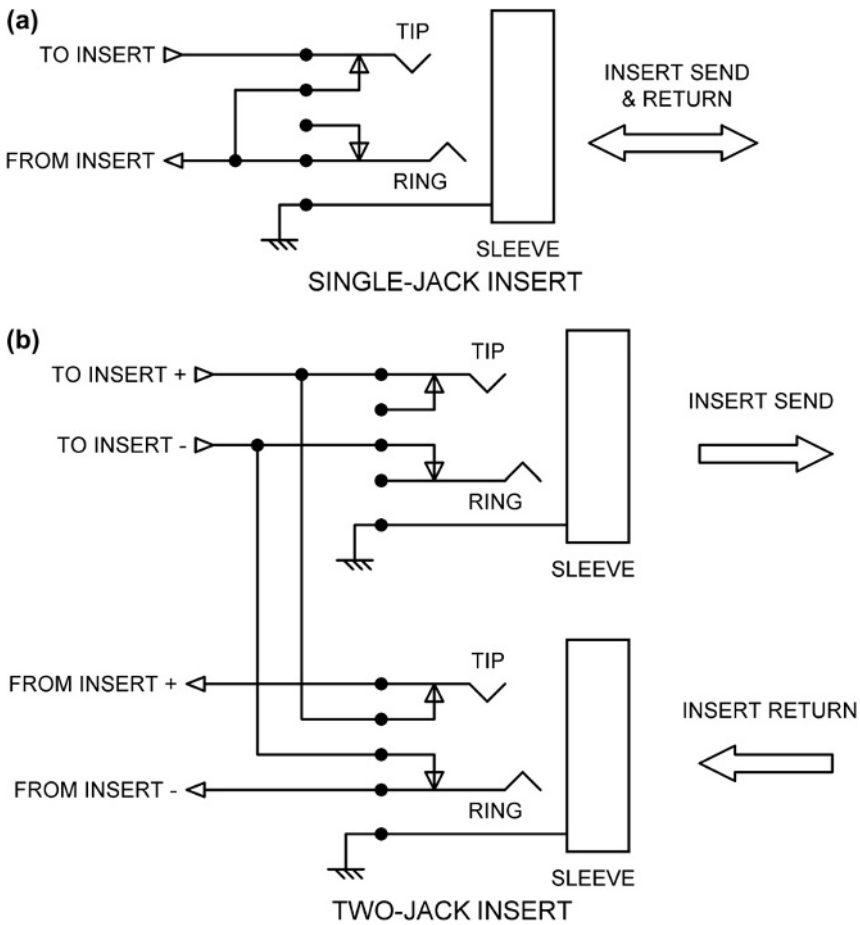


Figure 17.1: The single-jack insert (a) and the double-jack insert (b)

www.raguvaran.puzl.com

Things get worse when the layers of ribbon cable are laid together in a 'lasagne' format; this is very often necessary because of the sheer number of signals going to and from the patchbay. In some cases layers of grounded screening foil are interleaved with the cables, but this is rather expensive and awkward to do, and does not greatly reduce crosstalk between conductors in the same piece of ribbon. The only way to do this is to reduce the output impedance.

In a particular mixer design project, the crosstalk between the insert sends from the channels, with an output impedance of 75 Ω, was found to be −96 dB at 10 kHz. This may not sound like a lot, but I didn't get where I am today by designing consoles with measurable crosstalk, so something had to be done. An effective way to obtain a near-zero output impedance is shown in Figure 17.2. Here the main negative feedback for the op-amp goes through R1, from the outside end of isolating resistor R2 and so reduces the output impedance, while the stabilizing HF feedback is taken through C1 from the inside end, where it is not subject to phase-shift because of load capacitance. With this insert send stage the output impedance was reduced from 75 Ω to less than 1 Ω and the crosstalk disappeared below the noise floor. Very similar circuitry can be used with stages that have gain, see Chapter 15. Arrangements like this must always be carefully checked to make sure that HF stability with a capacitative load really *is* maintained; this circuit is stable when driving a 22 nF load, which represents 220 meters of 100 pF/meter cable.

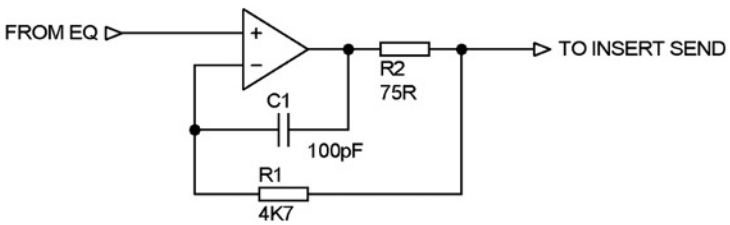


Figure 17.2: A typical insert-send amplifier with 'zero-impedance' output

This arrangement is sometimes called a 'zero-impedance' output; the impedance is certainly much lower than usual but it is not of course actually zero.

In a group module, an inverting insert send amplifier is often used to correct the phase inversion introduced by the summing amplifier. A zero-impedance version of this is illustrated in the section below on summing amplifiers.

How to Move a Circuit Block

In the more sophisticated and versatile mixers it is often possible to rearrange the order of blocks in the signal path. A typical example is the facility to move an insert point from before to after the EQ section, depending on what sort of external processing is being plugged into the insert. Dynamics sections are also often movable to before or after the EQ.

www.raguvaran.puzl.com

Figure 17.3 shows two ways to move a circuit block from one position to another in the signal path. The version in Figure 17.3(a) uses all the sections of a four-changeover switch to do the job. Normally the insert is before the EQ section, but when the switch, which would probably be labeled 'insert post' or whatever abbreviation of that can be fitted in, is pressed the signal passes along path A and reaches the EQ first; it then goes back through path B, through the insert, and then to the output along path C. However, the existence of unused switch contacts gives us a broad hint that there may be a more efficient way to do it.

There certainly is. Figure 17.3(b) shows how to do it with only three changeovers. When the switch is pressed the signal passes along path A and reaches the EQ first, goes back through path B, to the insert, and then out via path C. You might think that this economy is quite pointless because either way you will need to use a physical four-changeover switch, but in fact the spare switch section in Figure 17.3(b) comes in very handy indeed to operate a switch status indicator light-emitting diode (LED). Alternatively the spare switch section can be connected in parallel with one of the other sections in the hope of improving reliability, but …

Paralleling switch sections is not actually as useful for enhancing reliability as you might think. The classic switch problem is contamination of silver-plated contacts; the silver is converted to non-conductive silver sulfide by the action of hydrogen sulfide in the atmosphere. This can come from industrial pollution, but another source is diesel engine exhaust so

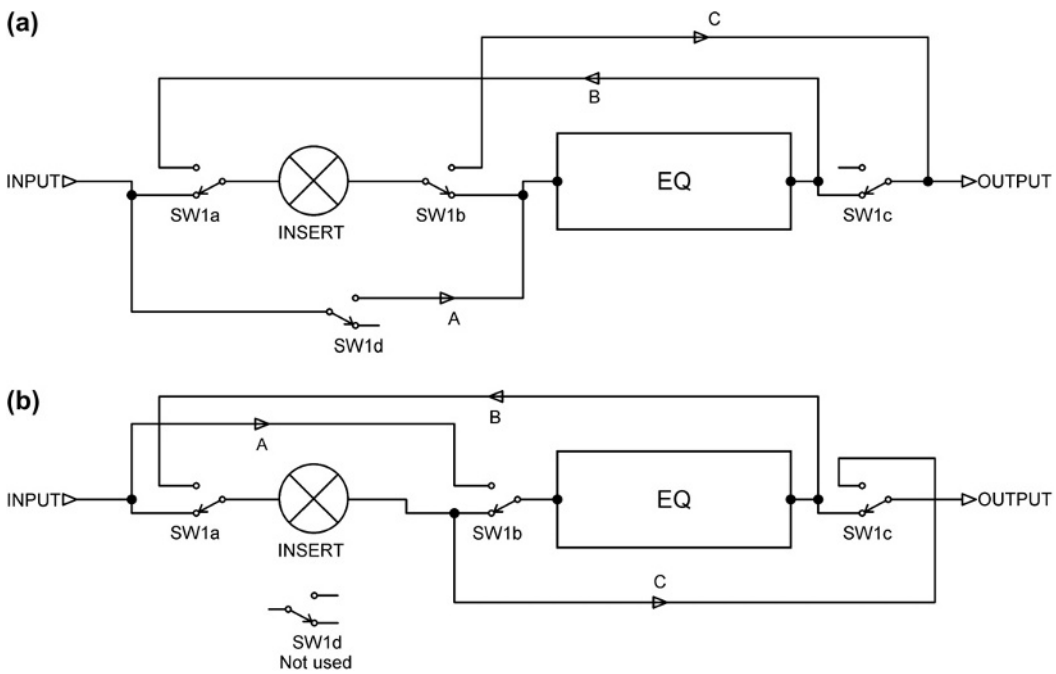


Figure 17.3: Two ways to move a block in the signal path. That at (b) requires only three switch sections instead of four

www.raguvaran.puzl.com

virtually nowhere can be assumed to be free of it. If one set of contacts in a switch is affected then the other set right next to it is certain to be affected also, and paralleling the contacts is actually of little or no use. Sad but true.

Faders

So far as circuit design is concerned, a slide fader can normally be regarded as simply a logarithmic potentiometer, though in fact its internal construction may be quite complicated. The technology of faders is dealt with in Chapter 9 on volume controls.

Postfade Amplifiers

The postfade amplifier (sometimes called just the post amp) is short for postfader amplifier. Its primary function is to allow the fader section to give gain. A +10 dB postamplifier, which is the most common amount of gain, allows the fader to be calibrated up to +10 dB. The 0 dB setting is then some way down the scale and it is much easier to adjust the channel level up and down with respect to the rest of the mix. Without this feature, if a channel was fully faded up at 0 dB, to increase its relative level it would be necessary to pull all the other faders down. When you consider that a large mixer may have 64 channel faders you can see that this would be something of a nuisance.

The second function of the post amp is to buffer the fader from the heavy loading of panpot, mix resistors, and aux sends. The fader control law is carefully chosen for optimal controllability, and that law would be seriously distorted if all that loading came directly on the fader wiper.

It is important that the post amp has good noise performance, because its noise contribution is not much reduced when the fader is pulled down, and it must also have a good load-driving capability at low distortion. For these two reasons, the 5534/2 op-amp found a home in post amps quite early in their history, when its high cost ruled it out for most other circuit functions, which were still handled by TL072s. (The other place that 5534/2s appeared early was in summing amplifiers.)

However, in budget designs TL072s or equivalent FET-input op-amps are still used because they are not only cheaper themselves but allow several other parts to be omitted. In Figure 17.4(a), the op-amp input can be directly connected to the fader wiper because the input bias current is negligible and does not cause significant noise when the wiper is moved along the track. Note that the gain-determining resistors R2, R3 have to be quite high to minimize the loading on the op-amp output. C3 improves stability.

In a more sophisticated design, as in Figure 17.4(b), the use of a 5534 with its significant bias current means that DC-blocking capacitor C1 is required to prevent unpleasant sounds when

www.raguvaran.puzl.com

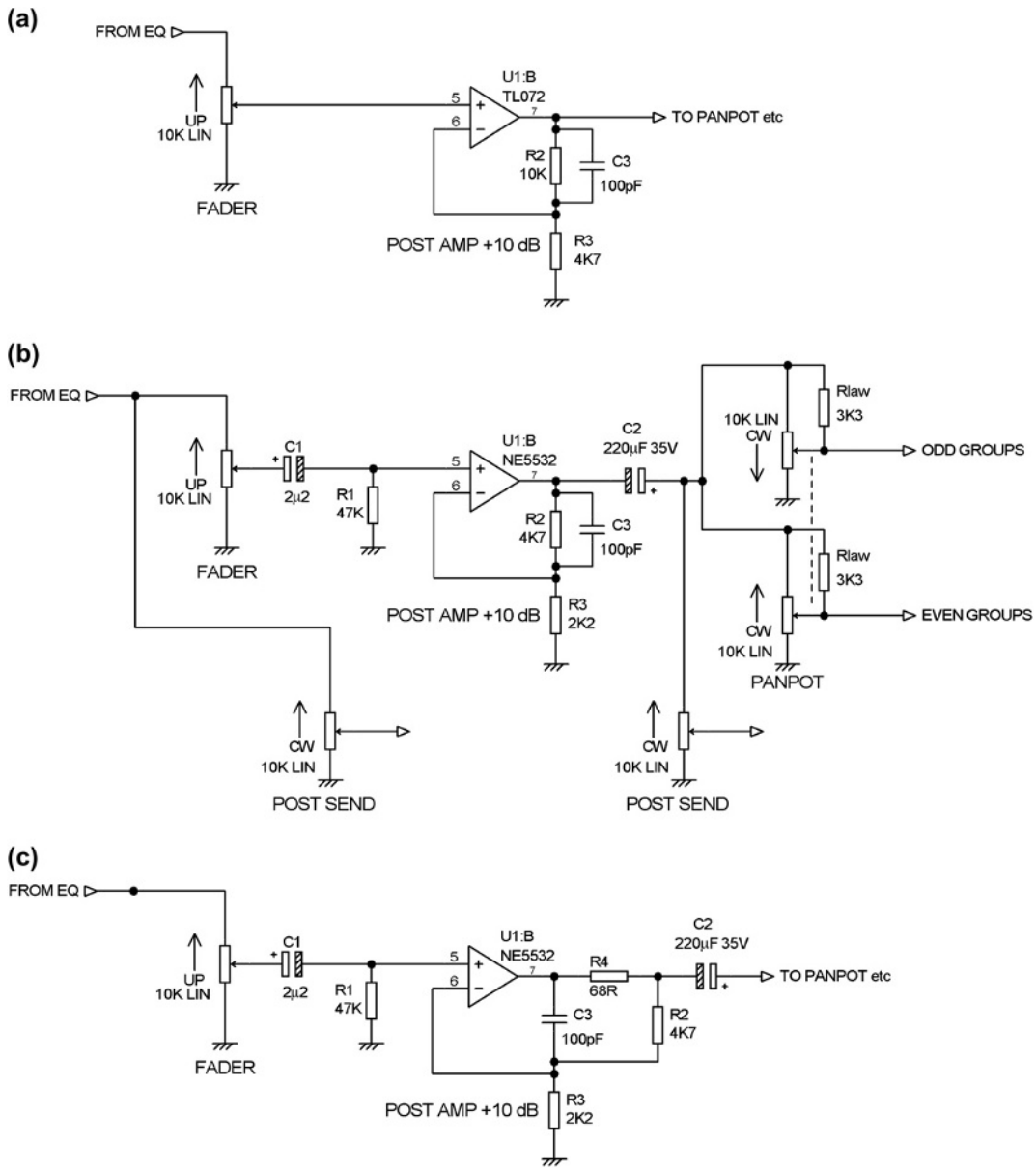


Figure 17.4: Three versions of a postfade amplifier. That at (a) is the cheap and cheerful version, while (b) is the more up-market version, with pre and post sends and the panpot section of a mixer channel shown. (c) protects the op-amp against capacitative loading without sacrificing the 'zero-impedance' output

www.raguvaran.puzl.com

the fader is adjusted; in turn this means that R1 is required to bias the op-amp. Another consequence flows from this: R1 has to have a relatively high value so it does not load the fader and distort its control law, and so the 5534 input bias current causes a relatively large offset voltage at the op-amp input. The 5534 max bias current is 800 nA, which means a possible drop across R1 of −37.6 mV. The 5534 has NPN input devices, so the bias current flows into the input pins, and this voltage is negative, and that is why C1 is the way round it is. This input offset appears multiplied by 3 at the post amp output, giving −113 mV. This is not enough to significantly affect the available output swing, but it does mean that C2 must be introduced to keep DC out of the panpot and postfade sends. Since the loading is quite heavy, C2 must be large to prevent it creating distortion (see Chapter 2). Note also the polarity of C2. We now have not only a more expensive op-amp with a higher power consumption, but three extra components. Their number is multiplied by *N*, the number of channels, so you can see that the use of the 5534 as a post amp is actually a more serious cost decision than it might appear.

You will note that in Figure 17.4(b) the resistors R2, R3 in the feedback network have been reduced in value by a factor of about 2 to improve the noise performance. However, they are still not particularly low in value, and it is true that the noise performance could be improved if they were reduced to, say, 680 and 330 Ω, which in itself would be well within the drive capability of a 5532. However, in simpler desks this op-amp has to drive not only its own feedback network, but also the panpot, routing matrix and the postfade sends, and most of its drive capability has to be reserved for this duty.

Figure 17.4(c) shows a post amp utilizing the 'zero-impedance' approach described above for insert send amplifiers. In some layouts the panpot, routing, and postfade sends are physically spread out so that the stray capacitance seen by the post amp output is enough to imperil its stability. An output-isolating resistor R4 cures the problem but if simply stuck in the output line causes the level after it to vary as changing panpot and send settings alter the load on it. The feedback resistor R2 is therefore fed from after R4, preserving the 'zero-impedance' output, but C3 is fed from before it, maintaining the HF stability.

Fader post amps almost invariably have a fixed gain; if they did not the carefully designed fader law would no longer be obtained. However, there are places where making the gain effectively variable by the use of positive feedback allows the post amp to have a low gain when its associated control is at a low setting; this minimizes the noise contribution of the post amp. There is more on this technique in the section on aux masters.

Direct Outputs

In more complex consoles, a Direct Out is available from each channel. This is a postfade signal, which can be fed directly to a recording device without sending it through the routing and summing systems. This gives a minimum signal path, which will exhibit less noise and

www.raguvaran.puzl.com

possibly less distortion. In all except the most elaborate consoles the Direct Out tends to be unbalanced, to save hardware, and reliance is placed on the recording device having balanced inputs.

Panpots

The word 'panpot' is short for 'panoramic potentiometer'. It is the control that places a monaural source in the desired place in the Left–Right stereo scene. It is an extremely fortunate property of human hearing that this can be done effectively simply by altering the proportion of the mono signal that is sent to the left and right channels of the stereo output. Changing the perception of up and down is considerably more complicated, and outside the scope of mixer design.

The earliest attempt at 'panpotting' dates back to before the use of electrical amplification. In 1903 the French engineer Leon Gaumont was granted patents for loudspeaker systems to go with his sound-on-gramophone-disc talking films. Gaumont was the first to suggest placing loudspeakers behind the screen, with men carrying them about to follow the images on the screen. This procedure has never found favor in the sound industry, especially amongst those who might be asked to do the carrying.

To give smooth stereo panning without unwanted level changes, the panpot should theoretically have a sine/cosine characteristic, as shown in Figure 17.5; note that with this law

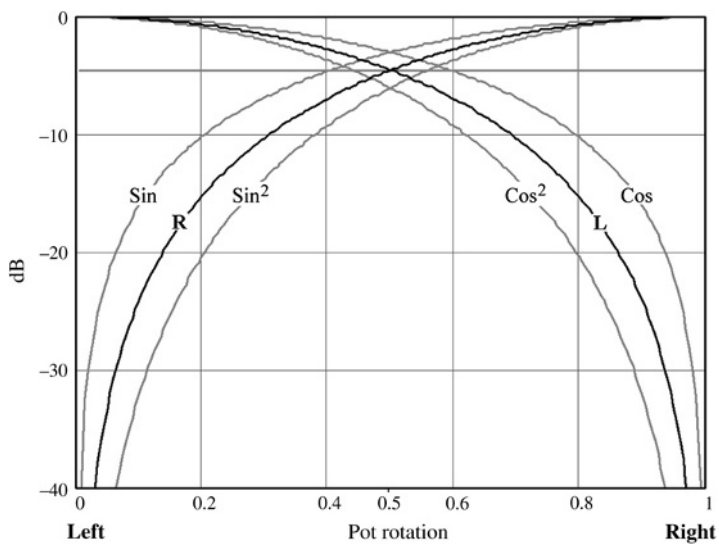


Figure 17.5: Panpot laws. The ideal compromise between the sine and sine-squared laws is the arithmetical mean of the two, shown here

www.raguvaran.puzl.com

both signals are attenuated by 3 dB when the panpot is set centrally. This is desirable because when listening to stereo over loudspeakers, the signals that sum at each ear are not correlated, and so with equal amplitude the two signals give a result that is only 3 dB above the level of each of the components. The summation of two uncorrelated white noise sources to give a combined signal that is 3 dB, and not 6 dB, greater works on exactly the same principle. The sine/cosine law is therefore appropriate for stereo.

However, when you are listening in mono, such as via AM radio, the left and right signals have been summed when they are still in the electrical domain, and therefore are in phase and sum together arithmetically. In this case the pan law needs to be −6 dB down at the center rather than −3 dB if level changes are to be avoided, so in this case a sine-squared/cosine-squared law is required. Since it is highly desirable to cope with both cases, the standard answer is to use a compromise between the two laws. This is entirely satisfactory in practice, not least because moving a component of the mix around during the performance is completely inappropriate for most forms of music, and I have yet to hear a music reviewer complain of 'bad panpotting'.

Figure 17.5 shows both the sine and sine-squared theoretical laws, and also the lines L and R, which indicate the ideal compromise law; it is the arithmetical mean of sine and sine-squared and so is −4.5 dB down with the panpot central. So, how do we obtain the compromise law we want?

Passive Panpots

Figure 17.6 shows three different ways to make a panpot. Panpot circuits can be made with single pots with the wiper grounded, as shown in Figure 17.6(a), but this has several disadvantages, not least the poor offness when panned hard over caused by the current flowing through the wiper contact resistance to ground.

Potentiometers with an accurate sine/cosine law used to exist, which in theory allowed the use of the simple panpot circuit in Figure 17.6(b), but they were always bulky and prohibitively expensive for audio use, and seem to have disappeared altogether now that analog computers are no longer at the cutting edge of technology. Hi-fi preamplifiers sometimes use a control with a 'balance law' for adjusting channel balance (see Chapter 9 for more on that), but this usually has no attenuation at all at the center point and would make a poor panpot.

The traditional answer to the problem has always been to use dual linear pots and bend the linear law into an approximation of the compromise law we want by the use of fixed pull-up resistors connected to the wipers, as shown in Figure 17.6(c). Here the panpot wipers are loaded with two mix resistors of 10 kΩ, which being connected to virtual-earth buses are connected to ground as far as the panpot is concerned, and the pull-up resistors R_{law} are made 3k9 to give the desired 4.5 dB of attenuation at the center position of the panpot.

www.raguvaran.puzl.com

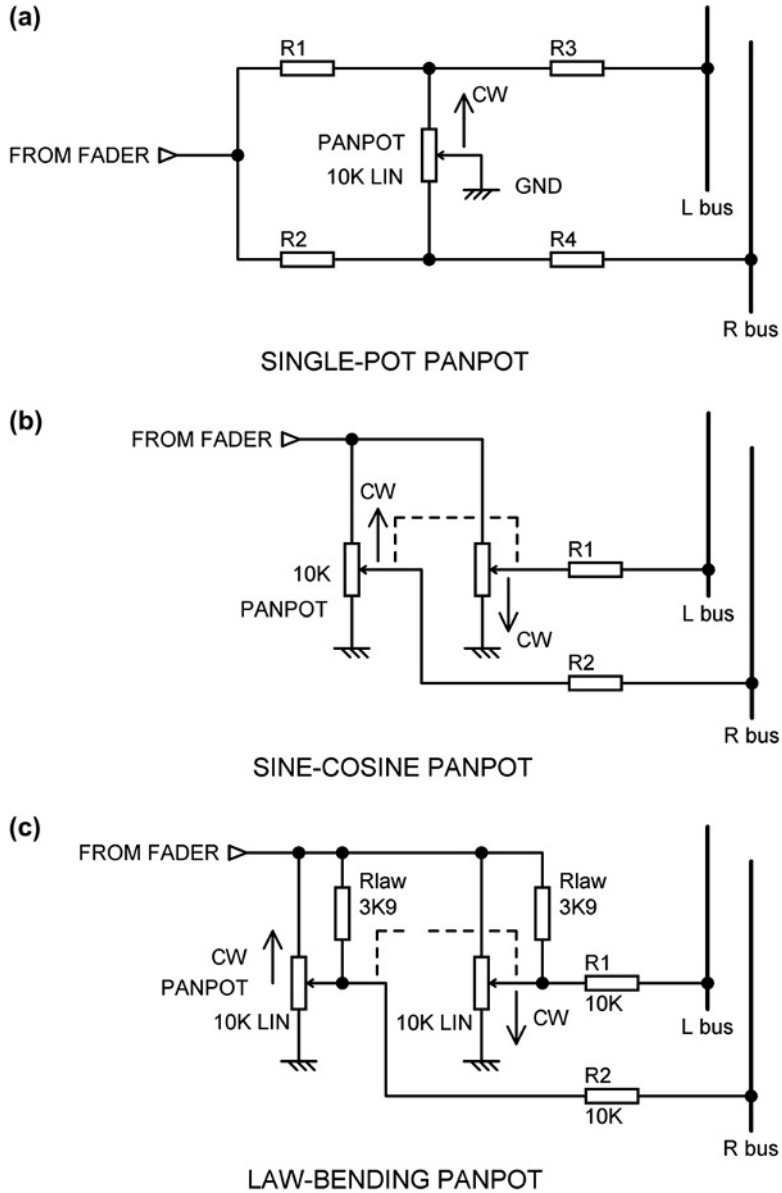


Figure 17.6: Three ways to make a panpot; (c) uses a linear pot with law-bending resistors, and is the most popular

There is an important factor in the design of panpot circuitry, and here a significant difference between low-cost and high-cost mixers intrudes. A low-cost mixer will drive the routing resistors directly from the panpot wipers, as shown in Figure 17.7(a), and the panpot law is affected by the loading to ground represented by the routing resistors; they work in opposition

www.raguvaran.puzl.com

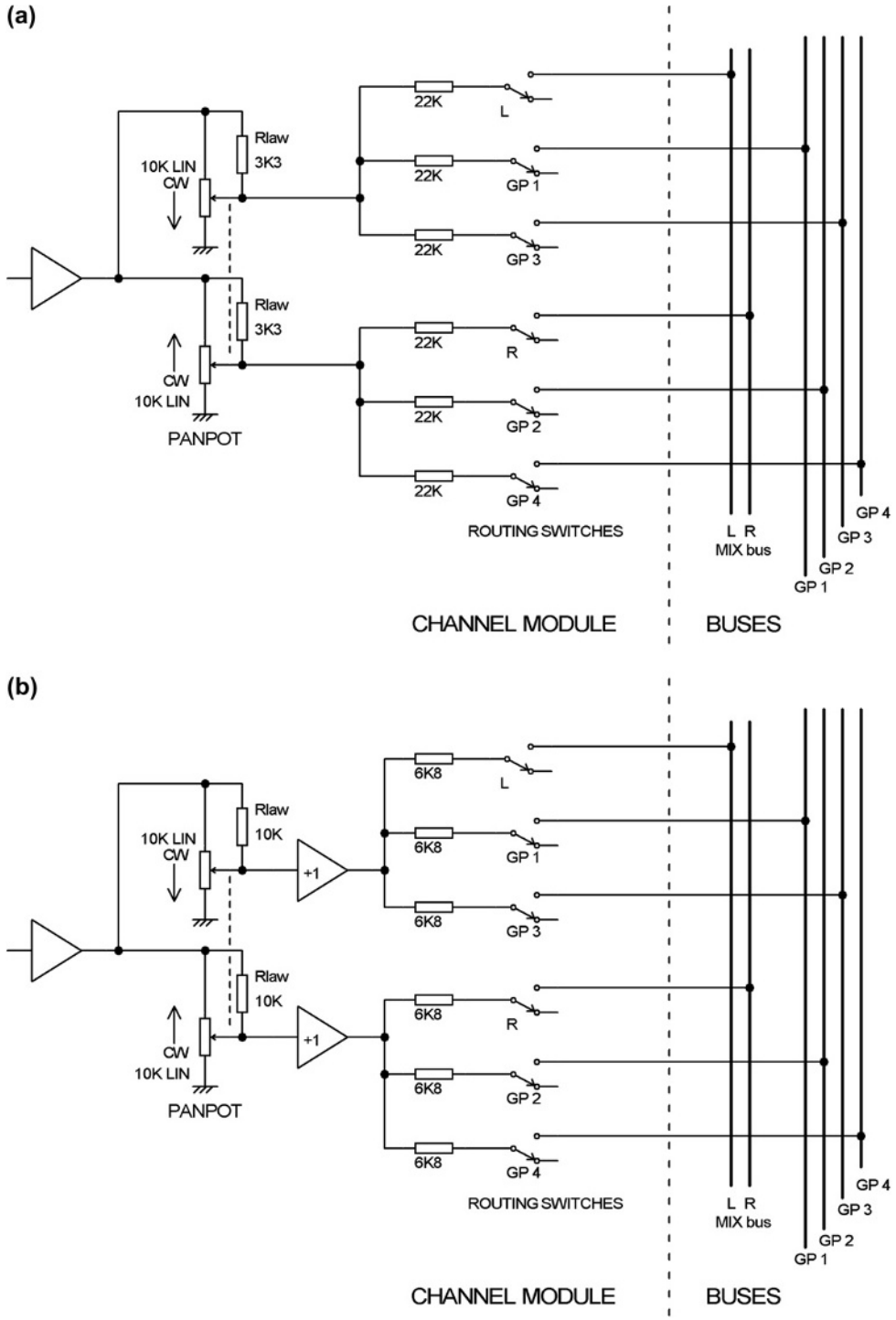


Figure 17.7: A panpot with law-bending resistors: unbuffered and buffered

www.raguvaran.puzl.com

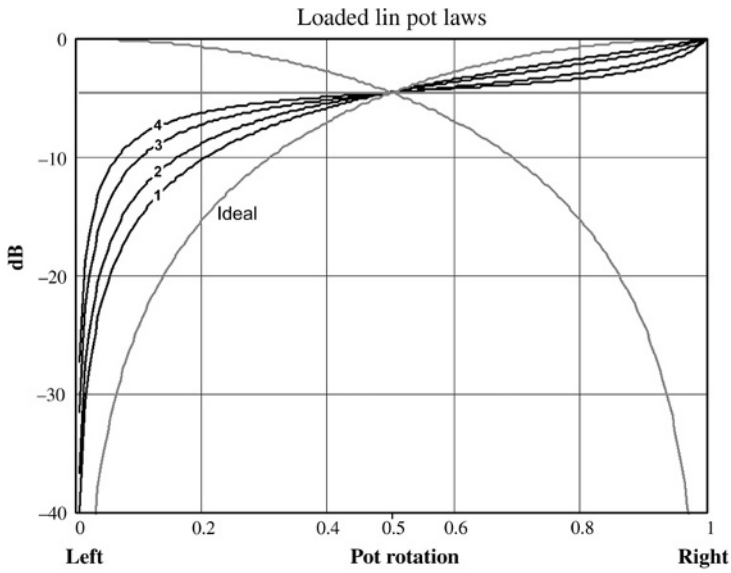


Figure 17.8: How the law of a panpot varies with loading on the wipers; in each case the law-bending resistors have been adjusted to give −4.5 dB at the center (see Table 17.1)

to the law-bending pull-up resistors, which therefore have to be reduced in value (here to 3k3) to get the mid-point attenuation back to −4.5 dB. As the wiper loading increases and the value of the pull-up resistors is decreased with respect to the panpot track, the more the resulting law tends to flatten out around the central position, as shown in Figure 17.8. Table 17.1 shows how the law-bend resistors change with loading when a 10 kΩ pot is used, and the resulting traces are shown in Figures 17.8 and 17.9.

TABLE 17.1 Wiper loading and pull-up resistor values in Figures 17.8 and 17.9

Trace in Figures 17.8 and 17.9	Wiper loading	Pull-up resistor
1	15k	5k
2	4k	2.2k
3	3k	1.8k
4	2k	1.2k

The distorted law that results from excessive loading gives a panpot that does little as it is moved through the central position, and has an unpleasant 'dead' feeling; there are also level variation problems, as previously described. In addition, the lower the pull-up resistors, the greater loading they place on the stage upstream, which is usually the fader post amp; the loading is at its worst when the panpot is hard over to either side. This way of using a panpot presents a compromise on top of a compromise, and puts a limit on the number of mix resistors that can be driven directly from a panpot.

www.raguvaran.puzl.com

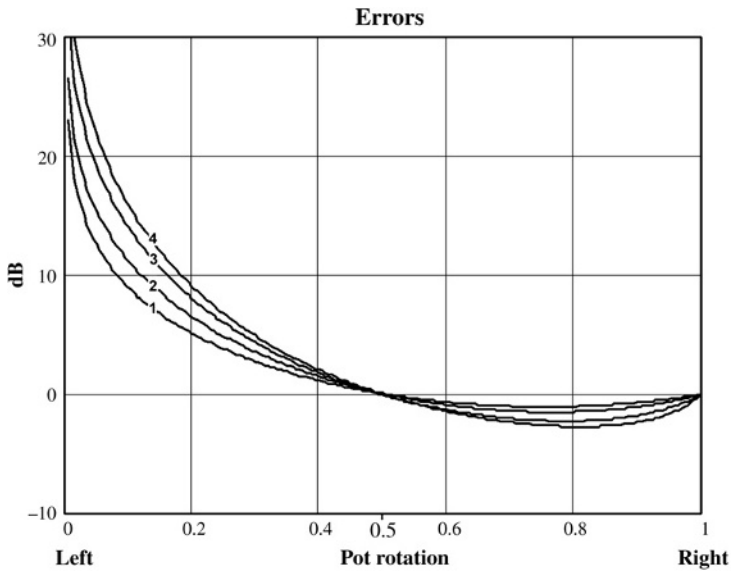


Figure 17.9: The difference between the ideal compromise law and the actual law that results from law bending (see Table 17.1)

Figure 17.9 shows the differences between the ideal compromise law and the actual law that results from law bending, for various wiper loadings as in Table 17.1. The amplitude error of the Right output gets enormous as the loading increases; when the panpot begins to move away from hard left the level shoots up much too quickly.

Since the value of the pull-up resistor required to get the best approach to the required law depends on the loading on the panpot wiper, this is a very good reason to use routing methods that place a constant load on the panpot, regardless of the number of buses that are being routed to, as this at least allows the value of the law-bending resistors to be optimized at one value. More on this later.

A more sophisticated approach is used in high-cost mixers, which puts unity-gain (or near-unity-gain) buffers between the panpot and the loading of the routing resistors, as shown in Figure 17.7(b). Here there are two points to note: the panpot wipers now have a negligible loading to ground, and the pull-up resistors therefore have a much higher value of 10 kΩ, which will improve offness as less current goes through the wiper contact resistance. Also, the routing resistor values can be reduced, as they no longer load the panpot, and this reduces the Johnson noise generated in the summing system.

This law-bending technique is a workable solution, and has been very widely used, but unfortunately the addition of the law-bending resistor introduces another problem.
A reasonable quality pot has an offness of about −90 dB with reference to fully up, due to the end-of-track resistance. The pull-up current from the law-bend resistors, however,

www.raguvaran.puzl.com

passes through the wiper contact resistance, which is usually greater than the end-of-track resistance, and this extra resistance severely limits the attenuation the panpot can provide when set hard left or right, degrading the offness of the panpot from approximately −90 to −65 dB when it is hard over. The problem is made worse because when the wiper is at the bottom of the track, the whole of the signal voltage across the pot is also across the law-bend resistor, which is lower in resistance than the pot track and so passes more current through the wiper contact resistance. The exact value of offness obtained depends on the component values and the construction of the pot. The way this works is shown in Figure 17.10.

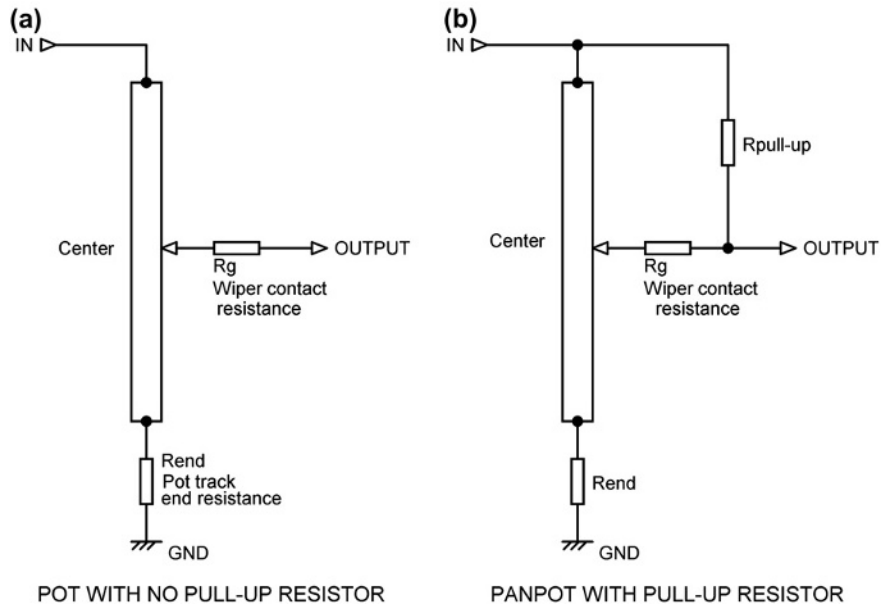


Figure 17.10: How the law-bending resistors degrade the offness of a panpot

This reduction in offness is not really a problem in stereo use, as an attenuation of −65 dB is more than enough to pan a sound so it is subjectively completely to one side (in fact about −20 dB will do that), but it is a very serious issue for mixers that route to groups in pairs, or in other words have one switch to route to groups 1 and 2, another switch for groups 3 and 4, and so on. Simpler mixers route to groups in pairs, as this makes good use of two-changeover switches (and, as mentioned before, the channels are by a long way the most cost-sensitive part of a mixer) but it means that if you are, say, routing to group 1 by pressing the 1–2 switch and then panning hard left, the signal being sent to group 2 will only be −65 dB down. In bigger and more expensive consoles each group has its own routing switch, with the spare switch section typically used for an indicator LED, so panpot offness is not such an important issue.

www.raguvaran.puzl.com

The Active Panpot

I think that it has been made clear that a conventional panpot is a bundle of compromises, and I decided in 1989 to do something about it. It was one of those ideas that strikes with such force that you remember exactly where and when it happened: in this case on the fifth floor of a tower block in Walthamstow.

The active panpot shown in Figure 17.11 works by replacing the simple pull-up resistor with a negative-impedance converter that modulates the law-bending effect in accordance with the panpot setting, making a closer approach to the sine law possible. This sounds intimidating but is actually very simple. The wiper of the left half of the panpot is connected to a series-feedback amplifier U1:A that gives a modest gain of about 2 dB, set by R2 and R3; the exact value shown in Figure 17.11 is 1.73 dB. The law-bending resistor R1 is driven from the output of this amplifier rather than the top of the panpot and so, when the panpot wiper is at the lower end of its travel, there is little or no signal to the amplifier, and therefore no pull-up action when it is not required. There is therefore no signal current flowing through the wiper contact resistance, so the left–right isolation using a good-quality pot is greatly improved from approximately −65 to −90 dB, being now limited solely by the end-of-track resistance as

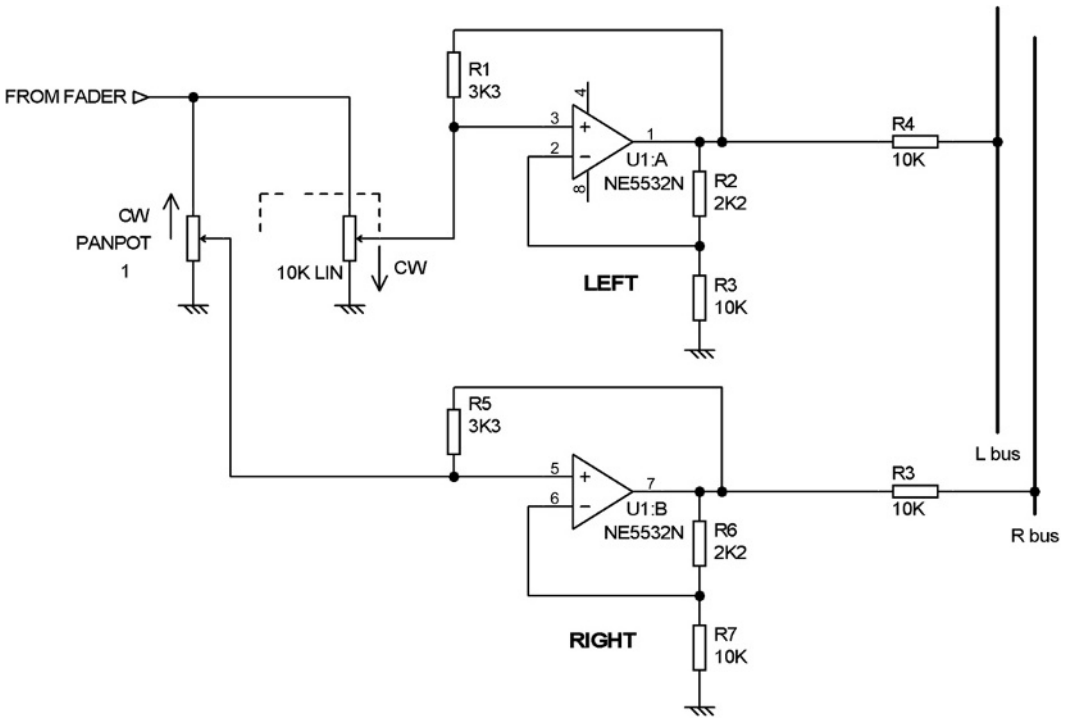


Figure 17.11: The active panpot: the law-bending resistor is driven from an amplifier stage with a gain of about 2 dB

www.raguvaran.puzl.com

shown earlier in Figure 17.10(a). As an extra benefit, if the amplifier gain and the law-bend resistor are correctly chosen, the pan law is much closer to the desired sin/sin^2 compromise law than can be obtained by the simple law-bending method described above; this is illustrated in Figure 17.12. This concept was protected by patent number GB 2214372 in 1989.

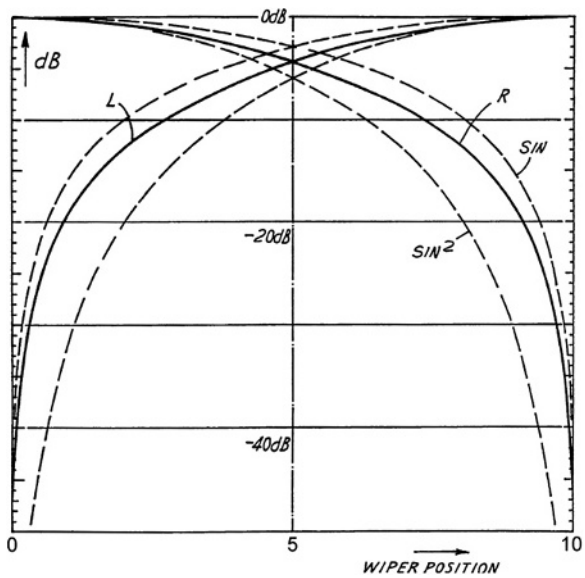


Figure 17.12: The active panpot: the resulting law is much closer to the desired sin/sin^2 compromise

The active panpot is obviously a more expensive solution, as we have added a dual op-amp and four resistors to every channel. However, we get other benefits for our money as well as better pan offness and a better pan law; if the dual op-amp is a 5532 (which it should be to keep the postfade noise down), then its good drive capability means that the value of the routing resistors can be much reduced, which reduces both capacitive crosstalk and noise in the summing system.

LCR Panpots

A relatively recent innovation in the world of panpots is the LCR panpot, where LCR stands for Left–Center–Right. This has three outputs: the Left and Right in general working the same way as for a normal panpot. As the panpot is moved from extreme left, the Center output rises from zero, reaches a maximum with the panpot central, and then falls to zero again as the panpot reaches extreme right. The LCR panpot is used with speaker systems that have a central cluster as well as left and right speaker banks. This means that centrally panned signals are always heard in the center of the sound field, regardless of listening position, and this is useful when vocals or speech are combined with music, as typically occurs in religious installations.

www.raguvaran.puzl.com

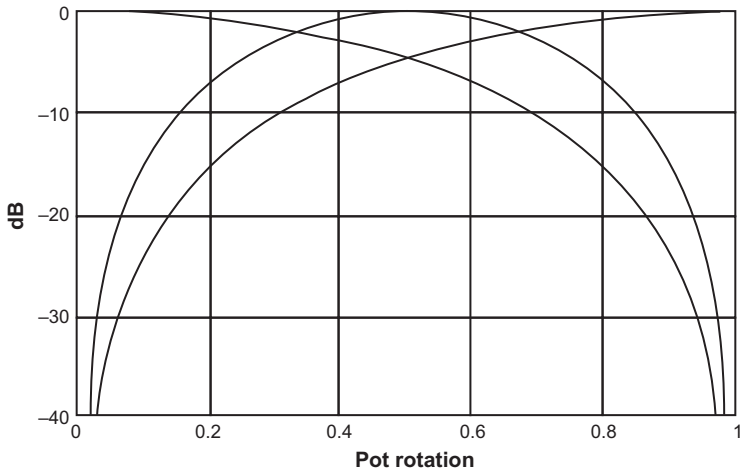


Figure 17.13: A full-width LCR panpot law. Left and Right cross over at −4.5 dB in the center as before, while C is 0 dB at the center, but zero at each extreme

LCR operation is also used for panning to a discrete center channel to create Surround and 5.1 mixes.

There are two possible ways to handle LCR panning. In what I call the 'full-width' method, as shown in Figure 17.13, the Left and Right panpot outputs work much as they do in a normal stereo panpot, only falling to zero at the extreme end of control rotation.

The alternative is the 'half-width' method, in which both Left and Right outputs fall to zero at the center setting, and stay there for further control rotation; this is shown in Figure 17.14. Both figures show laws which are a sin/sin^2 compromise as described earlier, and are therefore 4.5 dB down at the center – in the case of the half-width method these 'centers' are at pot positions of 25% and 75% rotation. A switch is normally provided by which the LCR pot can be converted to conventional L/R stereo operation. Which LCR panning method is most appropriate depends on the application; some LCR panpot systems can switch between the full-width and half-width modes.

Figure 17.15 shows a full-width LCR panpot arrangement. The L and R sections are as for a normal L–R panpot, but the C pot track is grounded at each end and fed with signal via its center-tap; an LCR panpot always requires specially designed pots with center-taps on the tracks. These are available from major manufacturers like ALPS but they are inevitably made in small quantities and are relatively expensive.

Figure 17.16 shows a half-width LCR panpot. The L and R sections now are also center-tapped, and that tap is grounded so that the output falls to zero at the central control setting. The C pot track is grounded at both ends and fed with signal via its center-tap as before.

www.raguvaran.puzl.com

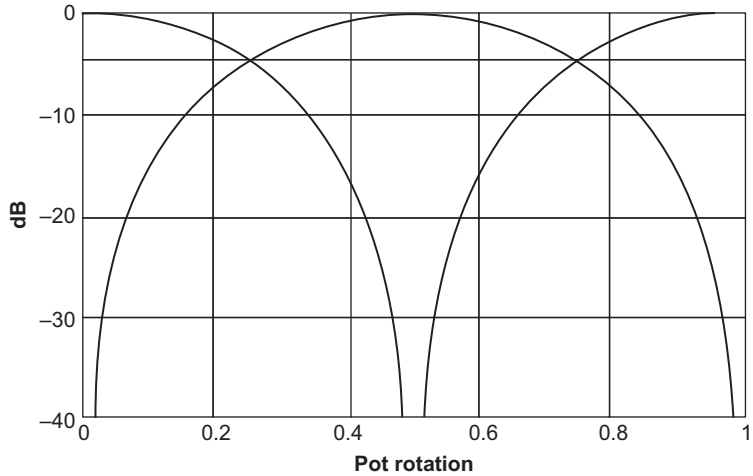


Figure 17.14: A half-width LCR panpot law. Left and Right fall to zero at the center setting and stay there for further control rotation. The crossover points between Left and Center and Right and Center are now at −4.5 dB

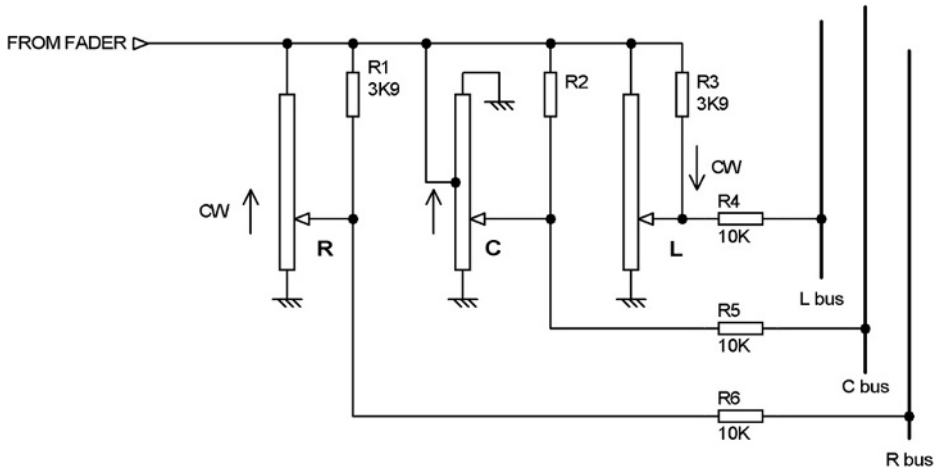


Figure 17.15: A full-width LCR panpot system

LCR panpots, like ordinary panpots, give much superior performance in terms of offness and a better pan law when the active panpot system is applied to them. This is straightforward to do, and is arranged exactly as for the stereo version.

Consoles with LCR panning systems include the Soundcraft Series 5, the Yamaha M2500-56, the D&R Orion, and the Amek 9098i. True LCR panning is not common as it is a relatively high-cost facility.

www.raguvaran.puzl.com

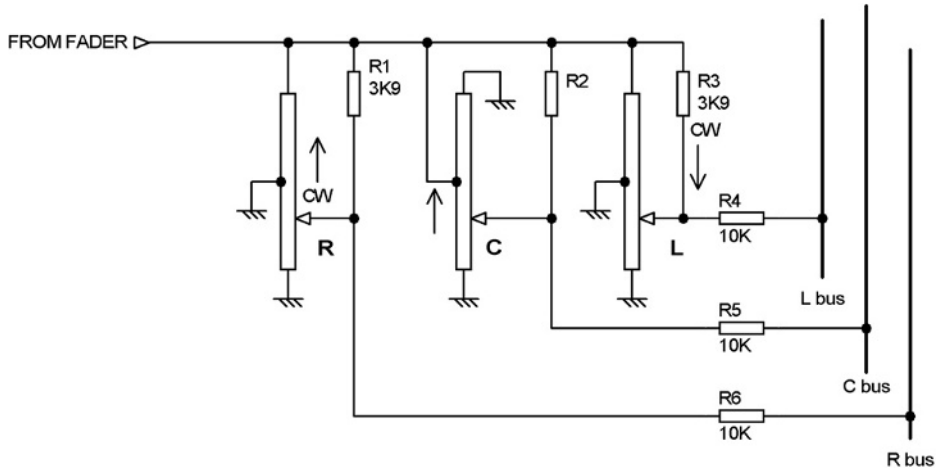


Figure 17.16: A half-width LCR panpot system

Routing Systems

The routing system selects which mix bus the channel signal goes to. In a simple *N* into 2 stereo mixer the panpot does all the routing that is required, but as soon as more groups are introduced some sort of switching is required for bus selection.

Figure 17.17(a) shows the conventional routing system as commonly used up until about 1980. There are two problems with this method. Firstly, the capacitance between the switch contacts when they are open is significant, and this severely limits the crosstalk performance of the console. To get a feel for the problem, look at Figure 17.18, which shows the crosstalk performance of an ALPS SPUN two-changeover switch working into a mix bus with a 10 kΩ feedback resistor in the summing stage. This is a conventional push-switch with two parallel sections.

At 10 kHz the offness is only −66 dB, and at 20 kHz it is barely 60 dB. The problem is capacitative crosstalk across the switch contacts in the off position. Note that grounding the second switch section in the pious hope of improving things only reduces the crosstalk by 2 dB.

This problem can be completely eliminated by using the arrangement in Figure 17.17(b), which is sometimes called the back-grounding system. The mix resistors are now connected to the channel ground when not in use, and so there is no possibility of capacitative crosstalk, so long as the switches, of which one contact still carries signal voltage, are not too close to the mix buses. But this plan has a terrible drawback. The crosstalk performance is good, but back-grounding the mix resistors means that since relatively few routing switches are on at a time, most of the resistors are grounded and the summing amp is working at or near maximum noise gain, and always picks up the maximum possible ground noise.

www.raguvaran.puzl.com

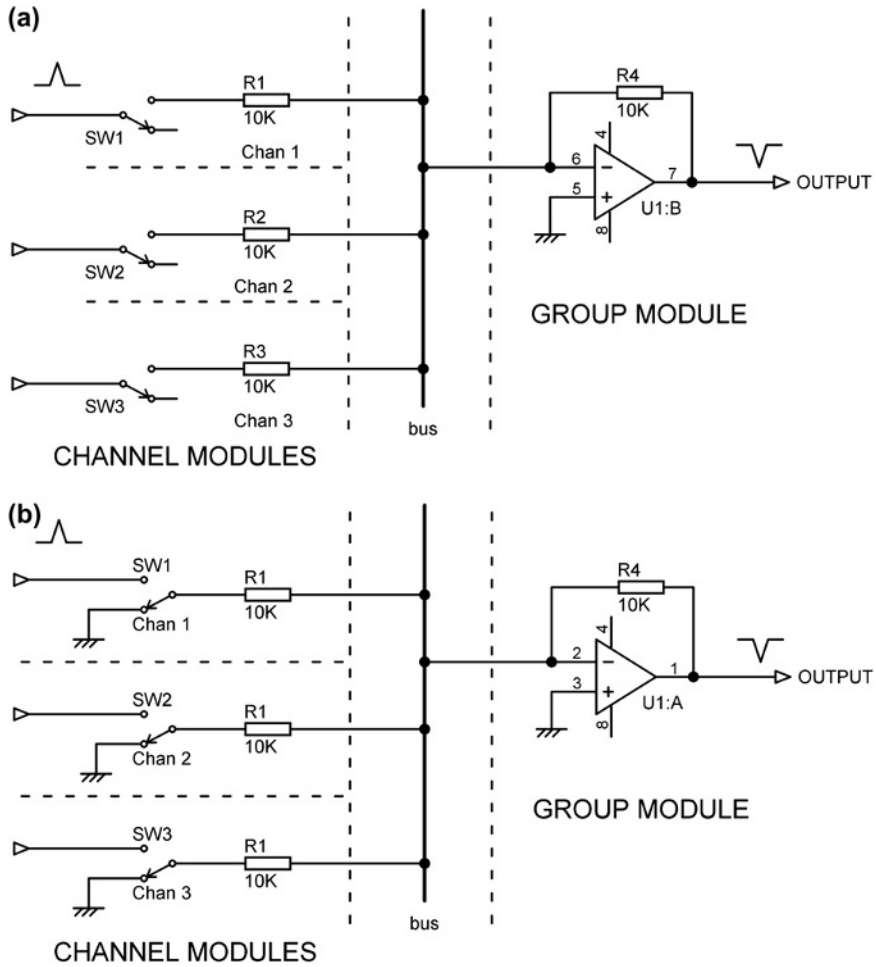


Figure 17.17: Two mediocre routing systems with virtual-earth summing, as used in the 1970s. (a) suffers from poor offness and (b) has bad noise performance

It might be wise to take a moment here to explain the noise gain of a summing system, though this material really belongs in the next section on summing systems. A simple inverting amplifier with equal input and feedback resistors has a noise gain of 2, or 6 dB, because the noise referred to the amplifier input sees effectively a non-inverting amplifier with a gain of 2, and so the noise at the output is twice that of the noise-generating mechanisms at the amplifier input. The important point is that the noise gain is greater than the signal gain, which is unity.

If a second mix resistor of equal value is connected to the summing point, the amplifier input now sees effectively a non-inverting amplifier with a gain of 3, and the noise gain is increased

www.raguvaran.puzl.com

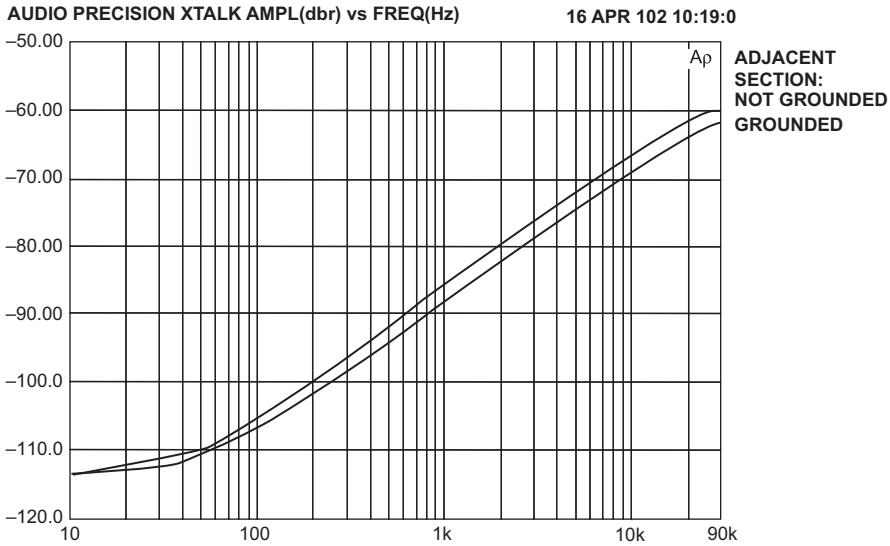


Figure 17.18: The poor high-frequency offness of a switch in a conventional routing system with a 10 kΩ summing feedback resistance

to almost 10 dB, though the signal gain via either mix resistor remains unity. The more channels that are routed to a mix bus, the worse the noise performance is, as summarized in Table 17.2. It is therefore clear that the arrangement in Figure 17.17(b) will be working at the worst-case noise gain all the time. Not only will the noise of the summing amp receive maximum amplification, but any ground noise in the console that puts the channel grounds at

TABLE 17.2 Noise gain of a summing amp versus the number of mix resistors connected

No. of mix resistors	Noise gain (×)	Noise gain (dB)
1	2	6.02
2	3	9.54
4	5	13.98
8	9	19.08
12	13	22.28
16	17	24.61
24	25	27.96
32	33	30.37
48	49	33.80
56	57	35.12
64	65	36.26

www.raguvaran.puzl.com

slightly different potentials will be picked up as effectively as possible. This is not a good plan, especially in large consoles.

Obviously a routing system that combines the advantages of the first two systems – good offness and minimum noise gain – would be a great improvement.

So, finally, here is the most satisfactory routing system. In the arrangement shown in Figure 17.19, the routing resistors are once more grounded when not being fed by the channel, but the topology has been turned around so that the grounded resistors are now connected to the channel rather than the mix bus. The switches rather than the routing resistors are connected directly to the summing buses, and when a routing switch is not active, the feed from the routing resistor is grounded. So long as the hot end of the mix resistor can be kept away from the mix buses, the offness is truly excellent. With a bit of care in the physical construction it can easily be better than 100 dB at 20 kHz. A possible objection is that all the grounded resistors put a heavier load on the channel circuitry upstream; this is quite true, but it has the countervailing advantage that since each mix resistor is either grounded directly or connected to a virtual-earth mix bus, the loading remains constant, and so if, for example, the resistors are driven from an unbuffered panpot, you will get a predictable panning law, if not necessarily a very good one. This routing method retains the advantage of working the summing system at the minimum noise gain.

I have a difficulty here. To the best of my knowledge I invented this system in the backroom of a shop in Leyton, London in 1979. I recall being nervous about the apparently iffy business of having the routing switch at virtual earth, because of increased stray capacitance to the virtual-earth bus, but in fact that caused no problems. However, since very little has been published

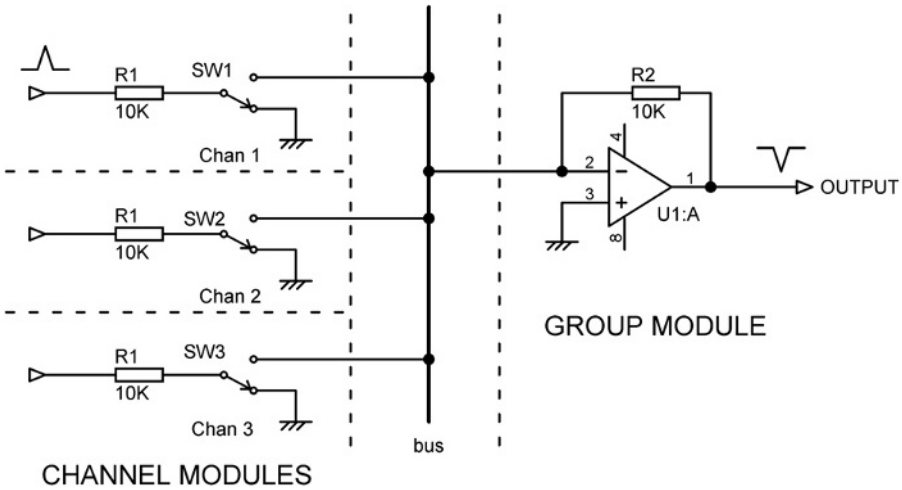


Figure 17.19: The routing system that I introduced in 1979, which gives both good routing offness and optimal noise performance. Virtual-earth summing system is also shown

www.raguvaran.puzl.com

about the details of mixer design, it being a proprietary art confined to a relatively small number of companies, it is not currently possible to say if I was the first to invent it.

The set of routing switches in each channel (often called the routing matrix) is relatively expensive and is of course multiplied by the number of channels. It is therefore common on all but the most expensive consoles for routing to be done in pairs; in other words the first switch routes the channel to groups 1 and 2, the second to groups 3 and 4, and so on. This is simple to do as many varieties of push-switch come in a two-changeover format as standard. If routing to only one group is required then the signal has to be steered to the appropriate bus with the panpot. The offness of the panpot when set hard over is then crucial to obtaining good crosstalk figures; the active panpot scheme described above greatly improves this aspect of performance.

Auxiliary Sends

These are straightforward. There is often a pre/post switch before the send pot so the send signal can be taken from before or after the fader. In many designs added flexibility is given in the form of push-on links so that other take-off points for the send signal can be used, such as before the EQ section. Changing all these links on a large console is naturally not a light matter.

If there is an on/off switch for the send then it should disconnect the mix resistor from the bus, just as with the group routing switches described above. This reduces summing noise dramatically if only a few auxes are in use, and gives an offness much better than that at the end of the aux pot travel.

Group Module Circuit Blocks

Most of the circuit blocks used in group modules carry out the same functions as they do in the input channel modules. The great exception is of course the summing amplifier, which has a technology all of its own. Before diving into the details of practical summing amplifiers, it is advantageous to look at the various methods of performing that apparently simple but actually rather demanding task – adding signals together. The most fundamental function of a mixer, as its name suggests, is to combine two or more signals in the desired proportion. As with many areas of electronics, an extremely simple definition of the job to be done (addition – how hard can that be?) rapidly shows itself to have ever-deeper levels of complexity.

There are several interesting techniques to look at:

- voltage summing
- virtual-earth summing

www.raguvaran.puzl.com

- balanced summing
- ground-cancel summing
- distributed summing.

Summing Systems

Voltage Summing

The earliest mixers used voltage summing, or passive summing as it sometimes called, as shown in Figure 17.20; note that only half of each panpot is shown, for clarity. The great drawback with this system is that there is a significant voltage on the mix buses, so that a signal can be fed on to the bus by one channel, and it can get back into another channel, as shown in the figure, from where it may, depending on the control settings, find its way on to another mix bus where it is definitely not wanted. This is demonstrated by the arrow A in the diagram. Channel 1 is feeding a signal on to group 3 bus, and this is sidling back into channel 2, which has both routing switches engaged and the panpot central, allowing the signal to turn round and get on to group 1 bus. This sort of thing can of course be completely suppressed by suitable

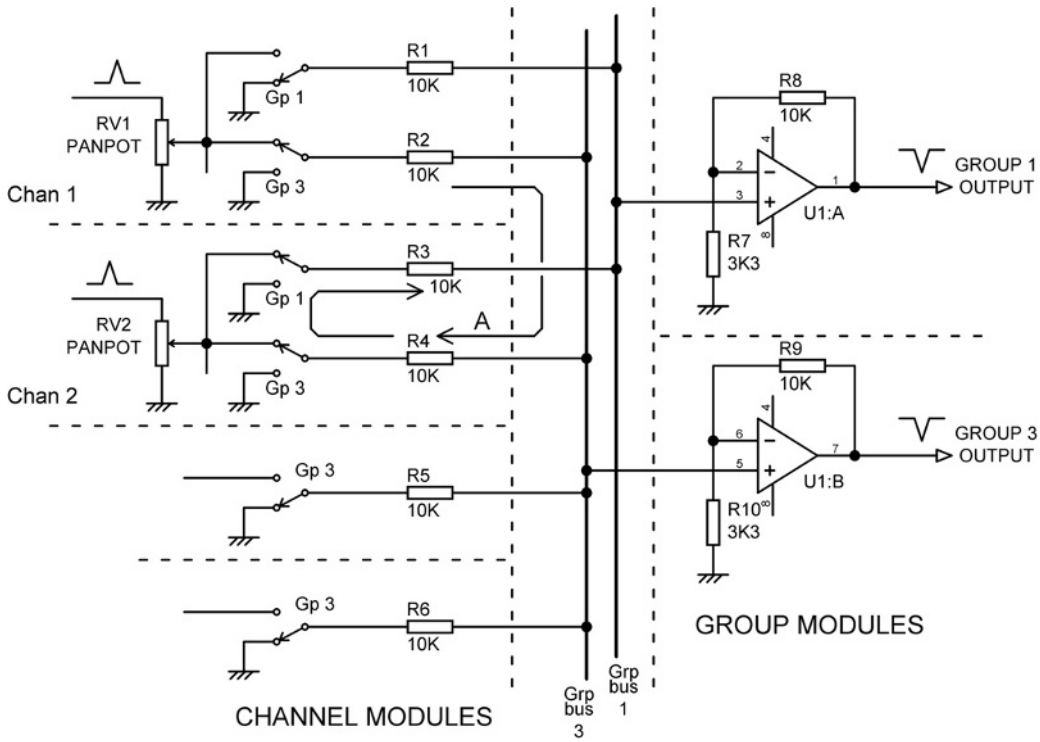


Figure 17.20: Voltage summing system as used in the 1960s

www.raguvaran.puzl.com

buffer stages after the panpot, and that is exactly what was done in the large consoles of the day. This means a lot of extra electronics, and this mixing system is not suitable for low- or medium-cost designs. Another big snag is that since the mix buses have significant voltages on them, they must be carefully and expensively screened from each other to prevent capacitative crosstalk. Running the buses in a piece of low-cost ribbon cable is simply not an option.

You will have noted that all the routing switches are back-grounded, so there is a constant impedance from the mix bus to ground; this is essential as otherwise the gain of the summing system would vary with the number of channels routed to a bus. This in turn means that the amplifiers that get the low level on the bus back up to the nominal internal level are working at substantial gain and are relatively noisy.

Virtual-Earth Summing

Summing today is done almost universally by virtual-earth techniques, as shown above in Figure 17.19. An op-amp or equivalent amplifier with shunt feedback is used to hold a long mixing bus at apparent ground, generating a sort of audio black hole; signals fed into this via mixing resistors apparently vanish, only to reappear at the output of the summing amplifier, as they have been summed in the form of current. The great elegance of virtual-earth mixing, as opposed to the voltage-mode summing technique in Figure 17.20, is that signals cannot be fed back out of the bus to unwanted places, as it is effectively grounded, and this can save massive numbers of buffer amplifiers in the input channels. The near-zero bus impedance also means that the gain does not alter as varying numbers of channels are routed to a bus. The appearance of the virtual-earth mixing concept was highly significant, as it essentially made low-cost mixing consoles viable.

There is, however, danger in being dazzled by this elegance into assuming that a virtual earth is a perfect earth – it is not. A typical op-amp-based summing amp loses open-loop gain as frequency increases, making the inverting input null less effective. The ‘bus residual’ (i.e. the voltage measurable on the summing bus) therefore increases with frequency and can cause inter-bus crosstalk in the classic situation with adjacent buses running down an IDC cable.

As we saw in the last section on routing, a virtual-earth summing system operates at quite a high noise gain when many inputs are routed to it. Minimizing the noise is therefore of prime importance, and one obvious step is to keep the impedance of the summing system as low as possible to reduce Johnson noise in the mix resistors. I have been involved in more mixing console designs than I care to contemplate, of all sizes and types, and it is interesting to recall that the mix resistors were always in the range 22 to 4.7 kΩ, which seems like a rather narrow range of 4.7:1. There is a reason for this: 22 kΩ mix resistors are used in small-budget consoles because a small number of them (say six) can be driven directly from a panpot without distorting its law unacceptably. Now consider a big expensive 32-group console where the panpot is buffered, but it has to drive 16 mix resistors on each side with the buffer, and making

www.raguvaran.puzl.com

them of very low resistance would present enormous loading. The mix resistors will therefore be something like 6.8 kΩ; the value has only been reduced by a factor of about 3, and Johnson noise will only be reduced by the square root of this, or 5.3 dB.

Mix resistors lower than this are only used for the critical L–R mix buses, where there are only two to be driven, and 4.7 kΩ is a common value. It actually only gives an advantage of 1.6 dB over 6.8 kΩ. The lowest mixing resistors I am aware of are in fact 4.7 kΩ, and this value was used in the Neve VR consoles, which in their largest formats had a possible 144 inputs going to the L–R mix bus (72 modules each with two paths to the mix bus).

This consideration of resistance demonstrates the difficulty of reducing summing Johnson noise beyond a certain point by reducing the bus impedance. Halving the mix resistors doubles the drive power required, but only improves the noise by a factor of the square root of 2. Johnson noise is of course only one factor; there is also the noise from the summing amplifier itself. The current-noise component of this is also reduced by reducing the bus impedance.

Apart from the noise performance, another advantage of a low summing system impedance is that it makes the mix buses less sensitive to capacitive crosstalk. Reducing mix resistors from 22 to 6.8 kΩ may only give a 5.3 dB improvement in noise, but it gives a 10 dB reduction in capacitative crosstalk.

Balanced Summing Systems

As a console grows larger, the mix bus system becomes more extensive, and therefore more liable to pick up internal capacitative crosstalk or external magnetic fields, which are poorly screened by the average piece of sheet steel. The increased physical size means longer ground paths with non-negligible signal voltage drops across them. An expensive but thoroughly effective answer to this is the use of a balanced summing system. The basic idea is shown in Figure 17.21; only half of the channel panpot is shown, for clarity.

Each mix bus is now double, having a hot (in-phase) and cold (out-of-phase) bus, exactly like the hot and cold wires in a balanced line connection. These run physically as close together as possible. At the channel end, each side of the panpot has an inverting buffer A2 that drives a second set of mix resistors that feed the cold mix buses. Two sections are now required for the routing switch SW1, which is fine as most push-switches are in a two-changeover format, but if you want switch status indication, which you normally will on a console large and sophisticated enough to employ balanced summing, four-changeover switches are needed – lots of them – and the cost jumps up. Each bus connects to a summing amplifier A3, A4 and their outputs are subtracted by A5 to cancel the common-mode signals.

A quite separate advantage of a balanced summing system is that it has a better noise performance, because having two buses gives you 6 dB more signal, but only 3 dB more noise because the noise from the two summing amplifiers is uncorrelated and partially cancels. This

www.raguvaran.puzl.com

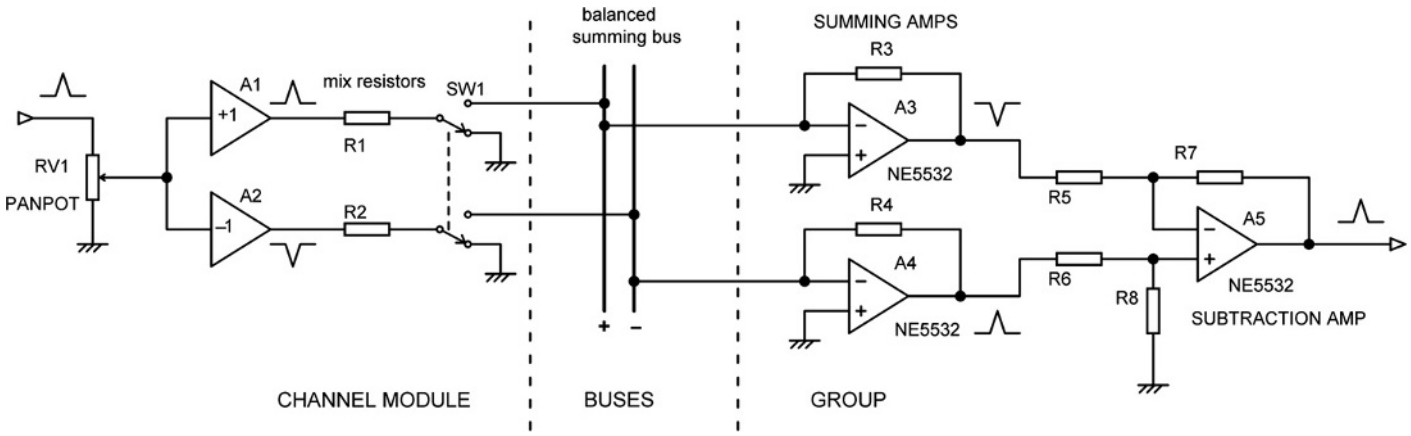


Figure 17.21: A balanced summing system

www.raguvaran.puzl.com

gives you a 3 dB noise advantage that it would be difficult to get by other methods. Against this must be counted the extra noise from the channel inverters.

Because of the cancellation of capacitative crosstalk that occurs, the offness of the routing matrix can be made very good indeed. On the last console of this type that I designed the routing offness was a barely measurable 120 dB at 20 kHz.

As I just noted, probably all consoles of this type will have routing switch status indication, and so use four-changeover switches. This leaves a spare switch section doing nothing, and if you have followed this book so far you will have gathered it is not in my nature to leave a situation like that alone. In the last balanced console I designed, this fourth switch section was used to detect when *no* routing switches were active; this happens more often than you might think on a large desk. In this condition an FET switcher removed the signal feed to the buffers driving the mix resistors. This prevented a large number of grounded resistors being pointlessly driven, reducing power consumption significantly, and further improving routing offness.

It may have occurred to you that a considerable simplification would result if the channel inverters were done away with, and the cold resistors simply connected to ground. The signal level in the summing system would be 6 dB lower, exacting a 3 dB noise penalty, but the rejection of ground voltages and the cancellation of capacitative crosstalk would be just as good. This is not a sound plan. Much of the cost of balanced summing lies in the doubled number of mix resistors, routing switch contacts, and mix buses. The saving on inverter cost is relatively small. The large currents flowing in the mix resistors are now unbalanced and are much more likely to cause troublesome voltage drops in ground connections.

A variation on this approach can, however, be very useful in specific applications (see the section on ground-cancel summing systems below).

Ground-Canceling Summing Systems

Ground-canceling summing systems are very useful in auxiliary send systems where all sends are routed permanently to the bus. It gives some of the advantages of a balanced summing system at a tiny fraction of the cost, and it is particularly useful in consoles where the modules are connected together by ribbon cabling with a relatively high ground resistance. The basic principle is shown in Figure 17.22, which represents a single channel with two aux sends. The ground potential of each channel is read by a low-value resistor R3, and summed into the GC bus; this is not a virtual-earth bus as such, but it is at a very low impedance and is essentially immune to capacitative crosstalk. The GC bus is then used as the ground reference for the aux summing system, and by subtraction it gives excellent rejection of the channel ground potentials. If a channel is at the remote end of the ribbon cable then its ground resistance is relatively high and the local channel ground will carry a large signal potential, which is fed

www.raguvaran.puzl.com

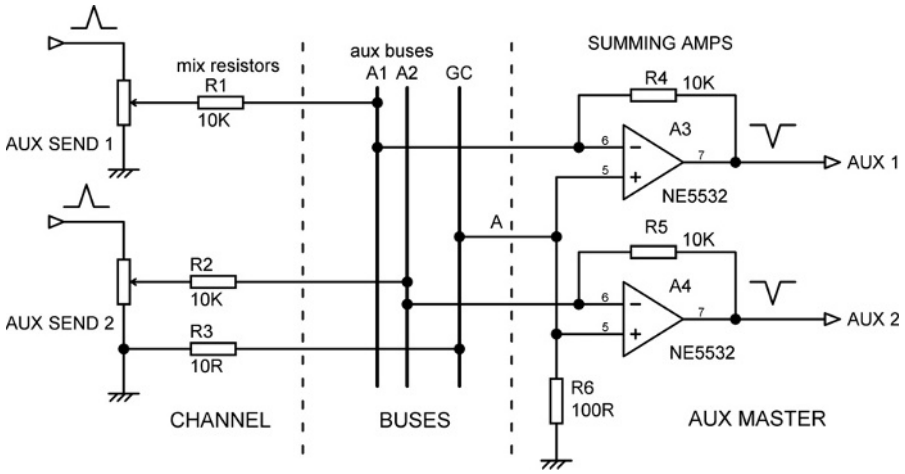


Figure 17.22: A typical ground-canceling summing system for aux sends

into the GC bus; conversely, a channel at the near end of the ribbon cable will have a much lower ground potential, also fed into the GC bus. The end result is that good aux offness figures are obtained for both those channels. In such a system the aux offness is defined not by the grounding arrangements but by the offness of the aux send pot itself, typically −90 dB with reference to fully up.

A couple of design points. The GC resistors R3 should be very low in value so that their Johnson noise contribution is negligible; the only requirement is that they are large with respect to the ground resistances to allow accurate subtraction. R6 is a safety resistor fitted in the aux master module so that it will keep working if the GC bus becomes disconnected. Without it the entire aux send system is dependent on the single connection A between the bus structure and the aux master module, and such a situation is not good practice. The only requirement is that R6 is high with respect to the total GC bus impedance so that the subtraction remains accurate. If an aux send has an on/off switch that disconnects the mix resistor from the bus, to minimize noise gain, the switch must also disconnect the GC resistor to maintain correct cancellation.

In some ribbon-cable consoles the grounding is reinforced by heavy (32/02) ground wires with push-on connectors plugged into every channel module, or in some cases every fourth or fifth module. Even in this case the use of a ground-canceling summing system will normally give a worthwhile improvement.

Distributed Summing Systems

Distributed summing (sometimes called devolved summing or devolved mixing) is a deeply cunning way of improving the noise performance of a summing system, with reference to

www.raguvaran.puzl.com

the noise from the summing amps themselves. In what follows the noise from an input channel and Johnson noise from the summing resistors are ignored for the time being.

In a distributed summing system the contributory signals are summed in two separate stages. Thus if you are summing together 24 channels, you might sum them in blocks of eight, to get three subgroups that are then summed together to get a single output, as is shown in Figure 17.23(a). Unlikely as it may seem, this two-layer summing system gives a definite noise advantage. Assume that every summing amplifier creates the same amount of input-referred noise V_n; the result of this at the amplifier output then depends only the noise gain at which it is working. If we sum our 24 channels into one in the normal virtual-earth manner, the output noise will be $25V_n$, and that is our reference case. If we sum eight inputs together, the output noise will be only $9V_n$; three of these sub-mix outputs are now summed together, and you may have seen this coming – the signals sum arithmetically but the noise partially cancels, so the combined noise output is $9V_n$ times $\sqrt{3}$, which is $15.6V_n$. To this we must add the noise of our second summing amp; this is working at a noise gain of only 4, and so its own output noise contribution is $4V_n$. Adding the uncorrelated $15.6V_n$ to $4V_n$ rms fashion gives a total of $16.1V_n$, which is 1.55 times or 3.83 dB quieter than the straightforward all-in-one-go summation of 24 inputs into one. Clever, eh?

The downside to this is of course that a bit more hardware is required – we are using four summing amplifiers rather than one. This is, however, a trivial extra expense compared with

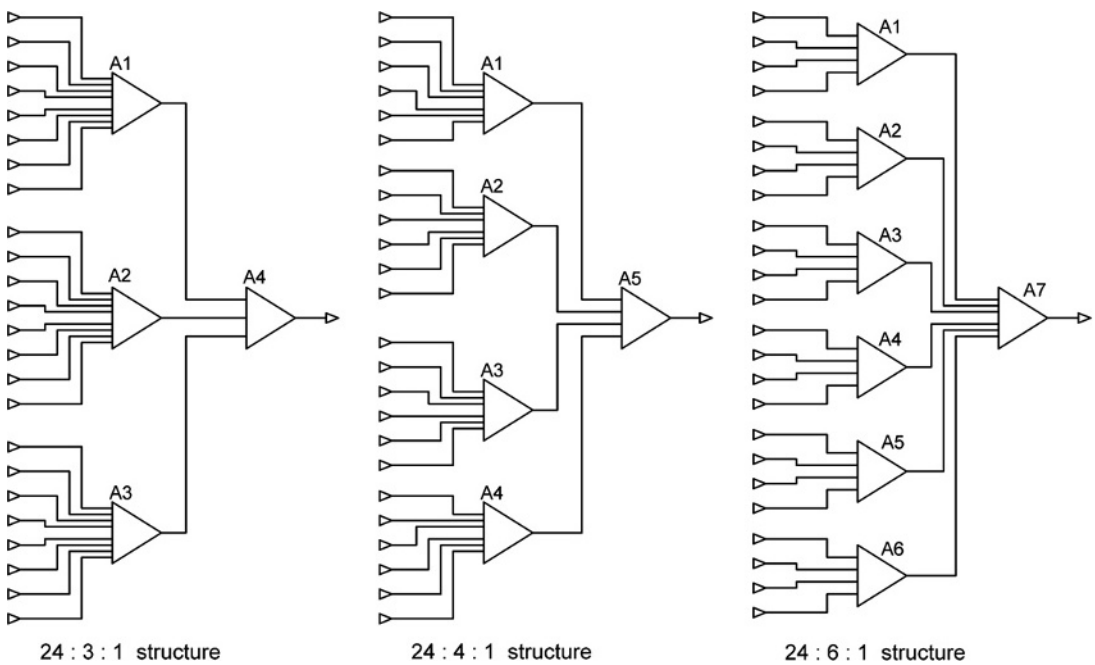


Figure 17.23: Three different distributed summing schemes for 24 inputs; 24:6:1 is the quietest

www.raguvaran.puzl.com

the electronics involved in 24 input channels. A more serious potential problem with this approach is that there is now a hidden layer of summing amplifiers, which might clip without there being any indication on the console control surface.

You are by now, I hope, pondering if there is anything special about the 24 into 3 into 1 (24:3:1) structure, or whether other variations might be better – the answer is yes. Three different possibilities are shown in Figure 17.23. Summing the 24 inputs in batches of six into four sub-mixes and then summing them to one (24:4:1) gives an advantage of 4.51 dB, while summing four at a time into six sub-mixes and then summing to one (24:6:1) gives an advantage of 4.97 dB, and this is the optimal solution for 24 inputs.

The results are summarized in Table 17.3, which shows that there is not much difference between 24:6:1 and 24:8:1, except that the latter uses a bit more hardware and is a negligible 0.2 dB noisier. It is hard to see why anyone would want to use the 24:12:1 structure, because of all those extra summing amps, but it is interesting to note that it still gives a theoretical noise improvement of 3.53 dB.

Table 17.4 illustrates the noise advantages with various forms of two-layer summing for 32 inputs; the maximum advantage is now a very useful 5.88 dB, and once again the best results come from first summing the inputs in batches of four. There are fewer options this time because 32 has fewer factors than 24.

TABLE 17.3 The noise advantage of two-layer distributed summing with 24 inputs

No. of inputs	No. of sub-mixes	Noise advantage (dB)
24	2	2.56
24	3	3.83
24	4	4.51
24	6	4.97
24	8	4.76
24	12	3.53

TABLE 17.4 The noise advantage of two-layer distributed summing with 32 inputs

No. of inputs	No. of sub-mixes	Noise advantage (dB)
32	2	2.68
32	4	4.94
32	8	5.88
32	16	4.01

www.raguvaran.puzl.com

Table 17.5 shows the optimal structures giving minimum noise for various numbers of inputs; at first it looks as if there is something special about doing the first layer of summing in batches of four inputs, but as the total number of inputs reaches 72 the optimal number suddenly starts to rise.

TABLE 17.5 The optimal two-layer summing structures for various numbers of inputs

Total no. of inputs	No. of inputs in each sub-mix	Optimal structure	Noise advantage (dB)
16	4	16:4:1	3.64
24	4	24:6:1	4.97
32	4	32:8:1	5.88
48	4	48:12:1	7.09
56	4	56:14:1	7.52
64	4	64:16:4	7.88
72	6	72:12:1	8.48
80	8	80:10:1	8.48
88	8	88:11:1	8.84
96	6	96:16:1	9.43
104	8	104:13:1	9.46
112	8	112:14:1	9.73
120	6	120:20:1	10.13
128	8	128:16:1	10.21
136	8	136:17:1	10.43
144	6	144:24:1	10.67

The figures in Table 17.5 do show a rather unsettling irregularity; for example, the optimal structure for 80 inputs is no quieter than that for 72 inputs. This seems to be because some numbers of inputs have more factors than others, and so give a greater choice of structures to choose from.

Two-layer distributed summing is particularly adapted to consoles built in sections, typically as bins of eight or 12 modules. It avoids virtual-earth buses extending over large physical areas, which makes them more susceptible to magnetic fields and ground voltage drops, and can cause difficulties with HF stability. Each bin can now be connected by balanced low-impedance line connections, with the final summing done at whatever point is convenient. Both Neve and Focusrite have produced consoles with distributed summing systems in the past.

As I'm sure you have appreciated by now, I am a great one for following a train of thought until it derails at the points and scatters passengers all over the track. If two layers of distributed summing give a better noise performance, *would three layers be even better?* The answer is generally yes, but not by much – basically it's not worth it. For example, the optimal two-layer

www.raguvaran.puzl.com

TABLE 17.6 The optimal three-layer summing structures for various numbers of inputs

Total no. of inputs	Optimal structure	Noise advantage (dB)	Advantage over two-layer summing (dB)
16	16:4:2:1	3.57	−0.07
24	24:8:2:1	5.24	0.27
32	32:8:2:1	6.24	0.36
48	48:16:4:1	7.99	0.90
56	56:14:7:1	8.33	0.81
64	64:16:4:1	9.06	1.18
72	72:18:9:1	9.27	0.79
80	80:20:5:1	9.97	1.49
88	88:22:11:1	10.00	1.16
96	96:32:8:1	10.74	1.28
104	104:26:13:1	10.59	1.13
112	112:28:14:1	11.33	1.60
120	120:40:8:1	11.62	1.49
128	128:32:8:1	11.87	1.66
136	136:34:17:1	11.51	1.08
144	144:36:9:1	12.35	1.68

structure 24:6:1 gives a noise advantage of 4.97 dB, while the optimal three-layer structure 24:8:2:1 gives 5.24 dB; the improvement over two-layer summing is a very small 0.27 dB. Noise performance is actually very slightly *worse* for the 16-input case, because you have added more summing amps creating noise, without enough partial cancellation going on to outweigh that.

See Table 17.6 for the optimal three-layer structures and their results. (Note that in the three-layer structures the second figure denotes the *total* number of first-layer summing amps rather than the number feeding each summing amp in the second layer.) The greatest advantage over two-layer distributed summing is 1.7 dB, and that is for a monster mixer with 144 paths to the mix bus; getting that improvement involves using $36 + 9 = 45$ summing amplifiers, when the optimal two-layer process uses 24. That's almost twice as many, and we must regretfully conclude that three-layer distributed summing is of marginal utility at best.

Four-layer summing anyone? Probably not; you would need a quite enormous number of inputs to make such a structure even faintly worthwhile.

Summing Amplifiers

Figure 17.24 shows a practical summing amplifier circuit using a 5534/2 op-amp; the accompanying insert send amplifier is included because it gives a good example of how the

www.raguvaran.puzl.com

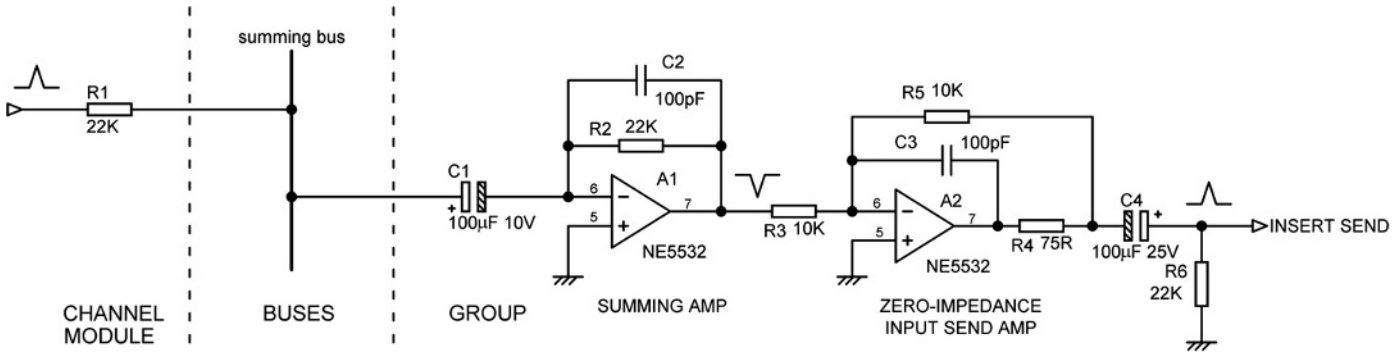


Figure 17.24: Simple op-amp summing amplifier, followed by an inverting zero-impedance insert send amplifier that corrects the phase

www.raguvaran.puzl.com

summing amp phase inversion needs to be corrected before the signal sees the outside world again, and it also demonstrates how to apply the 'zero-impedance' approach to an inverting stage. As mentioned before, the summing amplifiers were one of the first places that 5534/2s appeared in mixers, because the potentially high noise gain makes their noise performance critical. The circuit is very straightforward but there are a few points to watch. Firstly, the DC-blocking capacitor C1 is essential to prevent the bias current of the 5534/2 from making the routing switches clicky. It needs to be bigger than might at first appear, because the relevant LF time-constant is not C1 and the mix resistor R1 in the channel, but C1 and the parallel combination of all the mix resistors going to that bus. Thus for the 22k mix resistors and 100 μF capacitor shown, if there are eight input channels the LF −3 dB point will be at 0.59 Hz. Table 17.7 illustrates how this works, and you can see that in this case 100 μF is large enough for even a big mixer. If lower values of mix resistor are used then C1 must be scaled up proportionally.

It may appear that C1 is still rather oversized for the job. This is not so. The signal in a mixing console passes through a large number of RC time-constants on its journey from input to output, and even if each one causes only a tiny roll-off at the LF end, the cumulative effect on the frequency will be substantial. It is therefore essential that each LF roll-off is well below the normal audio band. There is also the important point that electrolytic capacitors distort with even small signal voltages across them, as described in Chapter 2, and so they must be much larger than might otherwise be required. The same considerations apply to HF roll-offs.

There is another reason to make C1 large. A summing amplifier aspires to appear as a short-circuit to ground, and even if it is itself perfect, the presence of C1 means that the bus residual voltage will rise at low frequencies, possibly causing increased LF crosstalk.

The capacitor C2 across the summing amp feedback resistor R2 is another vital component; without it the stage is pretty well guaranteed to be unstable at HF. This is because the mix bus has an appreciable capacitance to ground, and this will destabilize the op-amp by adding an

TABLE 17.7 How the LF frequency response of the summing amp in Figure 17.24 varies with the number of mix resistors connected to the bus

No. of mix resistors	Total mix res. to bus	−3 dB frequency (Hz)
1	22 kΩ	0.072
8	2.75 kΩ	0.59
16	1.375 kΩ	1.18
24	917 Ω	1.74
32	687 Ω	2.33
40	550 Ω	2.90
48	458 Ω	3.44

www.raguvaran.puzl.com

extra phase-shift to the feedback through R2. Adding C2 compensates for this; its value is normally determined by experiment. It clearly must not be so large that it gives an appreciable HF roll-off with R2.

Hybrid Summing Amplifiers

It is well known that at low source resistances, discrete bipolar transistors are normally quieter than op-amps, the exception being specialized op-amps like the AD797, which are usually ruled out on the grounds of cost. This applies to the normal run of low-cost transistors, though specialized types such as the 2SB737 are even better. While the advantage obtained varies with the number of summing resistors connected to the bus, and their value, in general a noise reduction of about 5 dB may be expected in a large console.

The hybrid combination of a discrete bipolar transistor input and an op-amp to provide open-loop gain for linearization and load-driving capability gives the arrangement shown in Figure 17.25. The summing bus is connected to the emitter of Q1 via the DC-blocking capacitor C1; the amplified signal from the Q1 collector is passed to the op-amp, which has a local feedback loop R3 that controls the DC conditions of Q1. The voltage set up on the non-inverting op-amp input by bias network R7, R9, C5 determines the voltage at the Q1 collector. C7 acts as a dominant-pole capacitor to give HF stability.

An outer layer of shunt feedback via R4 is used to minimize the effect of C1 in increasing the bus residual at LF, by putting C1 inside this outer feedback loop. C4 prevents the non-zero voltage at the op-amp output from getting on to the bus via R4, and in the same way C6 is a DC-blocking capacitor that prevents the following stage from putting any DC offset voltages on to the bus. This arrangement is effective at doing what it is supposed to, but there are clearly several LF time-constants involved because of the various blocking capacitors, and with such schemes it is essential to check that there is no peaking or other irregularity in the frequency response at the bottom end.

While the one-transistor version gives a good noise performance, it is susceptible to rail noise getting into the emitter or collector circuits of the transistor, and very careful filtering and decoupling are required to prevent this putting a limit on the noise performance. Another potential problem is that there will be LF signal-related voltages on the supply rails, due to the limited effectiveness of decoupling capacitors of reasonable size, and if these get into the summing amplifier they will compromise the LF crosstalk performance. An excellent fix for these issues is the use of a two-transistor circuit as in Figure 17.26. The two transistors are used in a balanced configuration that cancels out rail noise and makes the decoupling requirements much simpler; there is also no need for a biasing network as in the one-transistor version.

There are one or two subtleties to be observed here; the shunt feedback must go to the same transistor as the VE bus, to prevent large current swings in the transistors that would induce

www.raguvaran.puzl.com

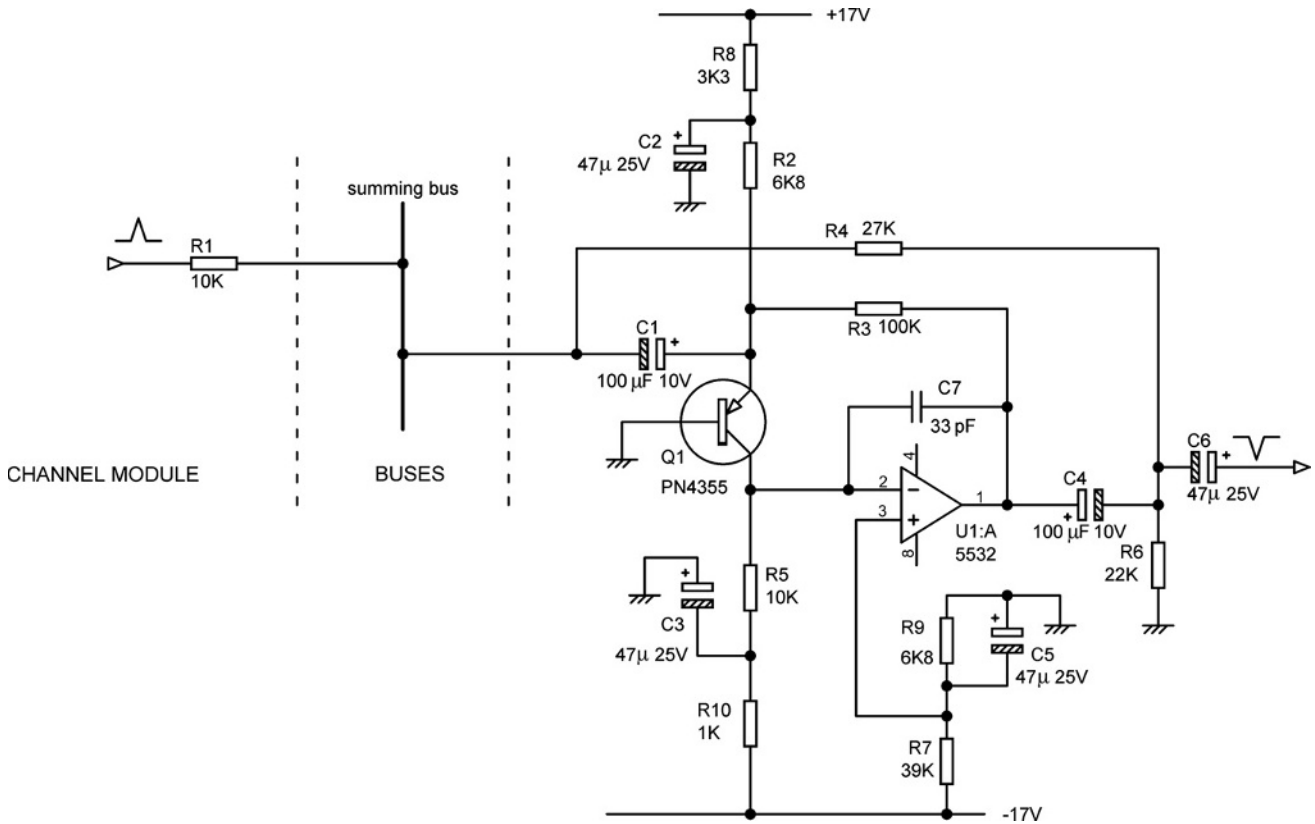


Figure 17.25: One-transistor hybrid summing amplifier

www.raguvaran.puzl.com

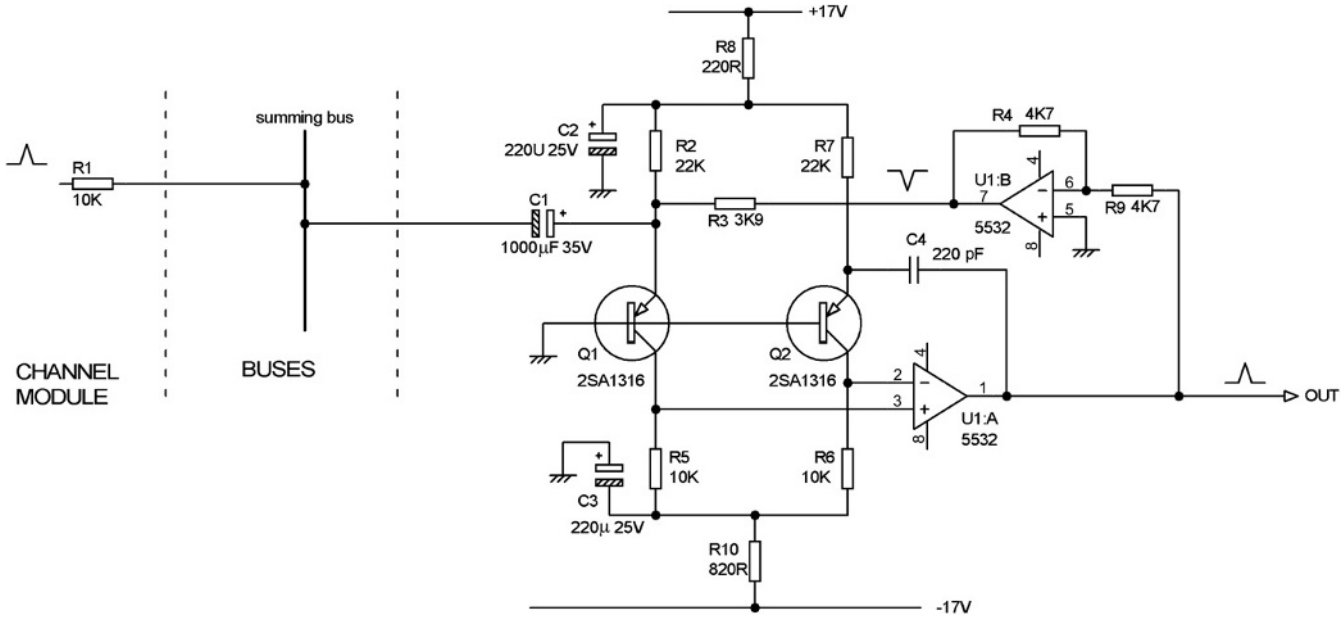


Figure 17.26: Two-transistor hybrid summing amplifier is better at rejecting noise and crosstalk from the supply rails

www.raguvaran.puzl.com

distortion. This means that the feedback must be inverted to be in the right phase, which is done by U1:B. This introduces stability-threatening extra phase-shifts, so HF stability capacitor C4 bypasses the feedback inverter and goes straight to the emitter of Q2. A very useful property of this arrangement is that the output is in phase and so can be fed directly to a group insert.

A more sophisticated version of this summing amplifier, which I thought up a while ago, is shown in Figure 17.27. Here the aux master level control RV1 not only acts as a normal gain control, but also alters the amount of negative feedback around the summing amplifier, effectively giving the stage variable gain. This can be extremely useful as aux summing amplifiers often have large numbers of inputs feeding them and are more liable to clipping than group summing amplifiers. It is extremely tedious to turn down dozens of aux sends to remedy the situation, so being able to reduce the aux summing amp gain is advantageous.

When RV1 is at minimum, the circuit is essentially that of Figure 17.26, with the main shunt feedback path through the inverter U1:B, and the HF stabilizing path through C7. The value of the main feedback resistor R3 is chosen so the sum amp gain is low, in this case −6.8 dB, so overload is unlikely. When RV1 is turned up, a positive feedback signal is sent to the non-inverting input of U1:B via R11 and R12, and the partial cancellation that results reduces the output of the inverter, decreasing the overall amount of negative feedback and increasing the gain of the summing amplifier. The need for a postamplifier with fixed make-up gain is therefore avoided, improving the noise performance.

This technique was used in the Soundcraft Delta mixing console, and I modestly propose that it might be one reason why it won a British Design Award in 1991. My collaborator in the design of the Delta was Mr Gareth Connor.

Balanced Hybrid Summing Amplifiers

In sophisticated consoles it is desirable to combine the benefits of balanced mixing with low-noise hybrid summing amplifiers. The obvious method of implementing this is to use two hybrid summing amplifiers and then subtract the result, but this uses quite a lot of hardware. A more elegant approach is to use a single balanced hybrid summing amplifier to accept the two anti-phase mix buses simultaneously; this reduces noise as well as minimizing parts cost and power consumption.

The circuit shown in Figure 17.28 is at first sight very similar to that in Figure 17.26, but with the crucial difference that two mix buses now feed into the stage, one to each transistor emitter. To prevent distortion-inducing variations in the transistor currents, and to maintain symmetry, it is therefore now necessary to also apply shunt feedback to both transistor emitters. The feedback to Q1 is taken via an inverter to get the phase right, as in Figure 17.26, but the other feedback path is simply resistor R12. As before, C4 gives HF stabilization.

www.raguvaran.puzl.com

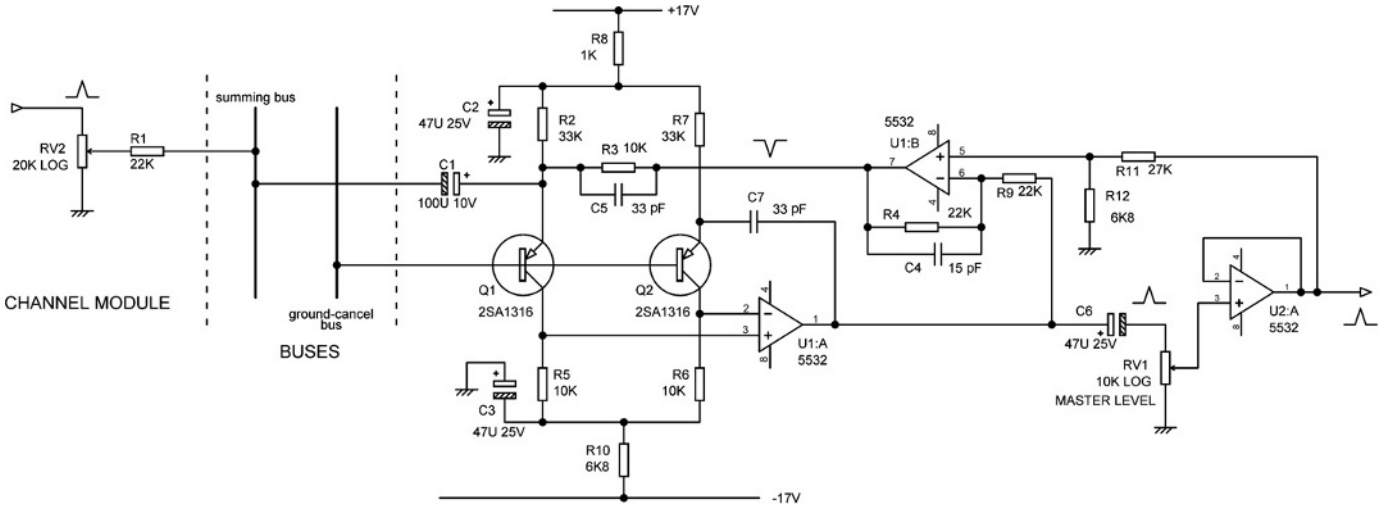


Figure 17.27: Two-transistor hybrid summing amplifier with positive feedback to enhance gain at high control settings only, improving headroom at low gain settings

www.raguvaran.puzl.com

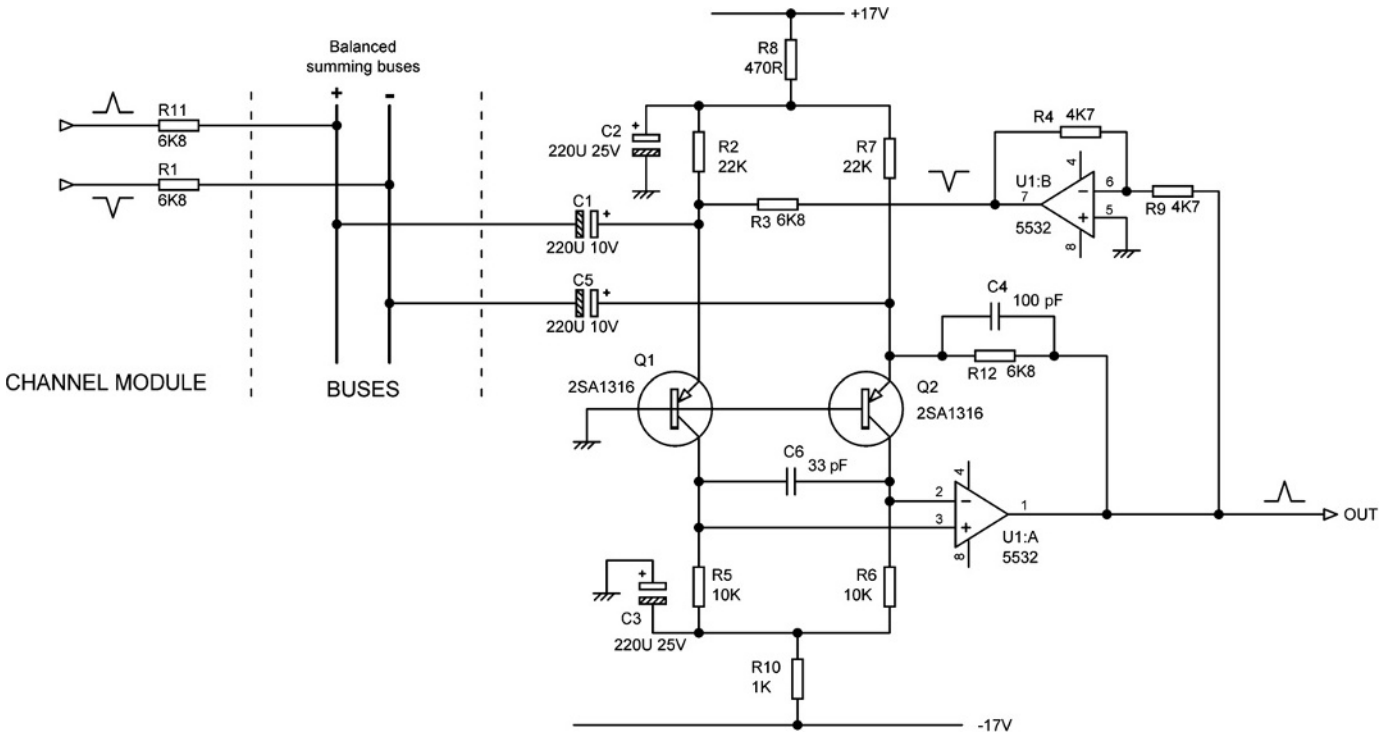


Figure 17.28: Two-transistor balanced hybrid summing amplifier. Note the dual feedback paths via R3 and R12

www.raguvaran.puzl.com

Note that the configuration is very similar to that of the balanced microphone amplifier described in Chapter 13, and gives low noise as well as excellent symmetry. This technology was used extensively in the Soundcraft 3200 recording console; once again, my valued collaborator in the design of the 3200 was Mr Gareth Connor.

PFL Systems

The prefade listen (PFL) facility goes back a long way in the history of mixers, almost certainly first appearing on broadcast consoles, where it proved extremely useful to be able to listen to a source before unleashing it on an unsuspecting public. It is also very handy in a recording environment, allowing you to listen to one source in isolation without undoing dozens of routing switches. It is invaluable for checking the level in a channel; the PFL feed takes over the main Mix metering, which is usually much bigger and better than metering incorporated on a per-channel basis, if indeed there is any at all. On most mixers there is only room for a Peak light. As the name suggests, a PFL feed is taken from before the fader, so that the signal is heard at full level even if the fader is down. In the USA it is often called the 'solo' facility.

'PFL' is often used as a verb, as in 'Could you PFL channel 23?' Interestingly, this request is always spelt out as P-F-L and the obvious pronunciation as 'piffle' has never achieved favor. When any PFL switch is pressed, a circuit block in the master section switches the left and right monitor outputs, so they reproduce the PFL signal in mono, as shown in Figure 17.29. The L/R meters are also taken over by the PFL system so they can be used to check channel prefade levels, allowing the adjustment of input gain when required. It is common to switch

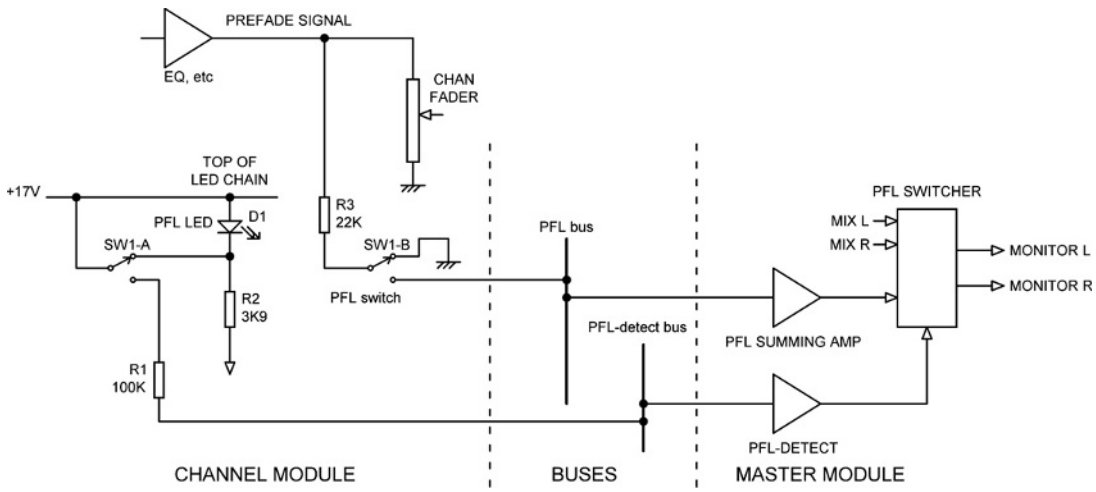


Figure 17.29: A PFL system, showing how to make a two-section PFL switch perform three functions at once

www.raguvaran.puzl.com

both meters over, though of course they both read the same thing. Historically PFL switching was done with a relay, but this was taken over by electronic means, specifically 4016 analog gates, at a relatively early date, as it is not essential for the switching to be absolutely click-free. The later availability of discrete FETs designed for analog switching improved the linearity and much reduced (but did not wholly eliminate) the switching transients.

All that is required on each channel is a switch section to route the prefade signal to the PFL bus, which is a simple virtual-earth bus that works in the same was as group buses, and another switch section to signal via a PFL-detect bus that a switch has been pressed and the monitor source must be changed over. The PFL-detect bus simply gathers together all the signals that indicate a PFL switch is pressed; it does not have to be a virtual-earth bus, but there are advantages in making it so, as described below. PFL switches are commonly fitted with indicator lights, even on quite small mixers, because otherwise it is necessary to hunt around the control surface if a PFL switch is inadvertently left pressed in. This does not make a third switch section (and therefore a four-changeover switch) necessary because the ingenious method shown in Figure 17.29 can be used. All the LEDs on a channel module are usually run in a chain from the positive to negative supply rails; so long as there are not too many LEDs a simple series resistor is usually all that is needed to give more or less constant-current operation of the chain. If the PFL indicator LED is put at the top of the chain, SW1-A removes the short-circuit from D1 and allows it to turn on, at the same time routing a DC signal into the PFL-detect bus via R1. (Note that the switch section SW1-A that routes the signal to the PFL bus uses the same high-offness configuration described earlier.)

This underlines the point that economy of design in a mixer channel is very important, because every extra component is multiplied by the number of channels. In the master section the need to design out every possible component is less pressing.

PFL Summing

There is nothing very difficult about the PFL summing amplifier. It is not necessary for it to have the best possible noise performance as it is only used for quick monitoring checks. The system shown in Figure 17.30 uses a dual op-amp, the second half being used to restore the correct phase after the inversion produced by the summing amplifier. C1 provides DC blocking, while R2 keeps the PFL bus at 0 V DC to prevent switch-clicks. R3 sets the gain from channel to PFL summer as unity, and C2 prevents the bus capacitance from causing instability. R4 and R5 set the gain of the re-inverting stage to unity. C3 provides DC blocking at the output to prevent clicks when the PFL switcher operates.

PFL Switching

When the PFL system operates, the normal L–R mix is removed from the monitor outputs and is replaced by a mono feed from the PFL summing system. This is usually

www.raguvaran.puzl.com

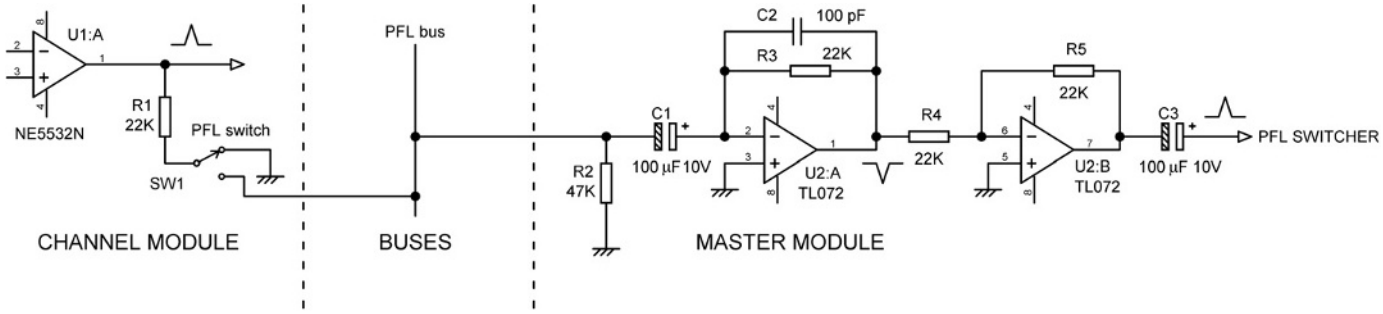


Figure 17.30: An example of a PFL summing system

www.raguvaran.puzl.com

done by some form of electronic switching (see Chapter 16 on signal switching for more details).

PFL Detection

In early consoles a DC voltage was simply switched into a PFL-detect bus that operated the PFL switching relay directly. Moving that amount of current around inside a mixer is simply asking for trouble with clicks crosstalking into the summing buses, and very soon a transistor relay driver was added so the voltage and current levels on the PFL-detect bus could be much lower. When solid-state PFL switching was introduced the transistor circuitry drove analog gates or discrete FETs instead.

Two simple methods of PFL detection are shown in Figure 17.31. Figure 17.31(a) demonstrates how not to do it. When the PFL switch SW1 is closed, +17 V is applied to the PFL-detect bus, turning on Q1 and illuminating the 'PFL Active' LED D1, and turning Q2 off. The problem is that the 34 V edge on the PFL-detect bus (which normally sits at −17 V) has an excellent chance of crosstalking into the other summing buses.

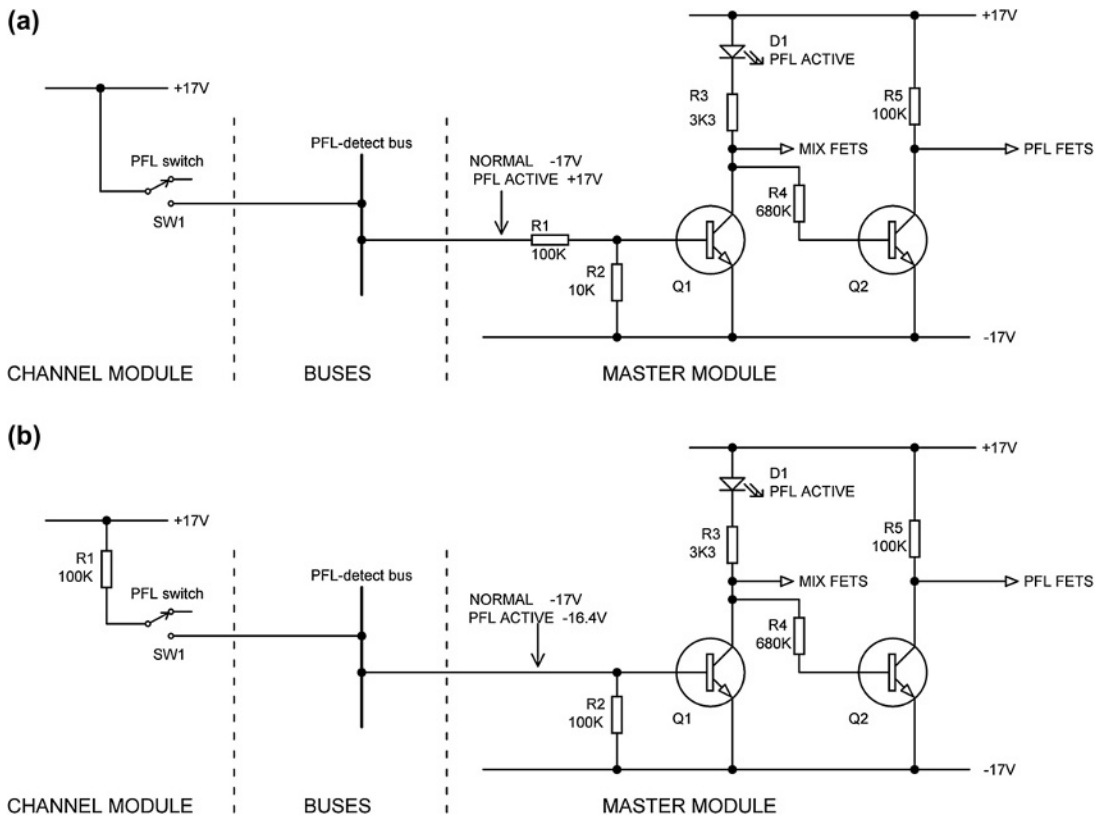


Figure 17.31: Two simple PFL-detect systems. The second version reduces the likelihood of clicks getting into other buses

www.raguvaran.puzl.com

A better method is shown in Figure 17.31(b). Here the resistor R1 has moved to the channel module, and the PFL-detect bus connects directly to the base of the detect transistor Q1. This ensures that the PFL-detect bus can never change by more than 0.6 V, which reduces the possibility of clicks crosstalking capacitatively into adjacent summing buses; there should be a 35 dB advantage. A countervailing disadvantage is that the base of Q1 is vulnerable to destruction if the PFL-detect line gets shorted to ground. It is also slightly more expensive to implement as now there has to be a resistor R1 for each PFL switch.

This method was workable when all mixers had motherboards, but the introduction of ribbon-cable interconnections between mixer modules put all the buses closer together, and increased the likelihood of PFL-detect clicks becoming audible.

Virtual-Earth PFL Detection

The ideal PFL detection system would have no voltage edges on the detect bus at all. This can be done by detecting current instead of voltage – the PFL-detect line becomes a virtual-earth summing bus for DC signals. I thought this up in 1983.

Figure 17.32(a) shows the principle. U1 is simply a summing amplifier that works at DC. When no PFL switches are pressed, the output of U1:A sits at +10 V because of the bias from resistor R2, Q1 is held on via R4, and Q2 is held off. Negative feedback through R3 keeps the detect bus at 0 V; C1 prevents the bus capacitance to ground from causing instability. The 'PFL Active' LED has moved to the collector of Q2 to allow for the logical inversion caused by the summing amplifier. When a PFL switch is pressed, the current injected into the detect bus by R1 overcomes that from R2 and the output of U1 goes below 0 V, and Q1 turns off. Note that the signaling current is only 170 μA, so there is little possibility of inductive crosstalk. It could probably be significantly reduced without problems.

This simple version of the system will silently detect a single PFL button down, but will come unstuck if more are pressed, because U1 output will hit the −V rail and will no longer be able to maintain a virtual earth on the detect bus. This situation is not uncommon, as it gives a rough sub-mix of the channels pressed, and so it is wise to cater for any number of PFLs at once. The circuit of Figure 17.32(a) overcomes the problem by adding diode D1, which clamps the op-amp output so that however many PFL switches are pressed, the op-amp output will not hit the −V rail and negative feedback will continue to keep the PFL-detect bus very close to zero volts. The values of R4 and R5 must be carefully chosen so that the level-shift down to the −V rail works reliably and Q1 will always switch correctly despite circuit tolerances.

Note that once again all the circuitry works between the +V and −V rails, so there is no chance of transients being injected into the ground system.

This detection method has another less obvious advantage. It can handle signals in both directions. The ribbon-cable systems referred to earlier come in a number of fixed widths, so

www.raguvaran.puzl.com

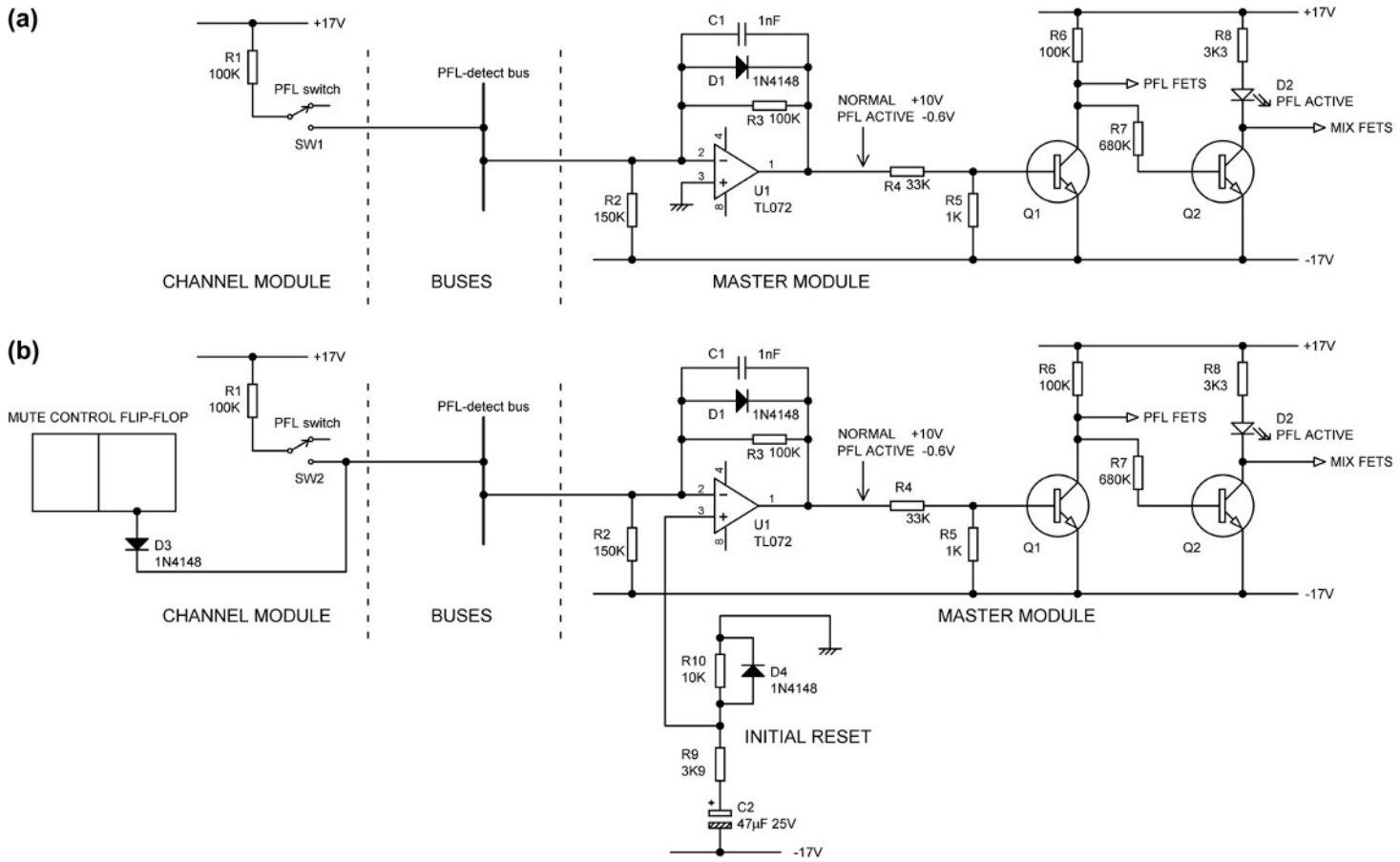


Figure 17.32: Virtual-earth PFL-detect systems. The second version uses the PFL-detect bus to reset channel mute status at power-up.

www.raguvaran.puzl.com

you can buy 40-way cable, but if you need 41 ways the next size up is 50-way. Since this increases the cost not only of the cable itself, but of all the connectors on it and all the mating module connectors, increasing the cable size is not a decision taken lightly. In one mixer design there was a need to send an initial-reset signal to mute-control flip-flops on each module, and all 40 ways of the ribbon cable were committed. The only way to avoid 50-way cabling was to make the PFL-detect bus dual-function.

The idea is that by changing the reference voltage applied to the PFL-detection op-amp, you can change the DC voltage that it maintains on the detect bus. In Figure 17.32(b), at power-up, the initial-reset network R9, R10, C2 pulls the reference voltage going to the non-inverting input of U1 low for a brief period until C2 fully charges, and the PFL-detect bus is therefore also pulled low via R3, and this signal is coupled to each channel mute-control flip-flop by a diode D3; the flip-flop works between positive and negative supply rails so it can drive muting FETs directly. After the initial-reset period the PFL-detect bus settles at its normal level of 0 V and D3 remains reverse-biased. D4 in the initial-reset network ensures rapid recovery when the power is switched off. Obviously the PFL system cannot function during the brief initial-reset period, but after that it works just as described before.

This is a good example of the ingenuity that is sometimes required to fit all the required functionality into a given size of ribbon cable. If there are spare conductors they are used to reinforce the audio ground.

AFL Systems

AFL is a very similar concept to PFL, the difference being that an AFL feed is taken from after the fader, so that if a selection of AFL buttons are pressed, the resulting mix has the sources in the right proportion. A stereo AFL system goes one better by taking a stereo feed from after the panpot, so that the sources selected can be heard in their correct positions in the stereo field as well. Figure 17.33 shows the differing points in the channel from where PFL, AFL, and stereo AFL are taken. Stereo AFL requires a four-pole switch; two sections route the stereo signal to the AFL buses, and one is required to signal the AFL-detect system. The detection of an AFL condition works in exactly the same way as PFL detection.

It is not common for mixers to be fitted with an AFL system only, as it is much less convenient for checking signal levels in the channel; it will only read the prefade level correctly if the channel fader is at 0 dB. More usually a solo system is fitted that can be switched from PFL to AFL at the master section. The channel module switch will route signals to both the PFL and AFL buses; these are summed by separate summing amplifiers, and which of the two resulting signals is switched to the monitors is selected by the PFL/AFL switch on the master section.

www.raguvaran.puzl.com

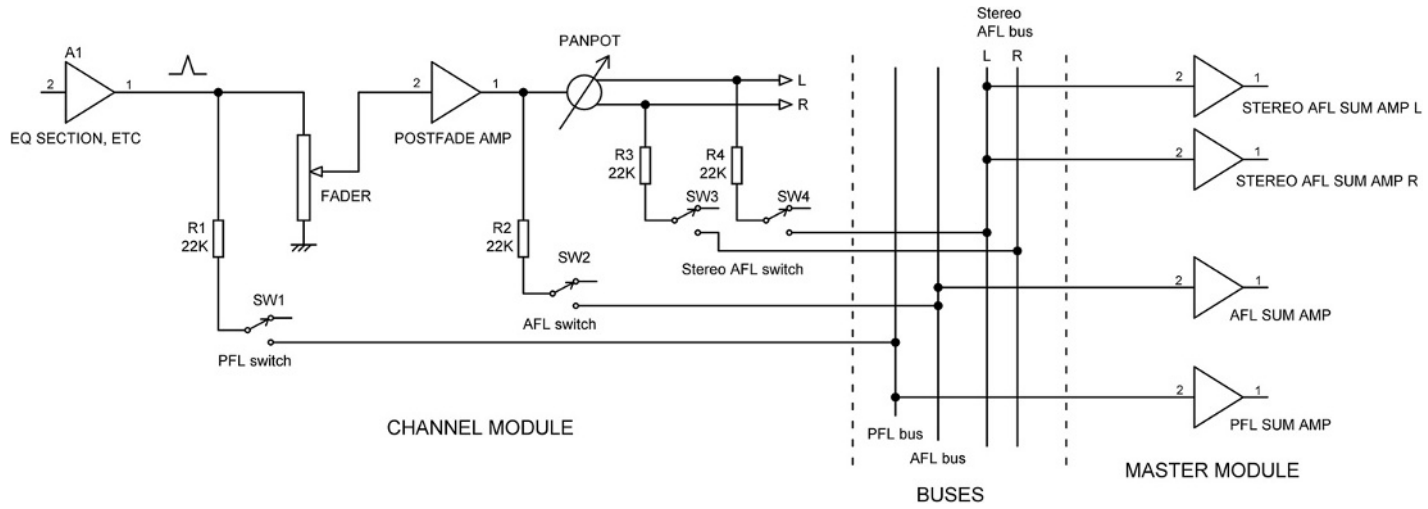


Figure 17.33: Comprehensive system with PFL and mono and stereo AFL

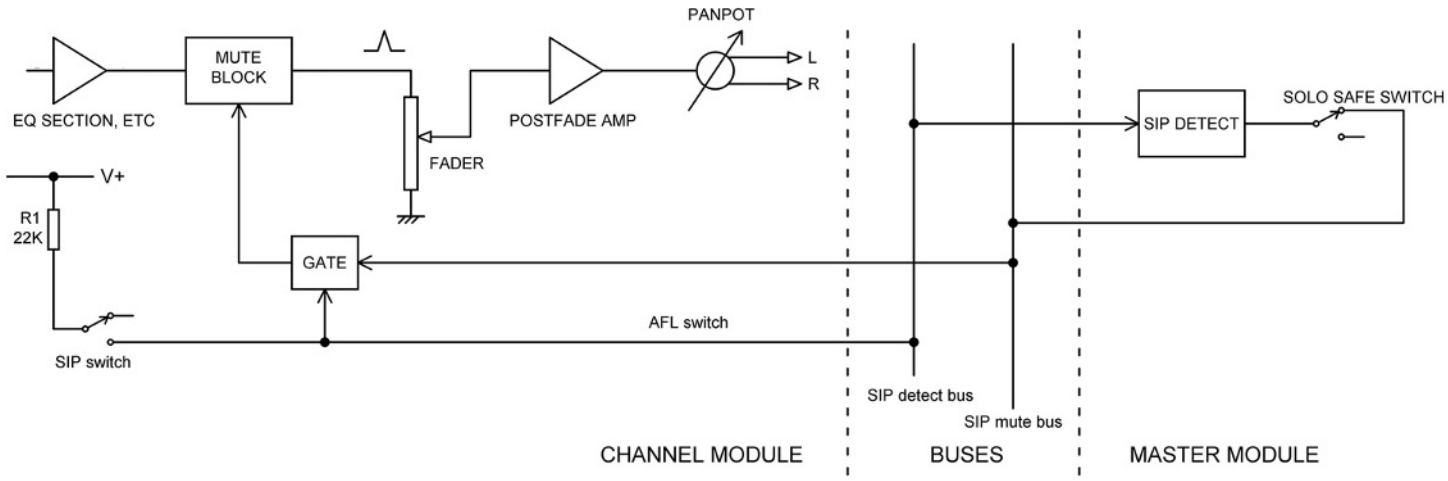


Figure 17.34: Solo-in-place system

www.raguvaran.puzl.com

Solo-in-Place Systems

PFL provides a mono look at a single channel. AFL preserves fader and perhaps pan position, but in both cases the signal that is heard is devoid of effects, which makes critical assessment difficult. A solo-in-place (SIP) system allows a channel or set of channels to be heard in isolation with all the effect sends working normally. This is done simply by muting all the channels except the ones you want to hear. This is obviously not applicable to PA work (except during set-up and sound checks) because it is what is called a 'destructive solo' in that the output of the mixer is disrupted while SIP is in use, unlike in PFL and AFL systems, which take over the monitor output but leave the group and LR mix outputs unaffected.

A SIP system requires an electronically controlled mute on each channel. When the SIP button on a channel is pressed, every *other* channel is muted except that one. If two or more channels are solo-ed, they are left unmuted while the rest of the channels are muted.

The basic SIP system therefore has a SIP-detect bus, which works in just the same way as a PFL-detect bus. This determines when one or more SIP switches are pressed and activates an SIP mute bus that potentially mutes all the channels; however, if the SIP switch on a channel is pressed it intercepts the mute signal and prevents it from operating on that channel. The basic scheme is shown in Figure 17.34.

SIP is a destructive solo, and you really don't want to use it during a live performance. Therefore a 'solo-safe' switch is usually fitted in the master section to act as a safety-catch or Molly-guard to prevent accidental operation. When it is engaged, the solo system usually acts as a PFL system instead, which is safe to use during performance.

Talkback Microphone Amplifiers

Most contemporary talkback systems have a small electret microphone mounted flush with the master module panel. The microphones used typically have an internal head amplifier that buffers the high impedance of the electret element, and all else that is needed is an amplifier stage with variable gain over a wide range. A typical arrangement is shown in Figure 17.35; gain is variable from 0 to +55 dB by RV1, with C4 keeping it to unity at DC. Note that the microphone is powered through the filter network R1 and C1.

Figure 17.35 includes a couple of subtle but important design features. Firstly, the second changeovers of the talkback routing switches are connected into the feedback network so that when none of them is pressed, the amplifier gain is reduced to unity. This prevents the microphone amplifier from clipping continuously if the gain is turned up but talkback is not in use; the resulting distorted waveforms would almost certainly crosstalk crunchily into other parts of the master module, such as the L–R mix summing amplifiers, and this would not be a good thing.

www.raguvaran.puzl.com

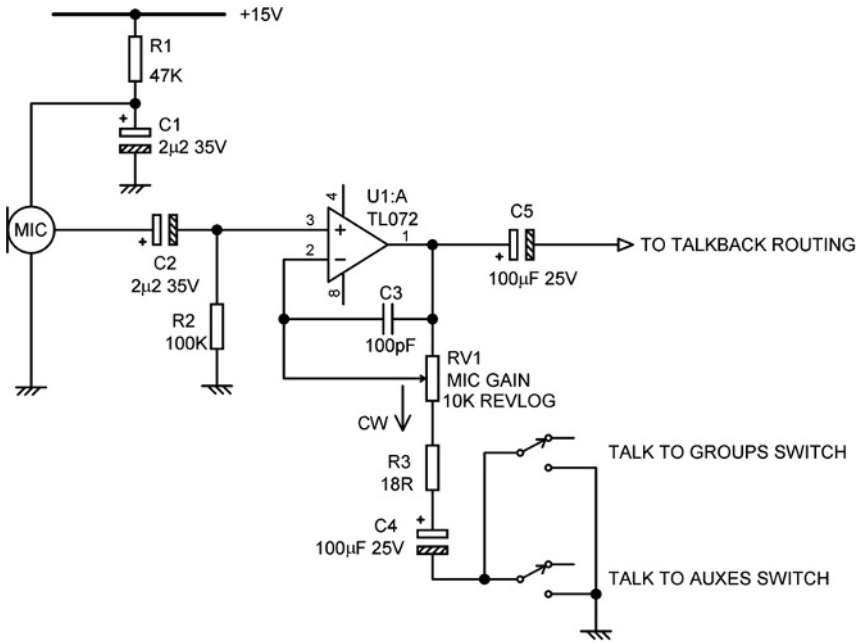


Figure 17.35: A typical talkback microphone amplifier, with gain variable from to 0 to +55 dB

Secondly, you will have noticed that the value of C4 looks rather small compared with the low value of the gain end-stop resistor R3, and in fact gives an LF roll-off that is −3 dB at 90 Hz. This is quite deliberate, and is intended to control the amplitude of LF transients when the gain is set high.

Line-Up Oscillators

The oscillator output is not required to be a perfect sine wave; it only has to be good enough so that meter calibration is not affected. It also needs to look like a sine wave on an oscilloscope, or customer confidence will be undermined. Figure 17.36 shows a simple but dependable arrangement for a 1 kHz oscillator that does not require either expensive and fragile thermistors or complicated leveling circuitry.

It consists of a feedback loop containing a bandpass filter and a soft clipping circuit. The waveform at the output of bandpass filter U1:A is amplified by U1:B and then soft-clipped by D1 and D2. This symmetrical clipping introduces odd-order harmonics only, which are more easily filtered out when the signal is fed back to the multiple-feedback bandpass filter U1:A. The distortion is approximately 0.5%. Note that when the oscillator is not in use, it is not routed to the buses, but is also firmly stopped from oscillating at all by the closure of SW1. This is essential to remove any possibility of the oscillator crosstalking into the audio paths

www.raguvaran.puzl.com

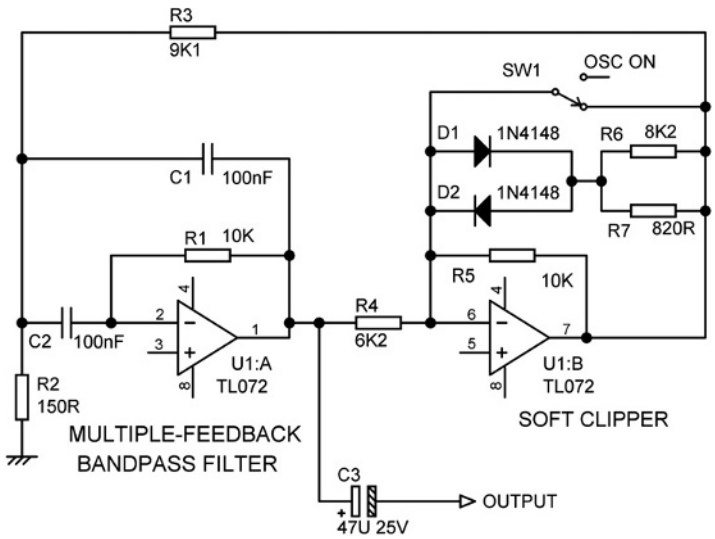


Figure 17.36: One-kilohertz oscillator with diode amplitude stabilization

during normal operation. When SW1 is opened this oscillator ramps up in amplitude in a very neat manner.

You will note the use of E24 resistor values to get the frequency as close to 1.0 kHz as possible, and the parallel combination of R6 and R7 to give 745 Ω; this resistance is fairly critical for obtaining the minimum distortion, and the combination shown here gave dependably lower distortion than the E24 value of 750 Ω. The output level is dependent on the forward voltage of the two diodes and so there is some variation with temperature, but it is insignificant in this application.

The next step up in oscillator sophistication is to provide two switched frequencies, usually 1 and 10 kHz. The higher frequency was historically used for checking the azimuth adjustment of tape machine heads (when the azimuth of a tape head is maladjusted, the high-frequency response falls off badly). Nowadays it is less useful but can still come in handy for quick frequency response checks. It is therefore very important that the output level is exactly the same for the two frequencies, and the oscillator shown in Figure 17.37 therefore has a more complex level-control loop.

The oscillator is based on the classic Wien bridge configuration, with the oscillation frequency controlled by R8, R9, C1 and R5, R6, C2. When running at 1 kHz, R8, R9 and R5, R6 are in circuit. For 10 kHz SW1 is pressed and R5 and R9 are now shorted out, raising the operating frequency. The level-control loop operation is as follows: when SW2 is pressed the short-circuit across R1 in the negative-feedback loop R1, R2 that prevents oscillation is removed and the amplitude of oscillation ramps up. When it reaches the

www.raguvaran.puzl.com

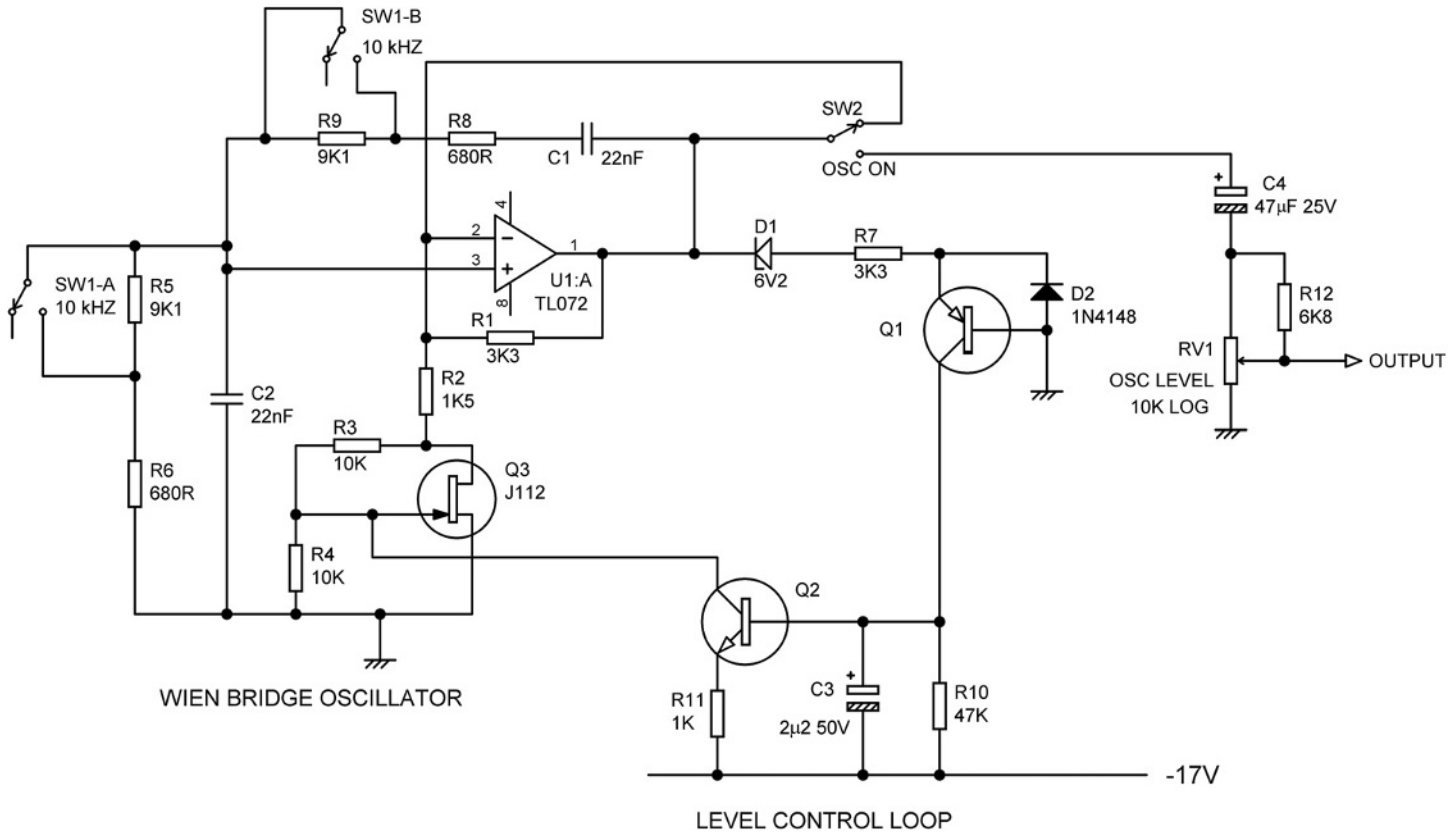


Figure 17.37: Switched 1 kHz/10 kHz oscillator with FET amplitude stabilization

www.raguvaran.puzl.com

desired level, Zener diode D1 begins to conduct on positive peaks and turns on the common-base transistor Q1 (D1 also conducts on negative half-cycles, but Q1 does not respond and is protected from reverse bias by clamp D2). The collector current of Q1 charges C3 and Q2 turns on, pulling down the gate of JFET Q3 and increasing its channel resistance, thus increasing the amount of negative feedback through R1, R2 and regulating the level. Q3 is a J112 FET, a type that is optimized for voltage-controlled resistance (VCR) operation. The network R3, R4 not only acts as the collector load for Q2, but also feeds half the V_{ds} of Q3 to its own gate; this is a classic method of reducing even-order distortion in JFETs and is dealt with in more detail in Chapter 19. C3 and R10 set the time-constant of the control loop, and their values have a strong effect on oscillator distortion. Q1 and Q2 can be any high-beta transistor types.

Larger mixers often have more sophisticated oscillators with fully variable frequency and sometimes a choice of a square-wave output. Oscillator design is a massive subject on its own, and there is no space to get deeper into it here.

Console Cooling and Component Lifetimes

Large mixing consoles have a lot of of electronics packed into a small space, both to reduce the size of the console so it is easier to install, and to bring all the controls within a reasonable reach. This means that natural convection cooling is often barely adequate. It is made more difficult by the fact that open cooling vents on the top of a console are not popular due to the possibility of various beverages being poured into the works, with an excellent chance of catastrophic results. Elevated internal console temperatures are rarely high enough to cause semiconductor failures, but have two bad consequences:

Firstly, over time the electrolytic capacitors will dry out and drop in value. This loss of capacitance can be dramatic: a 100 μF component can fall to 10 nF. This in turn causes two further bad consequences: the LF response of the console degrades as coupling capacitors drop in value, and op-amps go unstable as rail decoupling capacitors become ineffective. In a large studio console this can mean that its life is only 10 years before all the electrolytics need replacing. This ‘recapping’ procedure is, as you might imagine, a lengthy and expensive process that can be done only by companies with specialized skills.

The average electrolytic capacitor has a temperature rating of 85°C. This does not mean that no capacitor degradation occurs below that temperature; degradation is happening all the time, but it accelerates rapidly as temperature rises, roughly doubling in speed with each 10°C increase. 85°C is the temperature at which capacitor life has dropped to a nominal 1000 hours, which is only 49 days, and so obviously they must be operated at a much lower temperature than that to get a reasonable equipment lifetime. Bear in mind that large recording consoles are normally powered 24/7. Electrolytic capacitors rated at 105°C have recently fallen in price to

www.raguvaran.puzl.com

the point at which they are a viable alternative, and at a given temperature this increases lifetime by a factor of 4. The best recapping facilities, such as GJC Designs or Scotch McNeil Audio use 105°C capacitors exclusively for replacements.

Secondly, if silver contacts are used in switches or relays, the rate at which they are corroded by atmospheric hydrogen sulfide, creating non-conducting silver sulfide and causing the contacts to fail, also increases rapidly with temperature. Once again it roughly doubles in speed with each 10°C increase.

The obvious way to improve console cooling is to add fans to increase the air flow. Fans are very undesirable in consoles installed in recording studios because of the acoustic noise they generate, and are rarely if ever used. They are, however, sometimes fitted in PA consoles, though still with considerable care to minimize the noise they create. When this is done by running the fan at low speeds, at the low end of its specified voltage range, great caution is needed. After a period of use the fan may stop altogether, as the bearings become worn, and how long this takes is not readily predictable. Sleeve-bearings seem to be worse than ball-bearings for this. The use of fans to cool a mixing console is to a large extent a last resort, and every effort should be made to induce adequate cooling by convection alone.

www.raguvaran.puzl.com

CHAPTER 18

Level Indication and Metering

Signal-Present Indication

Some amplifiers and the more sophisticated mixers are fitted with a 'signal-present' indicator that illuminates to give reassurance that a channel is receiving a signal and doing something with it. The level at which it triggers must be well above the noise floor, but also well below the peak indication or clipping levels. Signal-present indicators are usually provided for each channel, and are commonly set up to illuminate when the channel output level exceeds a threshold 20 dB below the nominal signal level.

A vital design consideration is that since it may be active most of the time, the operation of the signal-present detector must not introduce distortion into the signal being monitored; this could easily occur by electrostatic coupling or imperfect grounding if there is a comparator switching on and off at signal frequency. A typical circuit would comprise just the bottom step of that of the LED bar-graph meter shown in Figure 18.4 below.

Peak Indication

A mixer has a relatively complex signal path, and the main metering is normally connected only to the group outputs. The mix metering can be used to measure the level in other parts of the signal path by use of the PFL system (see Chapter 17) if fitted, but this can only monitor one channel at a time. It is therefore usual to guard against clipping by fitting peak level indication to every channel of all but the simplest consoles, and sometimes to effect-return modules also.

The peak indicator is driven by fast-attack, slow-decay circuitry so that even brief peak excursions give a positive display. It is important that the circuitry should be bipolar, i.e. it will react to both positive and negative peaks. The peak values of a waveform can show asymmetry up to 8 dB or more, being greatest for unaccompanied voice or a single instrument, and this is of course very often exactly what goes through a mixer channel. This level of uncertainty in peak detection is not a good thing, so only the simplest implementations use unipolar peak detection. Composite waveforms, produced by mixing several voices or instruments together, do not usually show significant asymmetries in peak level.

Small Signal Audio Design; ISBN: 9780240521770
Copyright © 2010 Elsevier Ltd; All rights of reproduction, in any form, reserved.

www.raguvaran.puzl.com

Figure 18.1 shows a simple unipolar Peak LED driving circuit. This only responds to positive peaks, but it does have the advantage of using but two transistors and is very simple and cheap to implement. When a sufficient signal level is applied to C1, Q1 is turned on via the divider R1, R2; this turns off Q2, which is normally held on by R4, and Q2 then ceases to shunt current away from peak LED D1. C2 acts as a Miller integrator to stretch the peak hold time; when Q1 turns off again, R4 must charge C2 before Q2 can turn on again. Note that this circuit is integrated into the Channel On LED supply, with R5 setting the current through the two LEDs; the Channel On LED is illuminated by removing the short placed across it by SW1. R5 is of high enough value, because it is connected between the two supply rails, for there to be no significant variation in the brightness of one LED when the other turns off. If for some reason this was a critical issue, R5 could be replaced by a floating constant-current source. Other LEDs switched in the same way can be included in series with the Channel On LED.

This peak detect circuit has a non-linear input impedance and must only be driven from a low-impedance point, preferably direct from the output of an op-amp. The Peak LED illuminates at an input of 6.6 V_{peak}, which corresponds to 4.7 Vrms (for a sine wave)

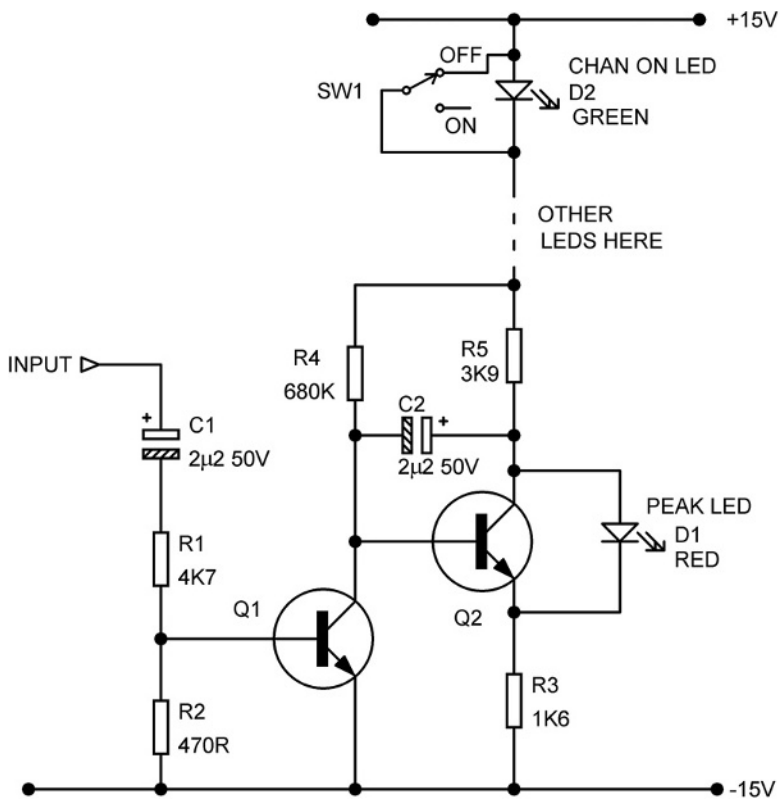


Figure 18.1: A simple unipolar peak detector, including powering for the Channel On LED

www.raguvaran.puzl.com

and +16 dBu. For typical op-amp circuitry running off the usual supply rails this corresponds to having only 3 or 4 dB of headroom left. The detect threshold can be altered by changing the values of the divider R1, R2.

Distributed Peak Detection

When an audio signal path consists of a series of circuit blocks, each of which may give either gain or attenuation – and the classical example is a mixer channel with multiple EQ stages and a fader and postfade amplifier – it is something of a challenge to make sure that excessive levels do not occur anywhere along the chain. Simply monitoring the level at the end of the chain is no use because a circuit block that gives gain, leading to clipping, may be followed by one that attenuates the clipped signal back to a lower level that does not trip a final peak-detect circuit. The only way to be absolutely sure that no clipping is happening anywhere along the path is to implement bipolar peak detection at the output of every op-amp stage. This is, however, normally regarded as a bit excessive, and the usual practice in high-end equipment is to just monitor the output of each circuit block, even though each such block (for example, a band of parametric EQ) may actually contain several op-amps. It could be argued that a well-designed circuit block should not clip anywhere except at its output, no matter what the control setting, but this is not always possible to arrange.

A multi-point or distributed peak detection circuit that I have made extensive use of is shown in Figure 18.2. It can detect when either a positive or negative threshold is exceeded, at any number of points desired; to add another stage to its responsibilities you need only add another pair of diodes, so it is very economical. However, if one peak detector monitors too many points in the signal path, it can be hard to determine which of them is causing the problem. In most applications I have used the circuit to keep an eye on the output of the microphone preamplifier, the output of the EQ section, and the output of the fader postamplifier. This means that the location of the clipping can be pinpointed quite easily. If you pull down the fader to 0 dB or below and the Peak LED goes out, the problem was at the postfade amplifier. If that doesn't do the trick, switch out the EQ; this assumes of course that the EQ in/out switch removes the signal feed to the unused EQ section. I always arrange matters so if possible, as removing the EQ signal reduces power consumption and minimizes the possibility of crosstalk. If that is not the case then you will have to back off any controls with significant boost and see if that works. Should the Peak indication persist, it must be coming from the output of the microphone preamplifier, and you will need to reduce the input gain.

The operation is as follows. Because R5 is greater than R1, normally the non-inverting input of the op-amp is held below the inverting input and the op-amp output is low. If any of the inputs to the peak system exceed the positive threshold set at the junction of R4, R3, one of D1, D3, D5 conducts and pulls up the non-inverting input, causing the output to go high. Similarly, if any of the inputs to the peak system exceed the negative threshold set at the junction of R2, R6,

www.raguvaran.puzl.com

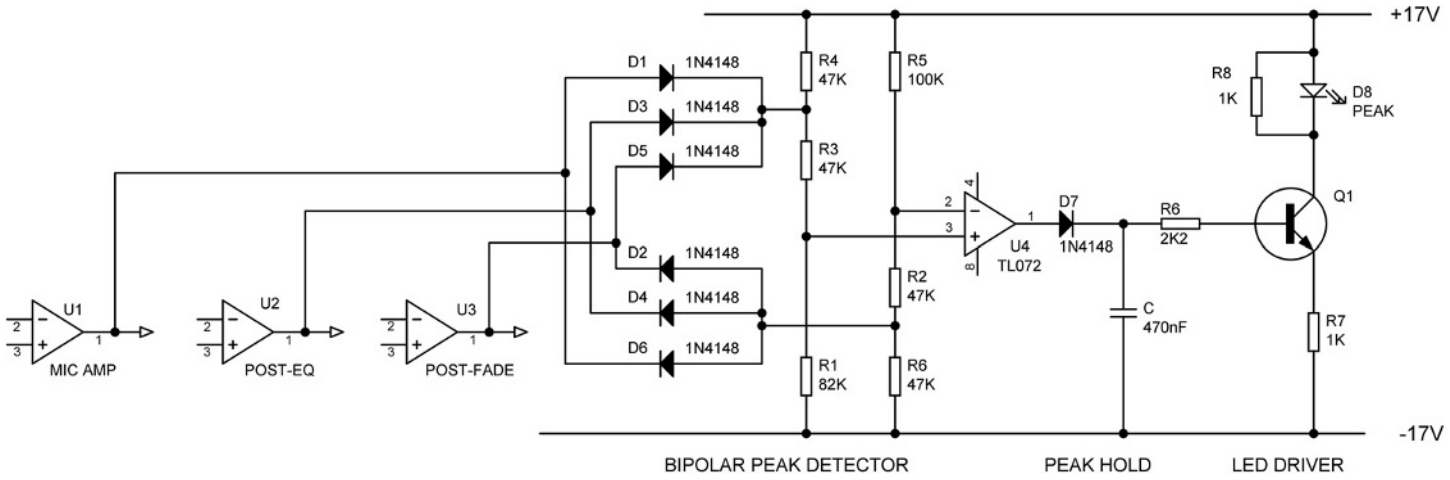


Figure 18.2: A multi-point bipolar peak detector, monitoring three circuit blocks

www.raguvaran.puzl.com

one of D2, D4, D6 conducts and pulls down the inverting input, once more causing the op-amp output to go high. When this occurs C1 is rapidly charged via D7. The output-current limiting of the op-amp discriminates against very narrow noise pulses. When C1 charges Q1 turns on, and illuminates D8 with a current set by the value of R7. R8 ensures that the LED stays off when U4 output is low, as it does not get close enough to the negative supply rail for Q1 to be completely turned off.

Each input to this circuit has a non-linear input impedance, and so for this system to work without introducing distortion into the signal path, it is essential that the diodes D1–D6 are driven directly from the output of an op-amp or an equivalently low impedance. Do not try to drive them through a coupling capacitor as asymmetrical conduction of the diodes can create unwanted DC shifts on the capacitor.

The peak-detect op-amp U4 must be an FET-input type to avoid errors due to bias currents flowing in the relatively high-value resistors R1–R6, and a cheap TL072 works very nicely here; in fact the resistor values could probably be raised significantly without any problems.

As with other non-linear circuits in this book, everything operates between the two supply rails so unwanted currents cannot find their way into the ground system.

Combined LED Indicators

For many years there has been a tendency towards very crowded channel front panels, driven by a need to keep the overall size of a complex console within reasonable limits. One apparently ingenious way to gain a few more square millimeters of panel space is to combine the signal-present and peak indicators into one by using a bicolor LED. Green shows signal present and red indicates peak. One might even consider using orange (both LED colors on) for an intermediate level indication.

Unfortunately, such indicators are hard to read, even with normal color vision. If you have red–green color-blindness, the most common kind (6% of males, 0.4% of females), they are useless. Combining indicators like this is not really a good idea.

VU Meters

VU meters are a relatively slow-response method of indicating an audio level in 'Volume Units'. The standard VU meter was originally developed in 1939 by Bell Labs and the USA broadcasters NBC and CBS. The meter response is intentionally a 'slow' measurement that is intended to average out short peaks and give an estimate of perceived loudness. This worked adequately when it was used for monitoring the levels going to an analog tape machine, as the overload characteristic of magnetic tape is one of soft compression and the occasional squashing of short transient peaks is hard to detect aurally. Digital recorders overload in

www.raguvaran.puzl.com

a much more abrupt and intrusive manner, making even brief overloads unpleasant, and the use of VU meters in professional audio has been in steady decline for many years.

The specifications, particularly of the dynamics, of a standard VU meter are closely defined in the documents British Standard BS 6840, ANSI C16.5-1942, and IEC 60268-17, but there are many cheap meters out there with 'VU' written on them that make no attempt to conform with these documents. The usual VU scale runs from −20 to +3, with the levels above zero being red; 1 VU is the same change in level as 1 dB. The rise and fall times of the meter are both 300 ms, so if a sine wave of amplitude 0 VU is applied suddenly, the needle will take 300 ms to swing over to 0 VU on the scale. A proper VU meter uses full-wave rectification so asymmetrical waveforms are measured correctly, but the cheap pretenders normally have a single series diode that only gives half-wave rectification. VU meters are calibrated on the usual measure-average-but-pretend-it's-rms basis, so a VU meter gives a true reading of rms voltage level only for a sine wave. Musical signals are usually more peaky than sine waves so the VU meter will usually read somewhat lower than the true rms value.

The 0 VU mark on the meter scale is 'zero-reference level', but what that means in terms of actual level depends on the system to which it is fitted. In professional audio equipment 0 VU is +4 dBu, whereas in semi-pro gear it will be −10 dBv (= −7.8 dBu).

VU meters consist of a relatively low-resistance meter winding driven by rectifier diodes, with a series resistor added to define the sensitivity. The usual value is 3k6, which gives 0 VU = +4 dBu. They therefore present a horribly non-linear load to an external circuit, and a VU meter must never be connected across a signal path unless it has near-zero impedance. This is particularly true for cheap ones with half-wave rectification. In practice a buffer stage is always used between a signal path and the VU meter to give complete isolation and to allow the calibration to be adjusted. Many mixers have a nominal internal level 6 dB below the nominal output level, because the balanced output amplifiers inherently have 6 dB of gain, and so the meter buffer amplifiers must also be capable of giving 6 dB of gain.

Figure 18.3 shows an effective design for a meter buffer stage designed to work with a nominal internal level of −6 dBu, and which must therefore give 10 dB of gain to raise the meter signal to the +4 dBu that will give a 0 VU reading. C1 provides DC blocking, because a VU meter will respond to DC as well as AC. R1 and R2 set the gain range to be +6 to +13 dB, which is an ample range of adjustment. With the preset centralized, the gain is 9.3 dB. The resistor values are unusually high because presenting a high input impedance is here more important than the noise performance. An amplifier that was noisy enough to register directly on a VU meter would probably be better fitted to a life as a white noise generator. R3 is an isolating resistor to make sure that the capacitance of the cable to the VU meter, which may be quite lengthy if the meter is perched up in an overbridge, does not cause instability in the buffer amplifier; its presence in series with the meter resistor R4 is allowed for when the calibration is set. The ground of the meter itself is labeled 'Dirty Ground' to underline the

www.raguvaran.puzl.com

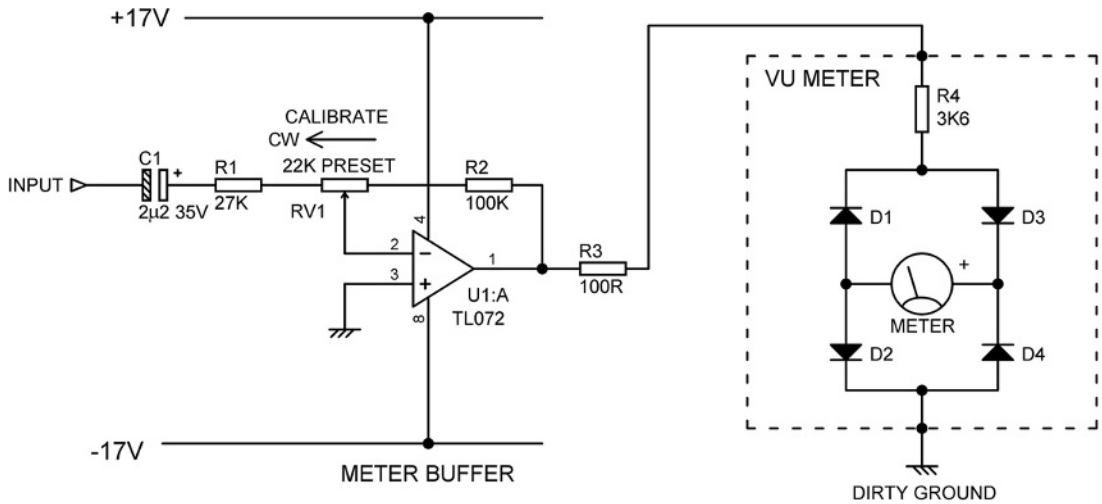


Figure 18.3: A very simple but effective design for a VU meter buffer stage

point that the current through the meter will be heavily distorted by the rectification going on, and must not be allowed to get into the clean audio ground.

Because of their slow response, VU meters are sometimes made with a Peak LED projecting through the meter scale. This is driven by a peak-detect circuit of the sort described earlier in this chapter.

PPM Meters

Peak program meters (PPMs) are essentially peak-reading instruments that respond much more quickly than VU meters. They are always a good deal more expensive, partly because of the precisely defined and rather demanding characteristics of the physical meter itself – for example, the needle has to be able to move much faster than a VU needle, but without excessive overshoot – and partly because they need much more complex drive circuitry. The PPM standard was originally developed by the BBC in 1938, as a response to the inadequacy of existing average-responding meters. PPMs have a distinctive scale with white legends and a white needle against a black background, and are marked from 1 to 7. This is a logarithmic scale giving 4 decibels per division, and the accurate and temperature-stable implementation of this characteristic is what makes the drive circuitry expensive.

Nonetheless PPMs are specifically designed not to catch the very fastest of transient peaks, and are therefore sometimes called ‘quasi-peak’ meters. They only respond to transients sustained for a defined time; the specs give ‘type I’ meters an integration time of 5 ms, while ‘type II’ meters use 10 ms. The result is that transient levels normally exceed the PPM reading by some 4–6 dB. This approach encourages operators to somewhat increase program levels, giving

www.raguvaran.puzl.com

a better signal-to-noise performance. The assumption (which is generally well-founded) is that occasional clipping of brief transients is not audible. The existence of both type I and II meters simply reflects differing views on the audibility of transient distortion.

PPMs exhibit a slow fallback from peak deflections, so it is easier to read peak levels visually. Type I meters should take 1.4–2.0 seconds to fall back 20 dB, while type II meters should take 2.5–3.1 seconds to fall back 24 dB. Type II meters also incorporate a delay of 75–150 ms before the needle fallback is allowed to begin; this peak-hold action makes reading easier.

LED Bar-Graph Metering

Bar-graph meters are commonly made up of an array of LEDs. An LED bar-graph meter can be made effectively with an active-rectifier circuit and a resistive divider chain that sets up the trip voltage of an array of comparators; this allows complete freedom in setting the trip level for each LED. A typical circuit that indicates from 0 to −14 dB in 2-dB steps with a selectable peak or average-reading characteristic is shown in Figure 18.4 and illustrates some important points in bar-graph design.

U3 is a half-wave precision rectifier of a familiar type, where negative feedback servos out the forward drop of D11, and D10 prevents op-amp clipping when D11 is reverse-biased. The rectified signal appears at the cathode of D11, and is smoothed by R7 and C1 to give an average, sort-of-VU response. D12 gives a separate rectified output and drives the peak-storage network R10, C9, which has a fast attack and a slow decay through R21. Either average or peak outputs are selected by SW1, and applied to the non-inverting inputs of an array of comparators. The LM2901 quad voltage comparator is very handy in this application; it has low input offsets and the essential open-collector outputs.

The inverting comparator inputs are connected to a resistor divider chain that sets the trip level for each LED. With no signal input, the comparator outputs are all low and their open-collector outputs shunt the LED chain current from Q1 to −15 V, so all LEDs are off. As the input signal rises in level, the first comparator U2:D switches its output off, and LED D8 illuminates. With more signal, U2:C also switches off and D7 comes on, and so on, until U1:A switches off and D1 illuminates. The important points about the LED chain are that the highest level LED is at the bottom of the chain, as it comes on last, and that the LED current flows from one supply rail down to the other, and is not passed into a ground. This prevents noise from getting into the audio path. The LED chain is driven with a constant-current source to keep LED brightness constant despite varying numbers of them being in circuit; this uses much less current than giving each LED its own resistor to the supply rail, and is universally used in mixing console metering. Make sure you have enough voltage headroom in the LED chain, not forgetting that yellow and green LEDs have a larger forward drop than red ones. The circuit shown has plenty of spare voltage for its

www.raguvaran.puzl.com

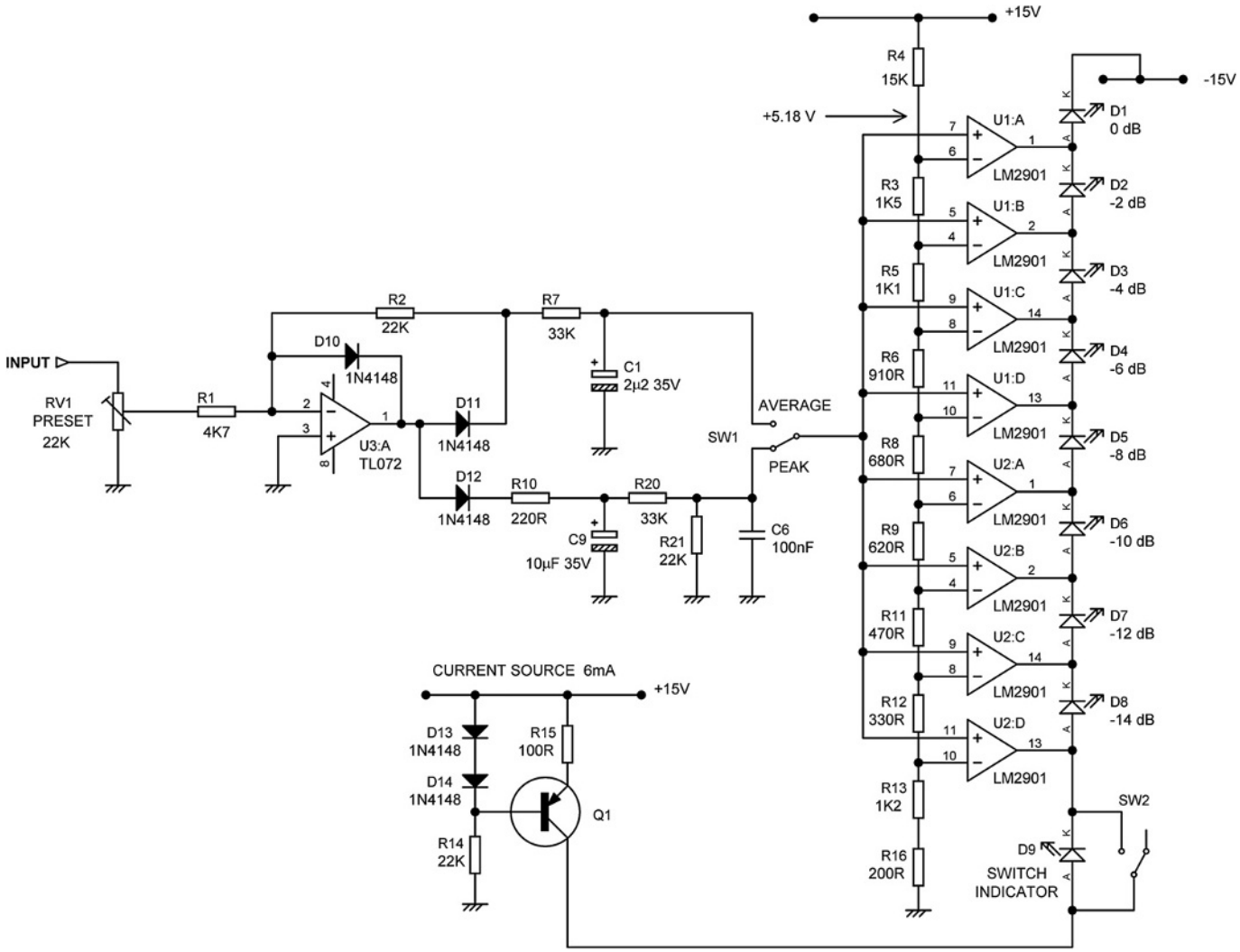


Figure 18.4: LED bar-graph meter circuit with selectable peak/average response

www.raguvaran.puzl.com

LED chain, and so it is possible to put other indicator LEDs in the same constant-current path; for example, D9 can be switched on and off completely independently of the bar-graph LEDs, and can be used to indicate Channel On status or whatever. An important point is that in use the voltage at the top of the LED chain is continually changing in 2-V steps, and this part of the circuit must be kept away from the audio path to prevent horrible crunching noises from crosstalking into it.

This meter can of course be modified to have a different number of steps, and there is no need for the steps to be the same size. It is as accurate in its indications as the use of E24 values in the resistor divider chain allows.

If a lot of LED steps are required, there are some handy ICs that contain multiple open-collector comparators connected to an in-built divider chain. The National LM3914 has 10 comparators and a divider chain with equal steps, so they can be daisy-chained to make big displays, but some law bending is required if you want a logarithmic output. The National LM3915 also has 10 comparators, but a logarithmic divider chain covering a 30 dB range in 3-dB steps.

A More Efficient LED Bar-Graph Architecture

The bar-graph meter shown in Figure 18.4 above draws 6 mA from the two supply rails at all times, even if all the LEDs are off for long periods, which is often the case in recording work. This is actually desirable in a simple mixer as the ±15 or ±17 V rails are also used to power the audio circuitry, and step-changes in current taken by the meter could get into the ground system via decoupling capacitors and suchlike, causing highly unwelcome clicks.

In larger mixers a separate meter supply is provided to prevent this problem, and this allows more freedom in the design of the meter circuitry. In the example I am about to recount, the meter supply available was a single rail of +24 V; this came from an existing power supply design and was not open to alteration, negotiation or messing about with. A meter design with 20 LEDs was required, and an immediate problem was that you cannot power 20 LEDs of assorted colors with one chain running from +24 V; two LED chains would be required and the power consumption of the meter, even when completely dormant, would be twice as great. I therefore devised a more efficient system, which not only saves a considerable amount of power, but also actually economizes on components.

The meter circuit is shown in Figure 18.5, and I must admit it is not one of those circuit diagrams where the *modus operandi* exactly leaps from the page. However, stick with me.

There are two LED chains, each powered by its own constant-current source Q1, Q2. The relevant current source is only turned on when it is needed. With no signal input, all LEDs are

www.raguvaran.puzl.com

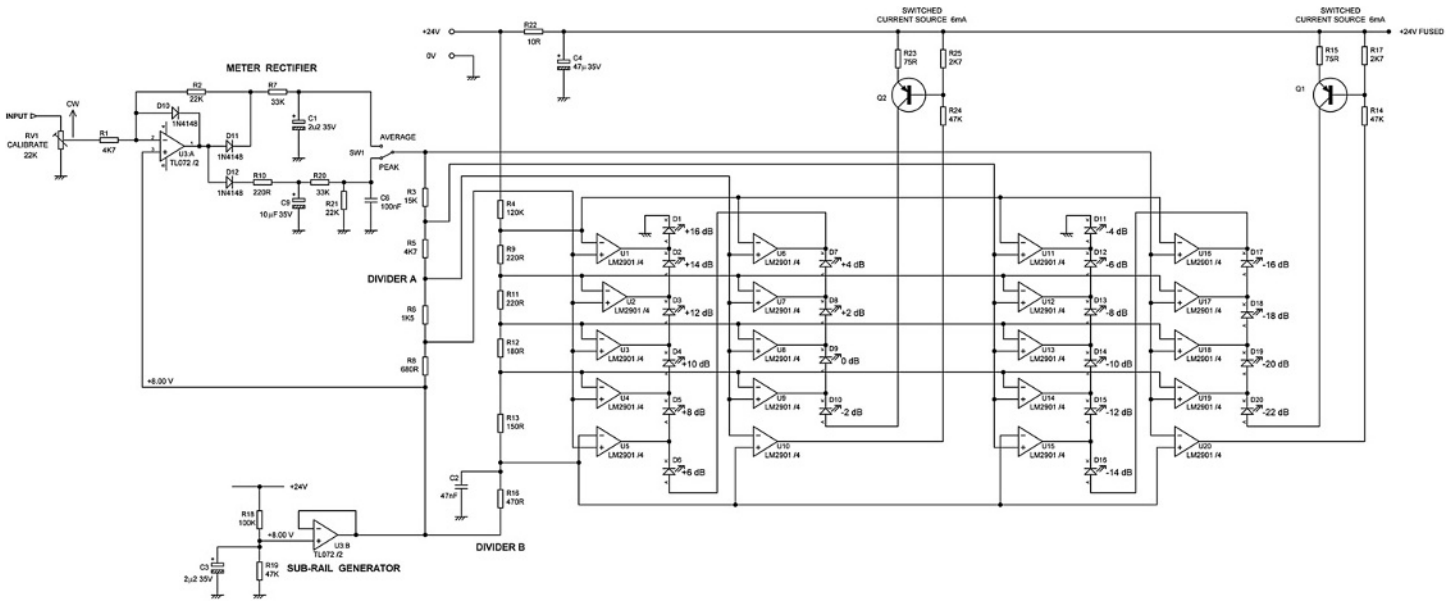


Figure 18.5: A more efficient LED bar-graph meter

www.raguvaran.puzl.com

off; the outputs of comparators U10 and U20 are high (open-collector output off) and both Q1 and Q2 are off. The outputs of all other comparators are low. When a steadily increasing signal arrives, U20 is the first comparator to switch, and LED D20 turns on. With increasing signal, the output of U19 goes high, and the next LED, D19, turns on. This continues in exactly the same way as the conventional bar-graph circuit described above until all the LEDs in the chain D11–D20 are illuminated. As the signal increases further, comparator U10 switches and turns on the second current source Q2, illuminating D10; the rest of the LEDs in the second chain are then turned on in sequence as before. This arrangement saves a considerable amount of power, as no supply current at all is drawn when the meter is inactive, and only half the maximum is drawn so long as the indication is below −2 dB.

There are 10 comparators for each LED chain, 20 in all, so a long potential divider with 21 resistors would be required to provide the reference voltage for each comparator if it was done in the conventional way, as shown in Figure 18.4. However, looking at all those comparator inputs tied together, it struck me there might be a better way to generate all the reference voltages required, and there is.

The new method, which I call a 'matrix divider' system, uses only 10 resistors. This is more significant than it might at first appear, because the LEDs are on the edge of the PCB, the comparators are in compact quad packages, and so the divider resistors actually take up quite a large proportion of the PCB area. Reducing their number by half made fitting the meter into a pre-existing and rather cramped meter bridge design possible without recourse to surface-mount techniques. There are now two potential dividers. Divider A is driven by the output of the rectifier circuit, while divider B produces a series of fixed voltages with respect to the +8.0 V sub-rail. As the input signal increases, the output of the meter rectifier goes straight to comparators U16–U20, which take their reference voltages from divider B and turn on in sequence as described above. Comparators U11–U15 are fed with the same reference voltages from divider B, but their signal from the meter rectifier is attenuated by divider A, coming from the tap between R3 and R5, and so these comparators require more input signal to turn on. This process is repeated for the third bank of comparators U6–U10, whose input signal is further attenuated, and finally for the fourth bank of comparators U1–U5, whose input is still further attenuated. The result is that all the comparators switch in the correct order.

Since in this application there was a only single supply rail, a bias generator is required to generate an intermediate sub-rail to bias the op-amps. This sub-rail is set at +8.0 V rather than V/2, to allow enough headroom for the rectifier circuit, which produces only positive outputs; it is generated by R18, R19 and C3, and buffered by op-amp section U3:B. The main +24 V supply is protected by a 10 Ω fusible resistor R22, so if a short-circuit occurs on the meter PCB the resistor will fail open and the whole metering system will not be shut down. This kind of per-module fusing is very common and very important in mixer design; it localizes a possibly disabling fault to one module, and avoids having the power supply shut down, which would put

www.raguvaran.puzl.com

the whole mixer out of action. A small but vital point is that the supply for divider B is taken from outside this fusing resistor; if it was not the divider voltages would vary with the number of LED chains powered, upsetting meter accuracy.

Once again the LM2901 quad voltage comparator is used, as it has low input offset voltages and the requisite open-collector outputs. Q1, Q2 can be any TO-92 devices with reasonable beta; their maximum power dissipation, which occurs with only one LED on in the chain, is a modest 128 mW. This meter system was used with great success.

Vacuum Fluorescent Displays

Vacuum fluorescent displays are sometimes used but require hefty tooling charges if you want a custom display, and their high-voltage operation makes driving them more complicated. The main power supply is usually 250 V DC, a voltage that requires considerable respect.

LCD Meter Displays

The ultimately versatile metering system is provided by making the meter display from a number of color LCD display screens. These can be of the size used in the smaller laptop computers, butted side-to-side to make something like a conventional meterbridge, or larger screens that can show three or more rows of bar-graphs at once. All the Calrec digital consoles currently have such metering. The advantages are obviously that you can display any kind of metering that you can think up, and the technology can be based on standard graphics cards.

www.raguvaran.puzl.com

This page intentionally left blank

www.raguvaran.puzl.com

CHAPTER 19

Level Control and Special Circuits

Gain-Control Elements

A circuit block that gives voltage control of gain is a difficult thing to implement. The basic function is simply that of multiplication, but most of the practical techniques introduce significant distortion. In every case distortion and noise can be traded off by altering the operating level, but in some techniques the non-linearity is so great that a compromise that meets modern performance standards is simply not possible.

A Brief History

The difficulty of implementing a good voltage-variable gain element is testified to by the number of technologies that have been tried. A variable attenuator could be made by varying the current flow through diodes, which varied their effective resistance; the control signal inevitably got mixed up with the audio signal and linearity was poor. Optical combinations of filament bulbs and cadmium sulfide photoresistors gave very good control-signal isolation but operation was slow because of the thermal inertia of the filament. Later versions used LEDs and were much faster but the linearity of the photoresistors remained a limitation. Chopper systems, which turned the signal hard on and off at an ultrasonic frequency, with a variable mark-space ratio, and then reconstructed the signal with a low-pass filter, gave fast response and reasonable linearity. However, it was difficult to get a wide control range and effectively removing all the switching frequencies required a fairly complicated filter. A long time ago I designed just such a variable attenuator, using 4016 analog gates for the switching, and a clock frequency of 160 kHz. The gain was difficult to control below −40 dB and it never made it to production, which on the whole was probably just as well.

JFETs

The introduction of JFETs (junction FETs as opposed to MOSFETs) as gain-control devices was a great advance. These devices promised good isolation between control voltage and signal, effectively instantaneous operation, and lower distortion than existing methods. These advantages do exist, but there are some less desirable features, such as a highly non-linear

Small Signal Audio Design; ISBN: 9780240521770
Copyright © 2010 Elsevier Ltd; All rights of reproduction, in any form, reserved.

www.raguvaran.puzl.com

gain/CV law that varies significantly from specimen to specimen, as a result of process tolerances. Signal-CV isolation is absolute in the DC sense, the gate looking like a very-low-leakage reverse-biased diode, but there is always some gate-channel capacitance, which means that fast edges on the control voltage can get through to the signal path.

The use of JFETs for on/off signal control rather than voltage-controlled attenuation is covered in Chapter 16 on signal switching.

A JFET used as a voltage-controlled attenuator is operated below pinch-off, i.e. at low values of V_{ds} that allow it to operate more like a resistor than a constant-current source. Some JFET types are better than others for this job, the 2N5457 and the 2N5459 being particularly favored in the mid-1970s.

Figure 19.1 shows the most basic voltage-controlled JFET attenuator circuit. When N-channel FETs are used, as is normally the case, the control voltage must go negative of ground to turn off the JFET and give the minimum-attenuation condition. The maximum attenuation is limited by the $R_{ds(on)}$ of the JFET to about 45 dB for practical circuitry. The attenuation law is highly non-linear, and variable between specimens of the same JFET; to some extent this can be trimmed out by applying a constant DC bias to the control voltage, but minor variations in the shape of the control law still remain. One of the advantages of this circuit is that when there is no gain reduction, the JFET is biased hard off and there is no extra distortion introduced, though there will be Johnson noise from the series resistor. This makes JFETs useful in limiter applications where gain reduction occurs only briefly and intermittently; JFETs are much cheaper than voltage-controlled amplifiers.

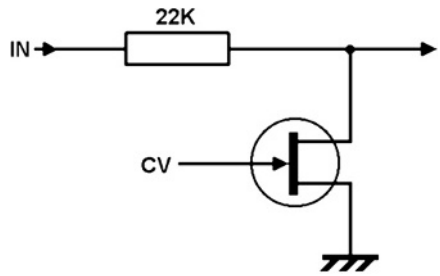


Figure 19.1: The basic voltage-controlled JFET attenuator circuit

With the JFET conducting, the major non-linearity is second-harmonic distortion, which can be much reduced by adding half the drain-source voltage to the gate control voltage, as in Figure 19.2. Note that it is half the drain-source voltage that is applied, not half of the input signal. The ratio is not critical – getting it right within 10% seems to give all the linearization available. A serious problem with this circuit is the control-voltage feedthrough into the signal path. One answer is to add a DC-blocking capacitor C as in Figure 19.3; this stops DC, but does nothing to stop fast transient feedthrough.

www.raguvaran.puzl.com

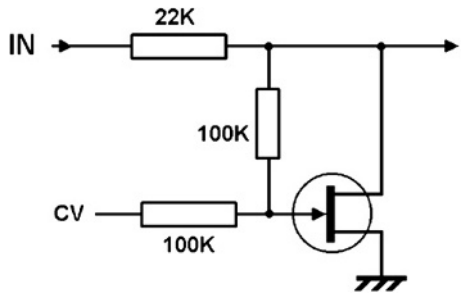


Figure 19.2: Second-harmonic distortion can be much reduced by adding half the drain-source voltage to the gate-control voltage

Another problem is that the presence of C adds an extra low-pass time-constant to the applied control voltage. In feedback limiters this can cause instability of the control loop, typically causing the gain reduction to suddenly snap to the maximum value, followed by a slow decay back to normal operating conditions. This is not a good thing. Both problems can be solved by using a unity-gain buffer amplifier to isolate the JFET gate network from the signal path, as shown in Figure 19.4.

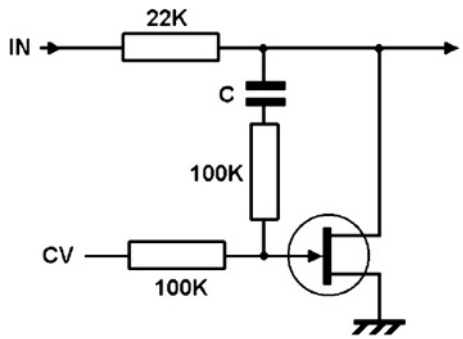


Figure 19.3: Adding blocking capacitor C prevents DC getting into the signal path, but does nothing to stop transient feedthrough

Because of their unhelpful control laws, JFETs were most useful in feedback-type compressors and limiters, where the feedback loop linearized the law (more on that later). I produced a compressor/limiter design for *Wireless World* [1] when the application of FETs in this way was relatively new; it was published some years after I designed it. The linearity limitation remains, and JFETs are now rarely used in compressors and limiters (except for those which deliberately embrace obsolescent technologies), having been replaced by voltage-controlled amplifiers (VCAs). They are, however, still useful in noise gates, because as shown above, they can be configured so that when the noise gate is open, the JFET is firmly off and introduces no signal degradation at all. A JFET is also much cheaper than a VCA.

www.raguvaran.puzl.com

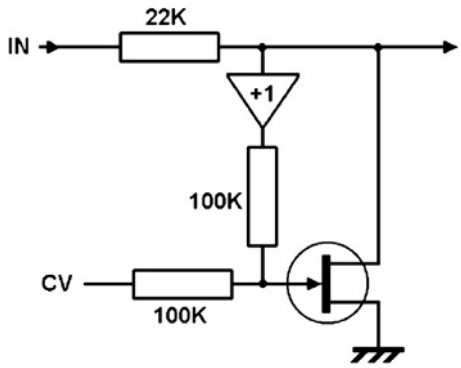


Figure 19.4: Using an active buffer instead of a blocking capacitor prevents both DC and transient feedthrough

Operational Transconductance Amplifiers

An operational transconductance amplifier (OTA) gives a current output for a voltage difference input, and the amount of current you get for a given input voltage is determined by the current that is made to flow into a control port, giving a variable-gain capability. The output stage is a high-impedance current source rather than the low-impedance voltage source of the conventional op-amp, and the output current must be converted to a voltage, usually using a simple resistive load. Buffering is then needed to give a low-impedance output. Despite the name – operational transconductance amplifier – this device is not used like a conventional op-amp when it is being used to give variable gain. There is no negative feedback around the device, as this would prevent the gain varying, and for acceptable linearity the differential input voltages should be kept to 20 mV or less. Distortion is mostly third harmonic, and comes from the input pair transistors.

The best-known operational transconductance amplifier was the CA3080E, and a typical voltage-controlled gain circuit for it is shown in Figure 19.5. A1 is the OTA, R1 and R2 reduce the input level so it is suitable for the device input transistors, R3 is the *I/V* conversion resistor, and A2 is a conventional op-amp acting as the output buffer. The THD is about 0.15% for a +5 dBu input level. Note the OTA symbol has a current source symbol attached to its output.

Another OTA called the LM13600, was introduced later by National Semiconductor; this is a dual part and has distortion-compensation diodes in the input stage, which are claimed to have improved linearity fourfold. I have to say that I found that the difference in practical applications was not that large. It also has built-in Darlington output buffers. The LM13600 has now been replaced by the LM13700; the CA3080E went out of production in 2005. The LM13700 is used in much the same way as the CA3080E.

www.raguvaran.puzl.com

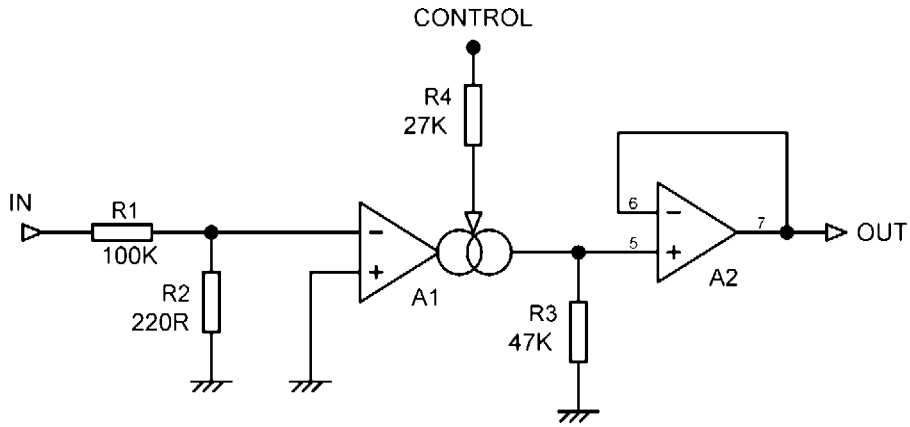


Figure 19.5: A typical gain-control circuit using the 3080E operational transconductance amplifier

You are probably thinking by now that transconductance amplifiers belong in the history section of this chapter; they are placed here because their flexibility has kept them very popular with builders of analog synthesizers.

Voltage-Controlled Amplifiers

Voltage-controlled amplifier (VCA) is the name given to a specific kind of variable-gain device. It is essentially a four-quadrant multiplier, and is a current-in, current-out device, which therefore needs some support circuitry if it is to interface with the normal voltage-driven world.

Since the input port is a virtual earth, converting the input voltage into an input current requires only a resistor (see R1 in the basic circuit of Figure 19.6). Converting the output current into an output voltage is less easy and requires a shunt-feedback amplifier A1, as it is essential to avoid signal voltages on the output port. This is a transadmittance amplifier (current in, voltage out). An important advantage of this circuit is that is does *not* phase-invert. It looks as if it

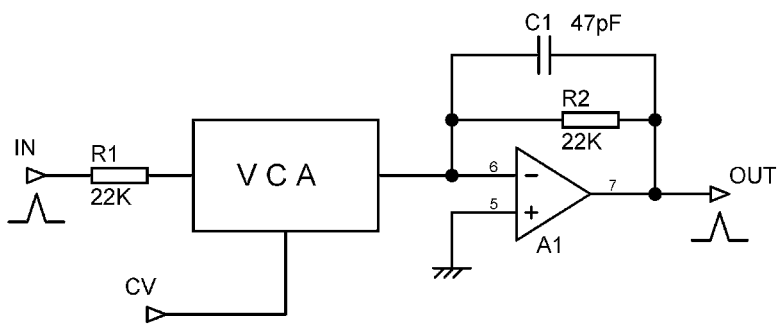


Figure 19.6: The basic gain-control circuit using a VCA

www.raguvaran.puzl.com

does – the amplifier A1 certainly inverts – but so does the VCA internally, so the output remains in phase, which is of course essential for mixing console use.

Because of their log/antilog operation VCAs are characterized by an exponential control characteristic, so gain varies directly in decibels with control voltage. This is apparently very convenient as it means that a linear fader can be used to control gain over a wide range; however, in practice this is less than ideal as it spreads out the less important higher attenuation range too much. Special 'VCA-law' faders are made that squeeze the high attenuation range so the normal operating range can be expanded.

The evolution of VCA technology starts in about 1970, based on the 'Blackmer gain cell' developed by David Blackmer of dbx Inc., and VCA history is a fascinating field in itself. Another of the very early models was the Allison EGC-101, which gave improved linearity through Class-A operation [3]. The study of the internal operation of VCAs is also a big subject, and regrettably I don't have the space to go into the details here. An article in *Studio Sound* [4] gives a good deal of information on the internals. Other useful references on the technology are Refs [2], [5], and [6].

For many years the 'standard' VCA was the DBX2150, of which I have deployed more than I care to contemplate in VCA subgroup systems and console automation; more modern ones are represented by parts like the THAT 2181. A typical application circuit suitable for both is shown in Figure 19.7.

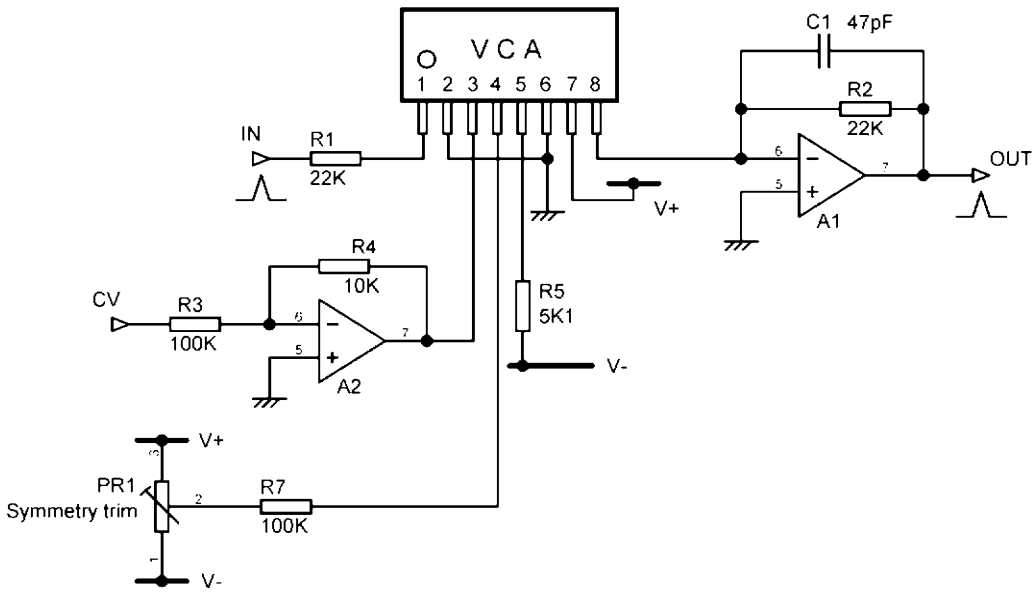


Figure 19.7: A typical gain-control circuit using the 2150 VCA. The eight-pin SIP package is almost always used for VCAs

www.raguvaran.puzl.com

You will note that a symmetry trim pot is required; this is set to minimize second harmonic generation. A THD analyzer is required to make this adjustment, but on the positive side, once set it stays set, and need never be touched again unless the VCA is replaced. Modern VCAs are very good; the THD at 1 Vrms with 0 dB gain can be as low as 0.002%. The off-isolation (or 'offness', as I prefer to call it) can be as good as −110 dB at 1 kHz, but will almost certainly be worse at higher frequencies as it depends on stray capacitance. This is why the SIP package is preferred; as shown in Figure 19.7, it keeps the input and output pins as far apart as possible. A crucial layout requirement is that the 'hot' end of R1 in Figure 19.7 is kept as far away as practicable from pin 8 of the VCA, even if it means extending the track between R1 and VCA pin 1. This track is at virtual earth and susceptible to capacitative crosstalk so keep unrelated signals well away from it.

The control law (which is set by transistor physics, and is therefore dependable) is 6 mV/dB at the actual VCA control pin, and this is inconveniently low. A scaling amplifier A2 is therefore used so the control voltage has a more useful range, such as 0–10 V. The presence of this amplifier also gives the opportunity for control-voltage filtering and changes of ground reference. The voltage applied to the control port must be at low impedance (basically an op-amp output is the only source that will do) and absolutely free from contamination by any sort of signal or avoidable noise. Contamination with even a trace of the signal being controlled will cause excess distortion.

The capacitor C1 shown across the *I*/*V* conversion op-amp feedback resistor is always required for HF stability, due to the destabilizing capacitance seen looking into the VCA output port.

In other places in this book I have described how noise can be reduced by using multiple transistors or multiple op-amps in parallel. This also works with VCAs; you simply put *N* of them in parallel and connect their outputs to a single shunt-feedback amplifier with a suitably reduced feedback resistor, as shown in Figure 19.8. Two VCAs give a 3 dB improvement and four a 6 dB benefit. There are of course limits as to how far you can go with this sort of thing. High-quality VCAs are relatively expensive, and to the best of my knowledge never more than four have been used in parallel. The technique is used in high-quality compressor/limiters such as the Connor.

Compressors and Limiters

Compressors and limiters are devices that control the dynamic range of a signal. The device acts like a rapid volume control that reduces the gain when the signal level becomes excessive.

A compressor reduces the general dynamic range of a signal for the majority of the time. A typical application is control of microphone levels when the talent does not make microphone technique their top priority. This means that it is applying some amount of gain reduction most

www.raguvaran.puzl.com

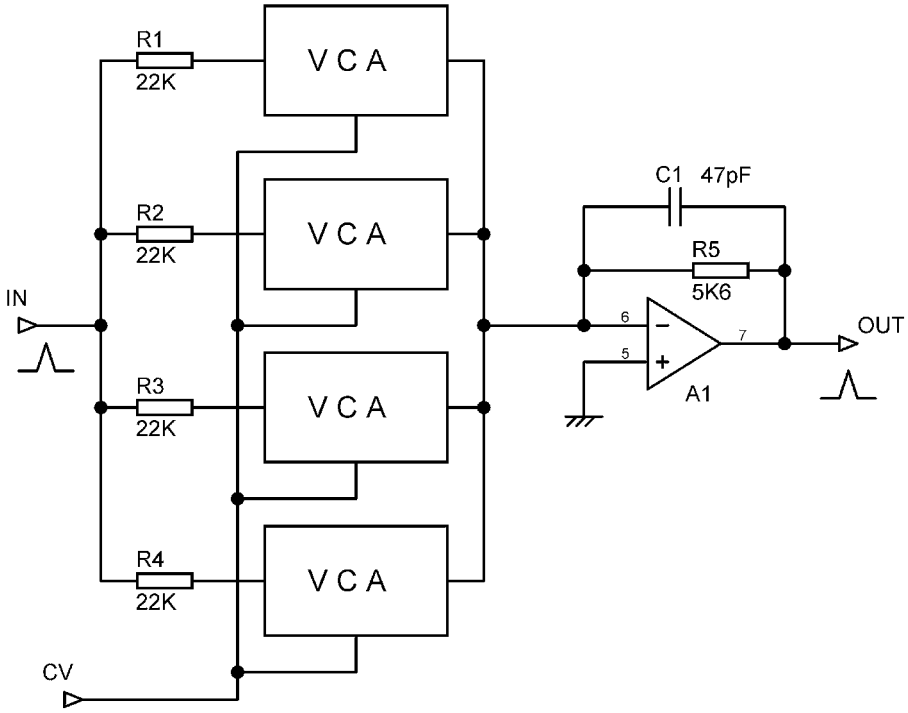


Figure 19.8: Using multiple VCAs in parallel to reduce noise. Each doubling of numbers theoretically reduces noise by 3 dB

of the time, so the gain-control element is active and it is important that it is relatively distortion-free and without other audible defects.

A limiter, in contrast, has as its main function the prevention of overload and horribly audible clipping. When loudspeakers and AM transmitters are being driven it may actually prevent expensive equipment damage. Transmitters require the extra protection of a clipper circuit (see the separate section on these). Since a limiter operates relatively rarely (assuming the system is being operated correctly) it is less important that the gain-control element is distortion-free.

Figure 19.9 shows the relationships between input and output levels for compression and limiting. The compressor law begins gain reduction at a lower threshold level, and has a moderate slope thereafter; the slope shown here is 3:1, which is typical. The limiter law has a higher threshold and only acts when overload is imminent; the slope is much flatter so that even very high signal levels cannot reach the overload point; the slope here is 10:1.

Compressors and limiters do a similar job, using much the same hardware with different parameter settings, so the functions are often combined in a compressor/limiter. Figure 19.10 gives the block diagram of a feedback compressor/limiter. The output signal is amplified, and if it exceeds a certain threshold, applied to a rectifier with a fast-attack, slow-decay

www.raguvaran.puzl.com

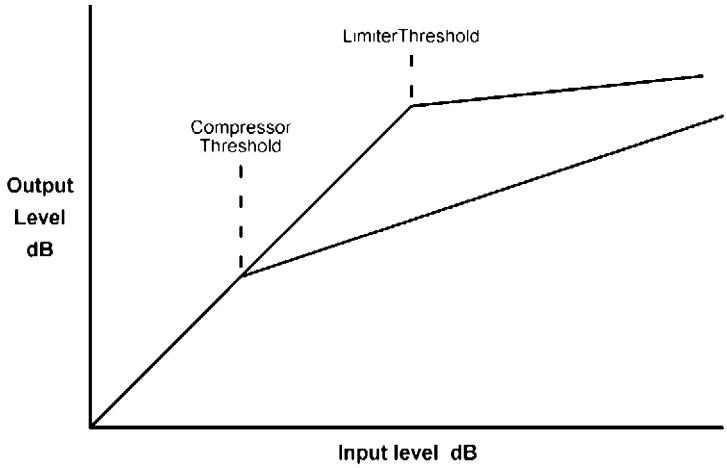


Figure 19.9: Compressor and limiter laws

characteristic. This part of the system is called the side-chain, to emphasize that the signal does not pass through it. The resulting control voltage is applied to the gain-control element and as the signal level increases the gain is reduced, reducing the variations in output level. The side-chain may use either peak or rms sensing; the latter is considered by some to relate better to our perception of loudness and give a less obtrusive effect.

The laws shown in Figure 19.9 have 'hard knees' as the gain laws change abruptly at the threshold. A 'soft knee' slowly increases the compression ratio as the level increases, giving a curve that gradually attains the desired compression ratio. A 'soft knee' is considered to make the change from uncompressed to compressed less audible, especially for higher compression ratios.

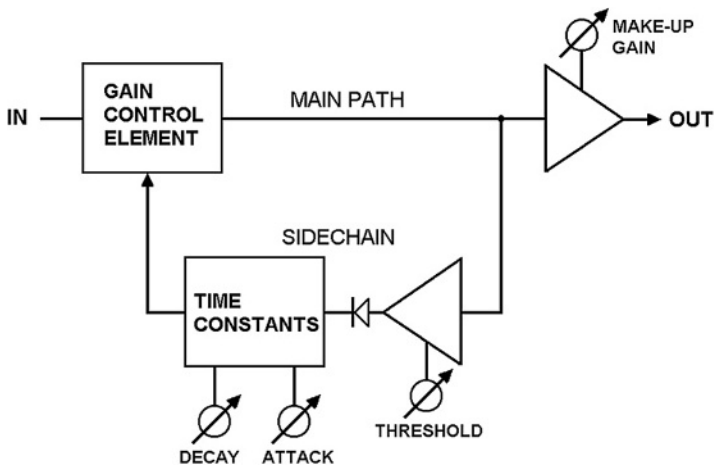


Figure 19.10: Block diagram of a feedback compressor/limiter

www.raguvaran.puzl.com

If the threshold is set low, and the side-chain gain also low, the system works as a compressor, the gain reduction acting to reduce the general dynamic range of the output signal, hopefully without obvious side-effects. If the threshold is set high, and the side-chain gain is also high, we have instead a limiter. The signal will be untouched until it exceeds the threshold, and then gain reduction is applied strongly to prevent the output level significantly increasing.

Compressor/limiters can work in either feedforward or feedback modes, each of which has its own advantages and problems. In the feedback mode the compression law tends to be inherently linear because of the feedback around the level-control loop. The feedback configuration has the advantage that the output level set does not depend on the control-voltage law of the gain element.

With feedforward operation, as in Figure 19.11 the compression law depends on the control-voltage law of the gain element. If this is a VCA with a very predictable law, no problem. However, other gain elements such as FETs sometimes have their uses – with FETs the advantage is that there is no signal degradation when there is no gain reduction. There is, however, also a highly non-linear control-voltage law to deal with, and it is not practical to bend this into a more desirable log law.

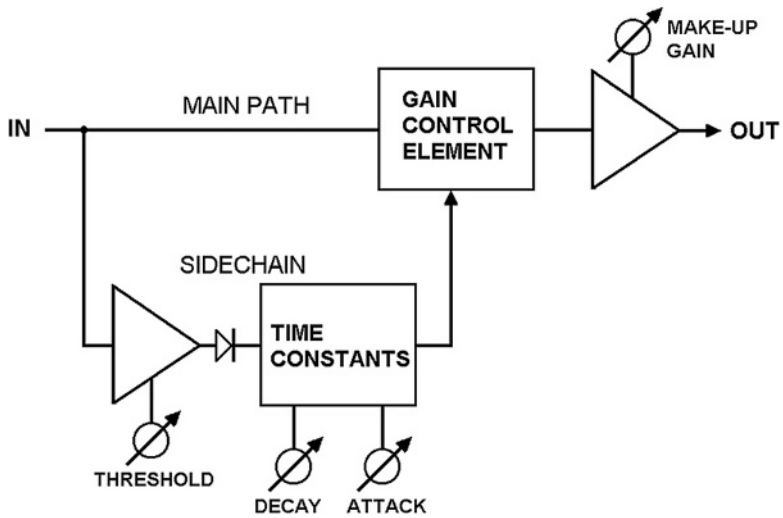


Figure 19.11: Block diagram of a feedforward compressor/limiter

The way to solve this is shown in Figure 19.12, which assumes that two matched gain elements are easier to make than one with the required law. The side-chain path contains a feedback limiter and the control voltage this generates is fed to the second gain element, which acts in feedforward mode. Typically the gain elements are matched as well as possible by trimming.

www.raguvaran.puzl.com

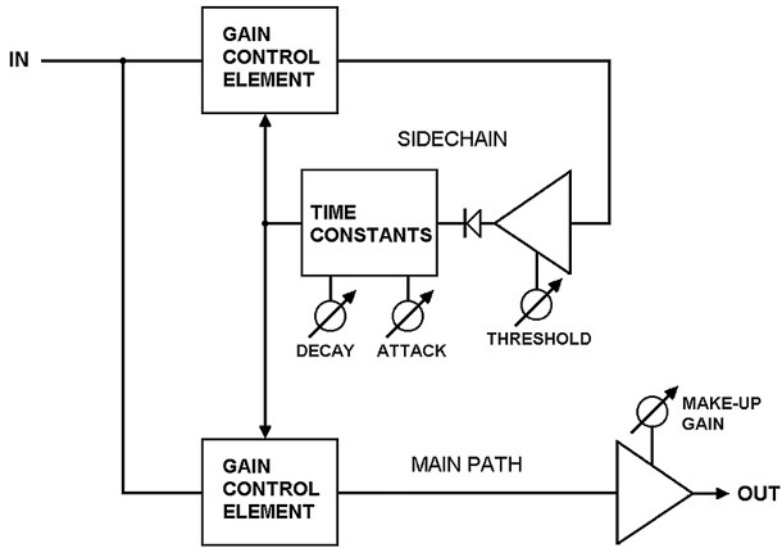


Figure 19.12: A combined feedforward/feedback compressor/limiter

So far we have assumed that the side-chain dynamics consist only of two simple time-constants for attack and decay. In many cases this will give rise to objectionable effects, and more sophisticated control measures have been developed to deal with the problems. Even if a gain-control element is completely linear when the control voltage is fixed, rapid CV changes put distortion into the waveform.

Attack Artefacts

When a fast attack is used to control level (typically when the unit is being used as a limiter) and prevent overshoot, it tends to bite chunks out of the controlled waveform, as a natural consequence of a rapid drop in gain. Isolated cases of this usually pass unnoticed, but repeated occurrences are perceived as a crackling noise. This can be controlled by the use of dual attack times, with the faster time switched in if the signal peak exceeds criteria for amplitude and rate of rise, but the only complete solution to this problem is a delay-line compressor/limiter, as seen in Figure 19.13.

A delay in the signal path between the side-chain take-off point and the gain element allows a feedforward side-chain to react relatively slowly and still have the gain reduced before the signal peak reaches the gain-control element. (This is not possible with the feedback compressor/limiter configuration because the side-chain feed is taken *after* the gain-control element, and therefore after the delay.) The difficulty is that the signal is delayed at all times; this will cause problems if it is being mixed with undelayed signals, and furthermore the delay section has to be of high quality as it passes the main signal, not just the side-chain information. A well-known BBC design of *circa* 1967 used a strictly analog 320 μs delay line

www.raguvaran.puzl.com

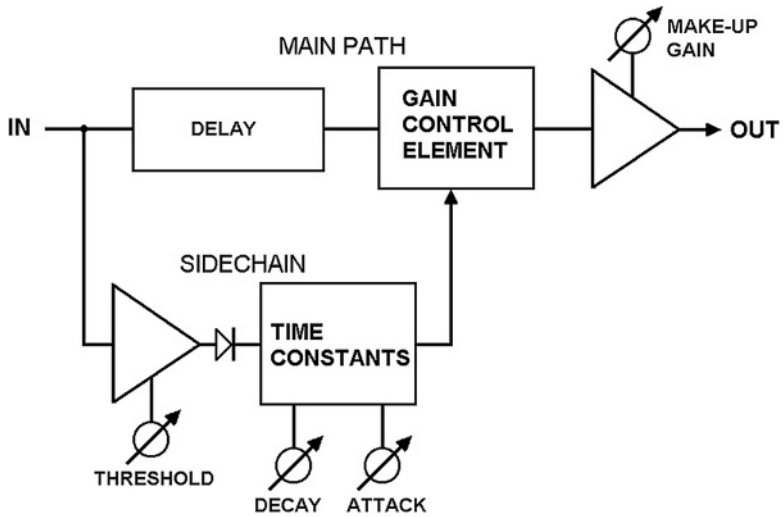


Figure 19.13: Block diagram of a feedforward compressor/limiter with delay

made up of 10 LCR second-order all-pass filter sections [7]. This worked very well but is obviously an expensive technique. An active filter delay line would be cheaper but still involves a significant amount of circuitry and a number of close-tolerance components. Nowadays a high-quality digital delay using 24-bit converters can be constructed relatively easily.

Decay Artefacts

Since the side-chain contains either a peak or rms detecting rectifier, the timing capacitor will have on it a ripple waveform resulting from cyclical charge and discharge, exactly as in a power-supply reservoir capacitor. This waveform modulates the signal path gain and generates distortion; the effect can be severe. In most compressor/limiter usage the attack time has to be fast, but often the decay time can be made much longer, reducing the amplitude of the ripple and reducing distortion. However, in some applications the decay has to be short for rapid recovery from brief level excursions, or the general level of the signal will be unduly depressed. An effective and widely adopted solution to this problem is the use of dual time-constants. See Figure 19.14 for a simple implementation, where the short time-constant R1, C1 responds to short-term level changes while the long time-constant R2, C2 copes with the general trend.

Another elaboration is a hold circuit that prevents the decay from beginning if any attack events have occurred in, say, the preceding 20 ms. This helps to prevent sudden and obvious gain changes, often called by the highly descriptive term ‘pumping’. If the noise background is significant, as with many live news recordings, then even slow gain changes produce a distracting effect called ‘breathing’ because that’s just what it sounds like.

www.raguvaran.puzl.com

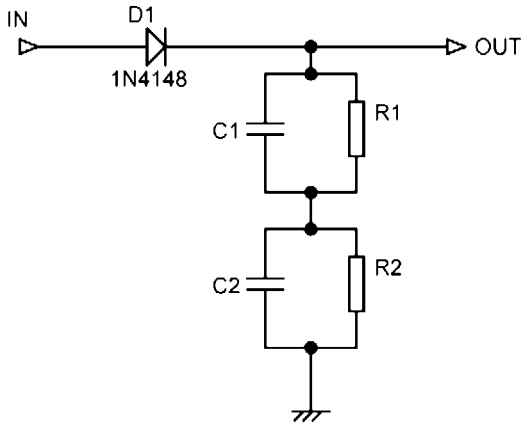


Figure 19.14: Dual decay time-constant circuit

Subtractive VCA Control

While modern VCAs are very good, they still introduce more noise and distortion than a typical well-designed chain of 5532 op-amps and, as mentioned above, FETs are therefore still favored for some compressor/limiter applications as they introduce no degradation when not bringing about gain reduction. However, the predictable characteristics of VCAs are very attractive, and it would be nice to combine the advantages of both.

This can be done; whether I invented the technique first I have no idea, but it was all a long time ago. The concept, as shown in Figure 19.15, is not to send the signal through a VCA with 0 dB gain when no gain reduction is required, but to have the VCA normally hard off, and then turning it on to reduce the gain. When the VCA is off, the circuit acts as a simple unity-gain inverting amplifier, and there is minimal signal degradation.

Note that point about inverting – this circuit phase-inverts although the standard VCA circuit does not. As the VCA gain is turned up, it lets more signal current through to the summing

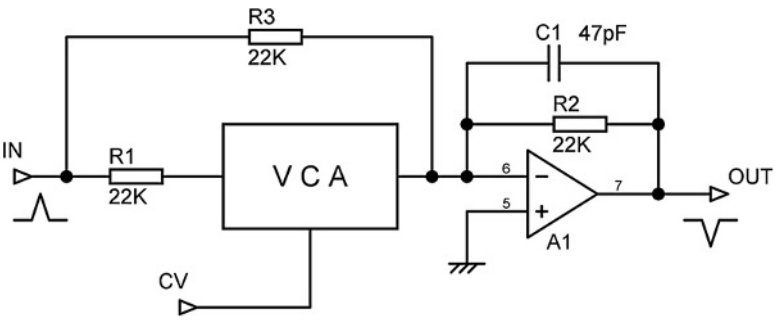


Figure 19.15: A VCA gain control arranged so that no signal passes through it when no gain reduction is occurring

www.raguvaran.puzl.com

point of A1, and since this is out of phase, due to the internal inversion of the VCA, there is partial cancellation and the overall gain falls. When the VCA is set to 0 dB gain, theoretically there is complete cancellation and the overall circuit is off. In practice the tolerances in R1, R3 and in the VCA set a limit to the practicable minimum gain, but the range is more than enough for effective dynamics control.

You may have spotted the potential snag. If the gain of the VCA goes above 0 dB (which corresponds with a control voltage of 0 V), the overall gain will begin to rise again. With a feedback-type compressor/limiter this could be disastrous as the side-chain will try to reduce the gain, but it will simply increase it further, and the system will latch up solid. This can be prevented by the simple expedient of using an active-clamp circuit to prevent the control voltage rising even a millivolt above 0 V. Note that because of the subtractive action the gain law is no longer linear in decibels, and so this plan is better suited to feedback compressor/limiters. I put this system into a broadcast console in 1990, and it worked very well.

Noise Gates

A noise gate allows a signal through only when it is above a set threshold: the gate is said to be open or on. When the signal level drops below the threshold the signal is either attenuated or stopped altogether, and the gate is closed or off. For effective noise reduction the level of the signal must be above that of the noise; the threshold is set above the noise level so when there is no signal the gate closes. Noise gates are not actually often used to discriminate against noise as such; in recording they are more likely to be used to, for example, increase the isolation between the signals coming from a multi-mic'd drum-kit. They are also used to provide the well-known 'gated reverb' effect where decaying reverb is cut off suddenly by a noise gate.

A noise gate is a rather different animal from a compressor/limiter. Its operation is much more on/off, and long time-constants are not normally used. This means that the non-linearity of an FET during intermediate degrees of gain reduction can be tolerated. In the traditional form of FET noise gate, as shown in Figure 19.16, and based on a design I did some years ago, FET distortion is reduced by attenuating the signal considerably – in this case by 32 dB – before applying it to the FET stage. This is done by R1, R4, and the signal is at a low impedance afterwards to keep noise down. The FET (actually two FETs here, to reduce $R_{ds(on)}$ and give a greater gain change) is at the bottom of the NFB network of a low-noise hybrid amplifier stage that restores the signal to its original level when the FET is on. When the FET is off, the gain is reduced to unity. This would only give an offness of −32 dB, which is not enough, so the network R9, PR1, R8 is added, which lets a little signal through to what is effectively an inverting input to the amplifier stage; when PR1 is correctly adjusted there is effective cancellation and the offness can easily exceed −80 dB. FET distortion is further reduced by adding half the drain-source voltage to the gate-control voltage via buffer A3, as described earlier.

www.raguvaran.puzl.com

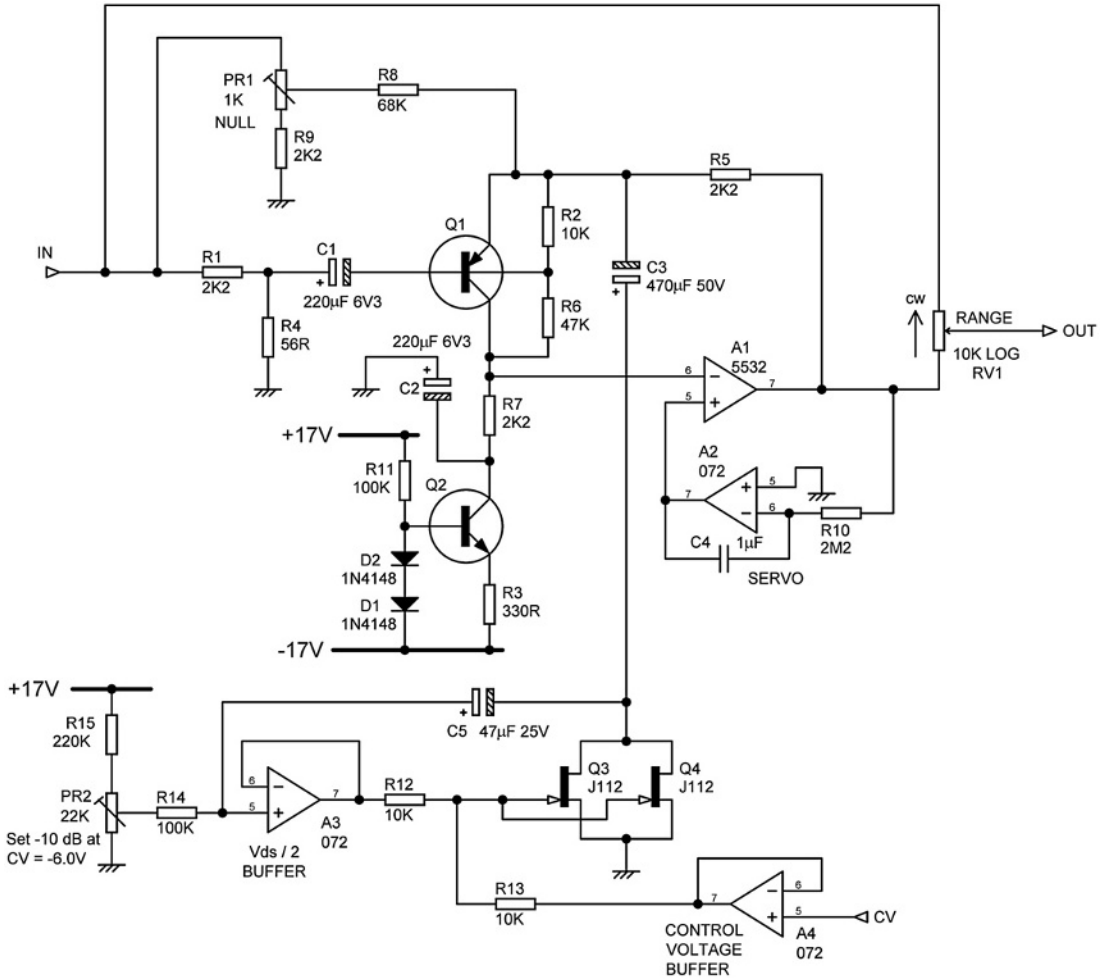


Figure 19.16: A typical FET noise gate

The low-noise amplifier is a hybrid stage, combining the low noise of the discrete input transistor Q1 with the open-loop gain and linearity of op-amp A1. R2, R6 set up the DC conditions for Q1, while servo integrator A2 defines the DC operating point of A1, keeping its output at 0 V on average. Q2 is a current source that helps keep rail noise out of the collector circuit of Q1. The alert reader will spot the similarity between this stage and the moving-coil preamplifier stage in Chapter 8. The noise from the amplifier with the gate off is −106 dBu.

Many noise gates have a 'range control' that allows the offness to be reduced if a more subtle effect is required. This is RV1 in Figure 19.16, an ordinary log pot that has the bottom of its track, which would normally be grounded, connected instead to the noise gate output. Offness is reduced by advancing the control so that some of the signal gets through even when the gate is off.

www.raguvaran.puzl.com

The side-chain of a noise gate is similar to that of a compressor/limiter, but there is a greater emphasis on fast attack times of 50 μs or less, and special circuitry is used to charge the timing capacitor as quickly as possible.

Noise gates can also be made with VCAs, and this is the more usual method when a 'dynamics section', which can act as either a noise gate or a compressor/limiter, is squeezed into a mixer channel.

Clipping

Since clipping and its attendant distortion is something we normally strive to avoid, it may be thought perverse to study ways in which to induce it. Clipping circuits do have their uses, however. One application is in the protection of AM transmitters, where even momentary excursions beyond 100% modulation are unacceptable, because they lead to legal problems (side-band splatter causes the regulator to close you down) and technical problems (the transmitter blows up). Either way you're off the air. Limiters are the first line of defence, but they are liable to overshoot and maladjustment. A fixed-level clipping circuit, however, can absolutely prevent any signal exceeding its threshold. Clipping circuitry also has specialized uses in power amplifier design, where it can emulate the performance of a much more complex and expensive regulated power supply [8] or increase the flexibility of bridged amplifiers [9].

To set down some of the requirements for the perfect clipping circuit:

1. Clipping must be at well-defined symmetrical levels, constant with frequency and not dependent upon signal history.
2. The top of the clipped waveform should be absolutely flat, to give tight level control.
3. The clipping level must be arbitrarily settable, without steps.
4. There must be absolutely no degradation of the signal below the clipping threshold – or at any rate no more than would be caused by going through an ordinary single op-amp stage.

This last requirement is actually much more demanding than it appears, but it is essential. Let us examine the problem. I appreciate that clean audio clipping may be a bit of a minority interest, but it will be highly instructive to see just what difficulties arise, and how they can be overcome.

Diode Clipping

If you look up 'clipping circuits' in the average textbook, you will probably find something like Figure 19.17, where back-to-back diodes conduct when the signal exceeds their conduction threshold.

www.raguvaran.puzl.com

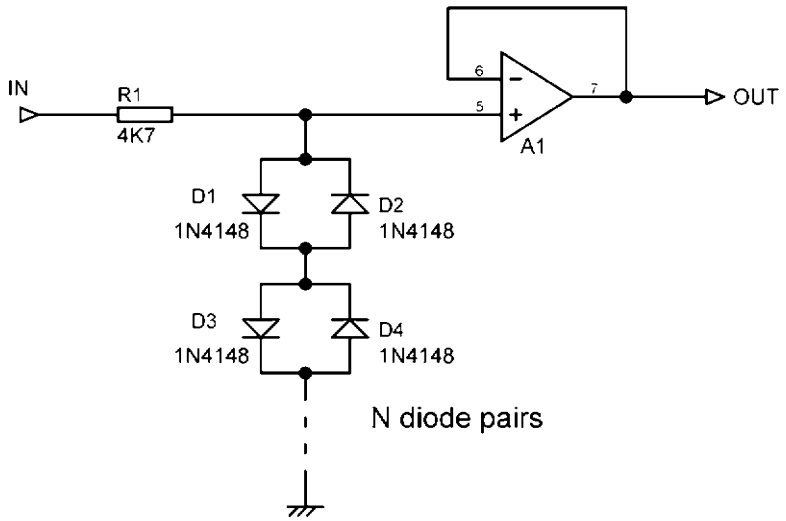


Figure 19.17: A simple passive clipping circuit using diodes

This fails to meet our requirements in several ways. Firstly, the gradual onset of conduction in the diodes means that the clipping is soft, infringing requirements 2 and 4. And finally, it breaks 3 as well, because the clipping level can only be altered by changing the number of diode pairs used, giving 0.6 V steps.

Figure 19.18 shows the force of this, the result of using various numbers of pairs 1N4148 silicon signal diodes. The lower line descending from left to right is the noise floor, getting relatively lower as the input level increases. With one pair of diodes distortion is already clearly above the noise floor for an input as low as 100 mVrms. The gradual increase in distortion after this as level increases shows that the clipping is distinctly soft, and so introduces unacceptable distortion long before it effectively controls the level. An obvious extra snag is that a passive circuit like this has a significant output impedance; in most applications some sort of output buffering like A1 will be required.

The very gradual onset of clipping shown above is due to the slow way in which the exponential conduction law of diodes begins. The clipping action can be made much sharper by connecting each diode to a suitable bias voltage so that it is firmly reverse-biased at low signal levels, and only conducts when the signal is significantly above the bias voltage. The bias chain must have a low impedance to give a steep clipping characteristic, and so diodes are also used to establish the voltages. It would of course be possible to generate near-zero impedance supplies by using op-amps to buffer resistive dividers (which also allow complete flexibility in setting the clipping threshold). However, if op-amps are to be employed, there are better ways to use them, as we shall see later in this chapter.

www.raguvaran.puzl.com

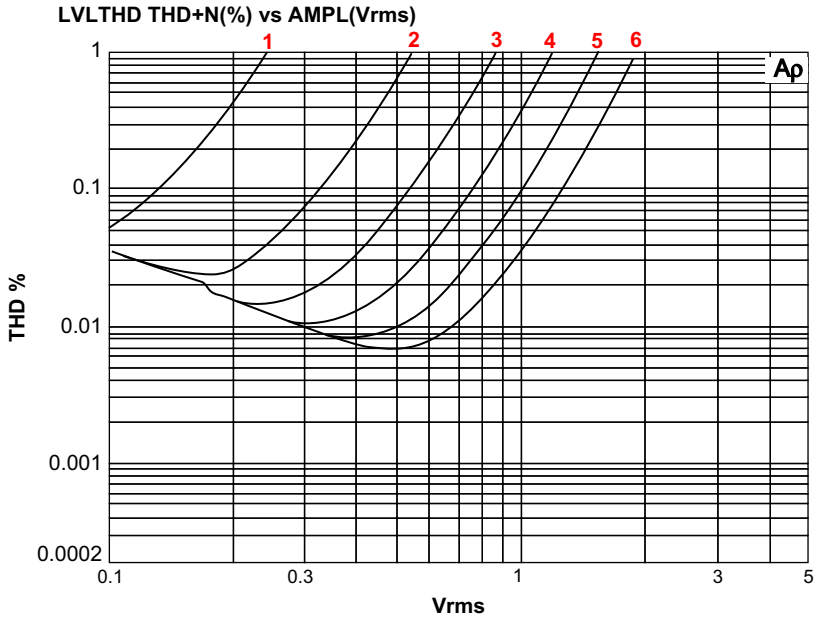


Figure 19.18: Distortion against input level for one to six pairs of clipping diodes. The uneven spacing of the curves as diode pairs are added is due to the logarithmic *X*-axis for input level

A biased clipping circuit is shown in Figure 19.19, and the resulting distortion performance in Figure 19.20, for both one and two pairs of diodes in the biasing chain. The much steeper rise in distortion shows that the onset of clipping is much sharper (compare trace 6 in Figure 19.18 with trace 1 in Figure 19.20).

The clipping action of simple circuits like this is, however, still some way short of perfect, in that the clipped part of the waveform is not a dead flat horizontal line. Diodes do not suddenly become short-circuits when they start to conduct, and the clipped part of the waveform bulges upwards somewhat. Clearly we need active circuitry to sharpen up the diode action.

Active Clipping with Transistors

The next step up in sophistication from simple diodes is the use of transistors for clipping. This is an active process, in that the voltage applied to the base turns on the collector current. It is not simply a matter of using the base-collector or base-emitter diodes to replace the simple diodes mentioned in the previous section.

Figure 19.21(a) shows a basic transistor clipper with NPN and PNP devices shunting the signal path, and Figure 19.22 shows the distortion performance. A buffer is required for a low-impedance output; trace A in Figure 19.22 shows the result without an output buffer and trace

www.raguvaran.puzl.com

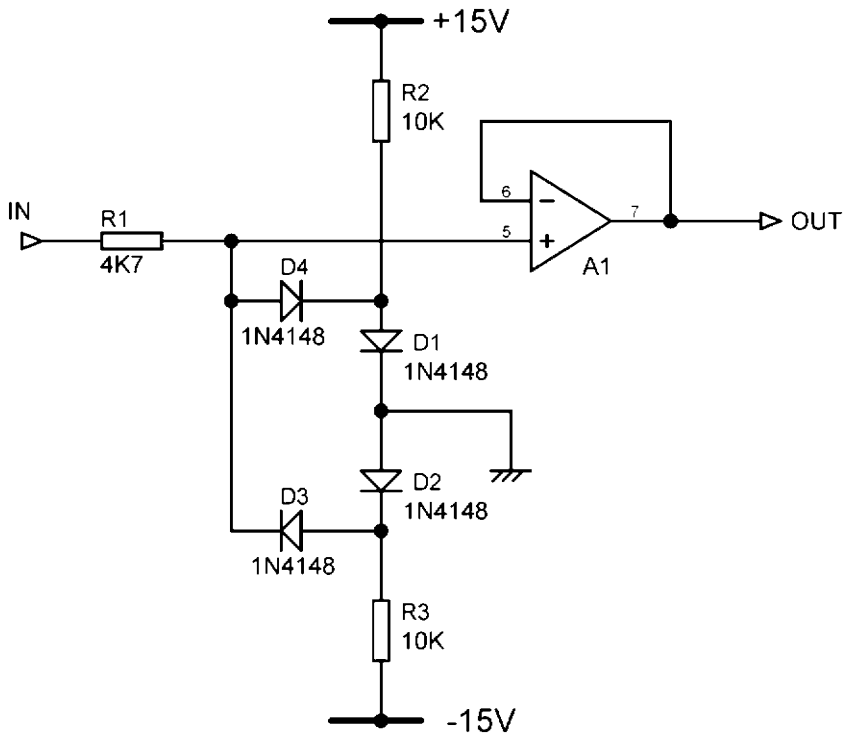


Figure 19.19: A biased diode clipping circuit

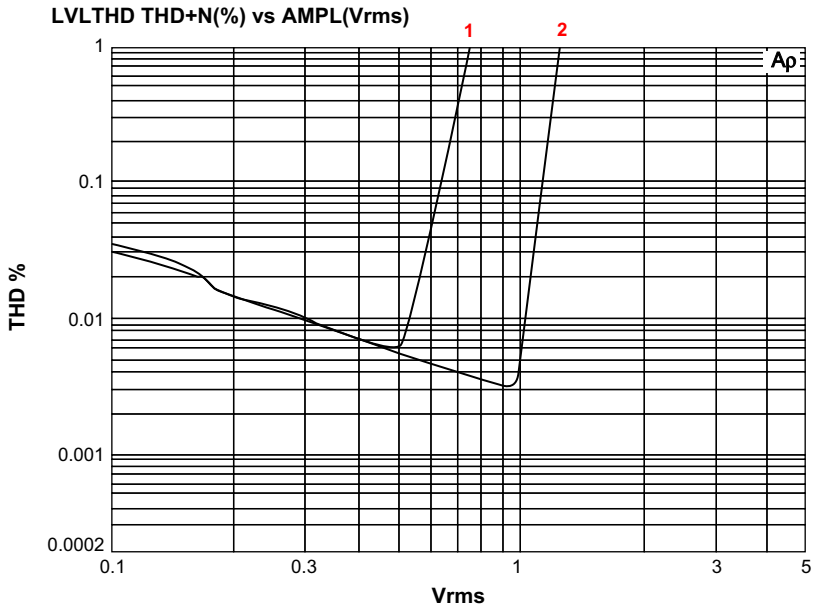


Figure 19.20: Distortion against input level for one or two pairs of diodes in the biasing chain

www.raguvaran.puzl.com

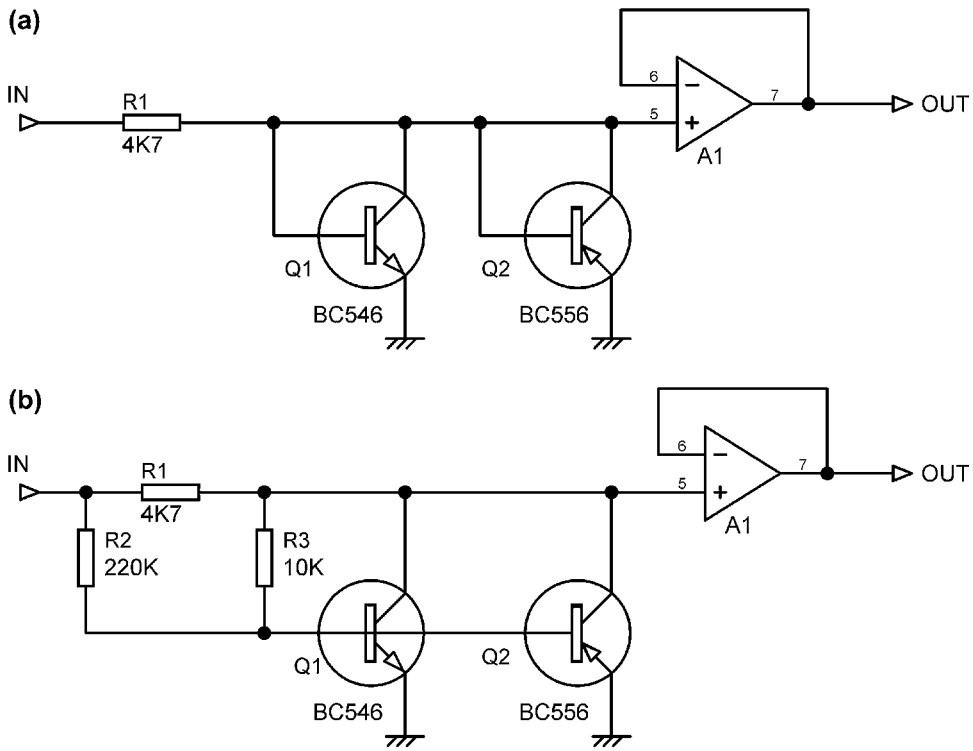


Figure 19.21: (a) A simple transistor clipper. (b) An enhanced transistor clipper with increased base drive via R2

B shows the result with it. In the latter case not only has the noise floor dropped, due to less pickup on what is now a low-impedance output, but the distortion at a given level has risen significantly, as the low-pass action of a capacitative screened cable driven from a medium impedance has been eliminated.

In Figure 19.21(b), the circuit is arranged so that extra drive is applied to the transistor bases via R2, because of the voltage drop across R1. This pulls down the clipped part of waveform in the center, and with a correct choice of values gives improved flatness. The circuit is based on a concept by Stefely. The distortion plot in Figure 19.23 is not very different.

Active Clipping with Op-Amps

Op-amps are noted for their versatility. However, using them to do precision clipping, with low distortion below the clipping point, is more difficult than it at first appears. Here I examine three ways of doing it, as the limitations of the first two approaches are highly instructive.

www.raguvaran.puzl.com

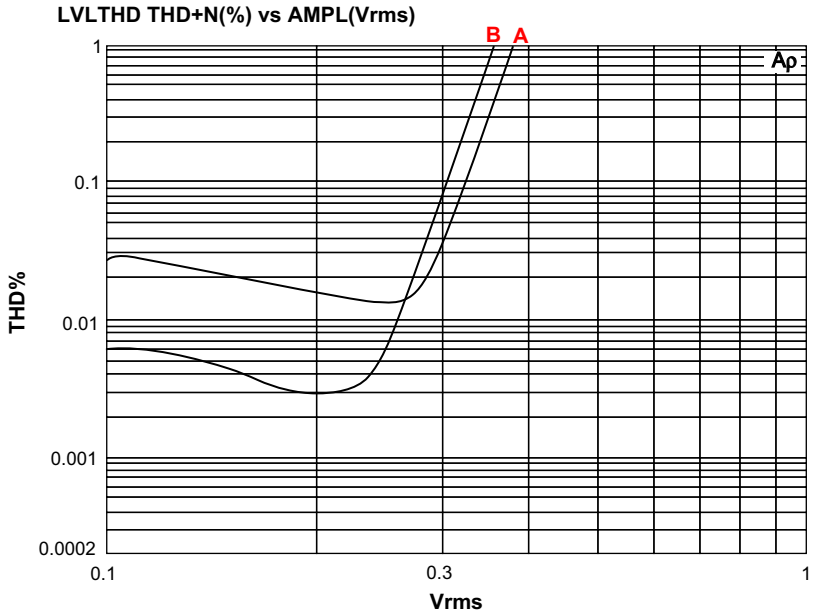


Figure 19.22: Distortion against input level for transistor clipper

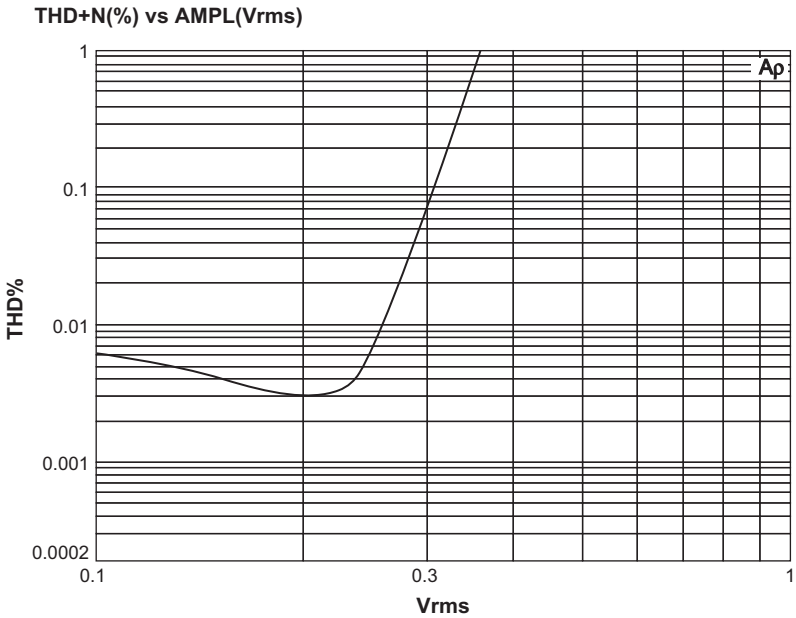


Figure 19.23: Distortion against input level for enhanced transistor clipper

www.raguvaran.puzl.com

Attempt 1: clipping by clamping

The first attempt is the direct descendant of the diode clipping circuits examined above, and uses two active clamps A1, A3 to constrain the voltage at point A, downstream of the 4k7 resistor as shown in Figure 19.24. Normally this point is between the two clipping thresholds, so the inverting input of A1 is positive of Vt−, and the op-amp output is saturated negatively. D1 is therefore firmly reverse-biased and A1 has no effect on the voltage at A. Likewise, normally the inverting input of A2 is negative of Vt+, and D2 is held off.

When the voltage at X tries to exceed Vt+, the output of A2 swings negative and D2 pulls down point A to prevent it. The diode imperfections are servoed out by the open-loop gain of A2, so the clipping threshold is exactly Vt+, neglecting op-amp offsets and other minor errors. In the same way, if point X tries to go below Vt−, the output of A1 swings positive and D1 conducts to clamp the output at this voltage. Like the passive diode clipper above, this circuit has a significant output impedance (4k7 in normal operation) and so buffer A2 is added to give a low output impedance.

The two clipping thresholds can be set to any desired voltage as they are derived from resistive dividers. They can of course be different, though an application where this might be useful is not that easy to visualize.

The clamping-clipper circuit gives very clean clipping, but falls down badly on the distortion it adds to signals below the clipping threshold of 2.2 Vrms. Figure 19.25 shows that distortion reaches an unacceptable 0.035% at 10 kHz and 2 Vrms; the 2 V trace is lower at LF because the relative noise floor is lower. This distortion occurs because point X is at a significant impedance, and is connected to two op-amps with JFET inputs, typically TL072s. These have non-linear capacitances to the IC substrate, and cause distortion that worsens with frequency and level, as shown in Figures 19.25 and 19.26 (see Chapter 4 for more details of this distortion mechanism). This effect could be eliminated by using op-amps with bipolar inputs but then the bias currents have to be dealt with. One advantage of this circuit is that it does not phase-invert, unlike most active clipper circuits.

Figure 19.26 shows that at 10 kHz distortion is beginning to appear when the input exceeds 300 mVrms, and rises slowly until clipping begins just above 2 Vrms.

Attempt 2: negative-feedback clipping

Having found that clipping in the forward path has significant problems, we move to a configuration where this takes place in the negative-feedback loop, so the clipped output can be taken from an op-amp output at low impedance. This also makes interfacing with the next stage simpler.

www.raguvaran.puzl.com

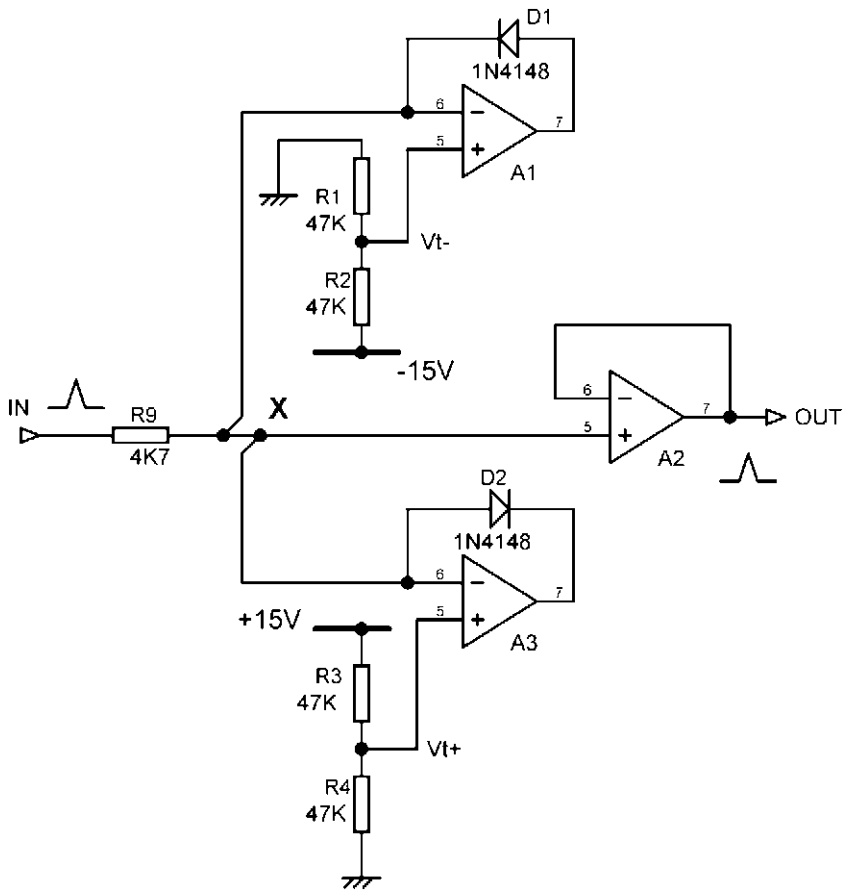


Figure 19.24: An active-clamp clipping circuit, with thresholds set at Vt+ and Vt−. Once again an output buffer A2 is essential to give a low output impedance

The circuit in Figure 19.27 works as follows. With no input, point X sits at +5 V and point Y sits at −5 V. When the output heads positive, eventually Y is pulled positive of 0 V at the A1 inverting input, and D2 starts to conduct, reducing gain and giving clipping. Similarly, on sufficiently large negative output excursions D1 will conduct. The clipping characteristic is rather soft; the gain cannot be reduced entirely to zero above the threshold because of the diode forward impedance in series with the source resistance of the 2k2 – 1k bias network. The latter can be reduced by using lower resistance values but this loads the op-amp output excessively, and draws more power from the supply rails.

This configuration gives less sub-threshold distortion than the previous circuit by a factor of roughly 3 at 10 kHz, but it is still a long way from our requirement for 'absolutely no degradation of the signal', which some of you are by now probably thinking was a bit ambitious.

www.raguvaran.puzl.com

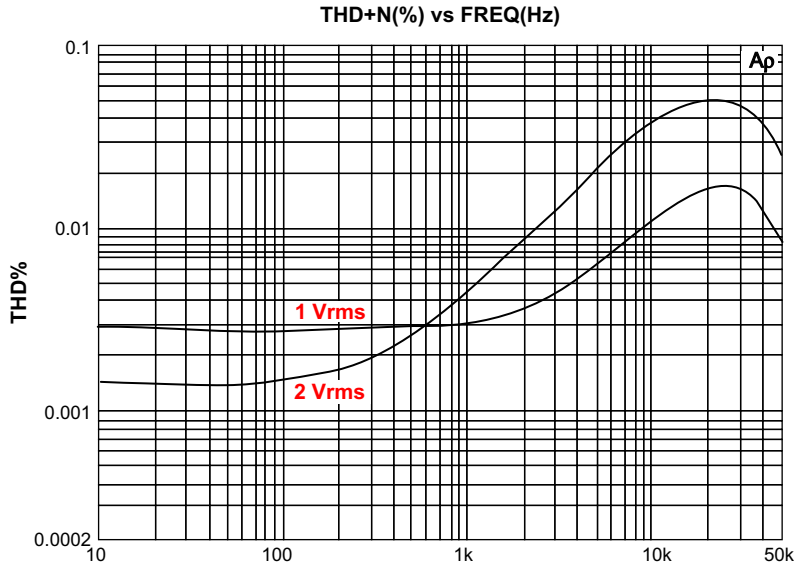


Figure 19.25: Distortion against frequency for the active-clamp clipping circuit

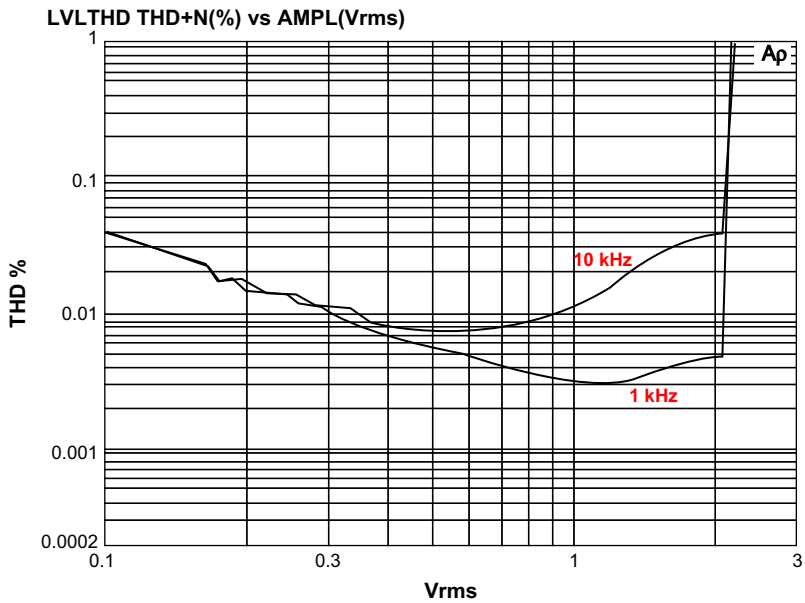


Figure 19.26: Distortion against input level for the active-clamp clipping circuit

The distortion for a 3 Vrms input, which is well below the clipping threshold for this circuit, is shown by curve A in Figure 19.28, and it is instructive to wheel out the scientific method to find out what is going wrong with the linearity. The previous clamping-clipper circuit ran into trouble by driving non-linear op-amp inputs from a non-zero impedance; that cannot be

www.raguvaran.puzl.com

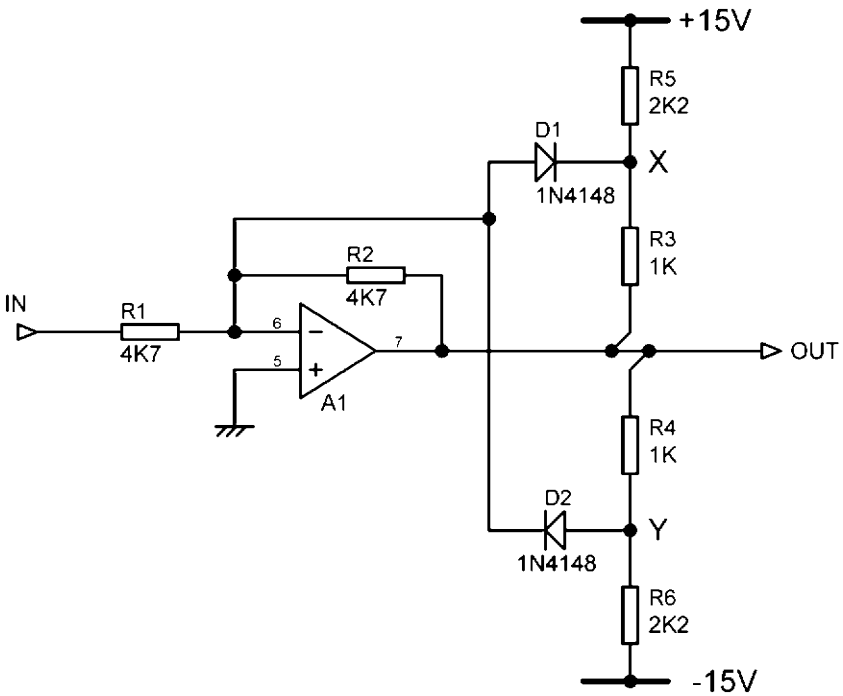


Figure 19.27: A negative-feedback clipper; the clipping threshold is set by the voltages at X and Y

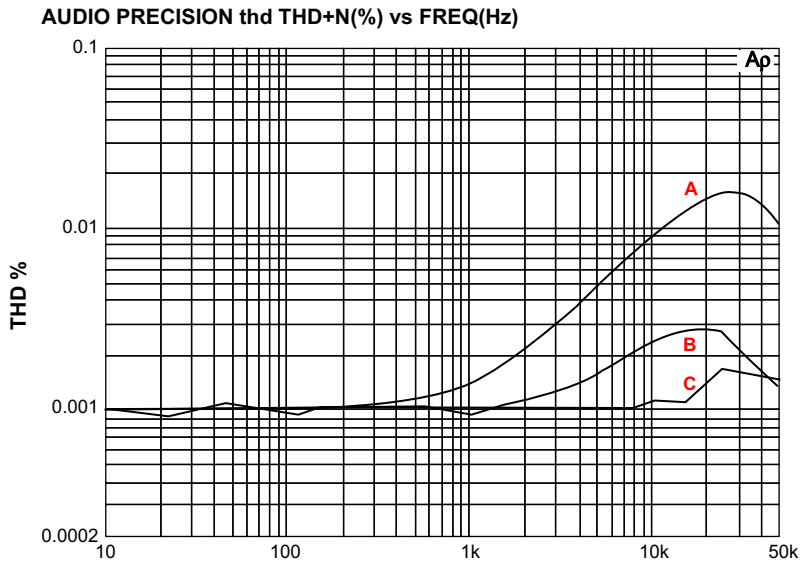


Figure 19.28: The distortion performance of the negative-feedback clipper against frequency for a 3 Vrms input. Plot A: basic distortion of Figure 19.27 Plot B: loading distortion eliminated by buffering A1 output. Plot C: with the diode chain disconnected

www.raguvaran.puzl.com

happening here as both op-amp inputs are at zero voltage, the inverting input so because there is always a healthy amount of shunt feedback.

So what might be the problem? We know that op-amps have their linearity degraded by excessive output loading, and the two bias networks, being connected to rails that are effectively at AC ground, represent a significant load of (2.2k + 1k) in parallel with (2.2k + 1k), which works out to a total load of 1.6k on the output. This is quite enough to degrade the linearity of most op-amps (see Chapter 4), but the idea is still only a hypothesis and needs testing.

This can be done experimentally by means of the circuit in Figure 19.29, which uses a separate voltage-follower A2 to drive the two bias networks, the clean output being taken off before it. Our hypothesis looks good, as eliminating the loading on A1 has much reduced the distortion, giving plot B in Figure 19.28. This, however, still falls somewhat short of perfect linearity.

The distortion remaining in plot B is due to negative feedback through the non-linear diode capacitances while they are still reverse-biased, i.e. below the clipping threshold. This can be demonstrated by disconnecting the output of A2 so clipping is disabled and points X and

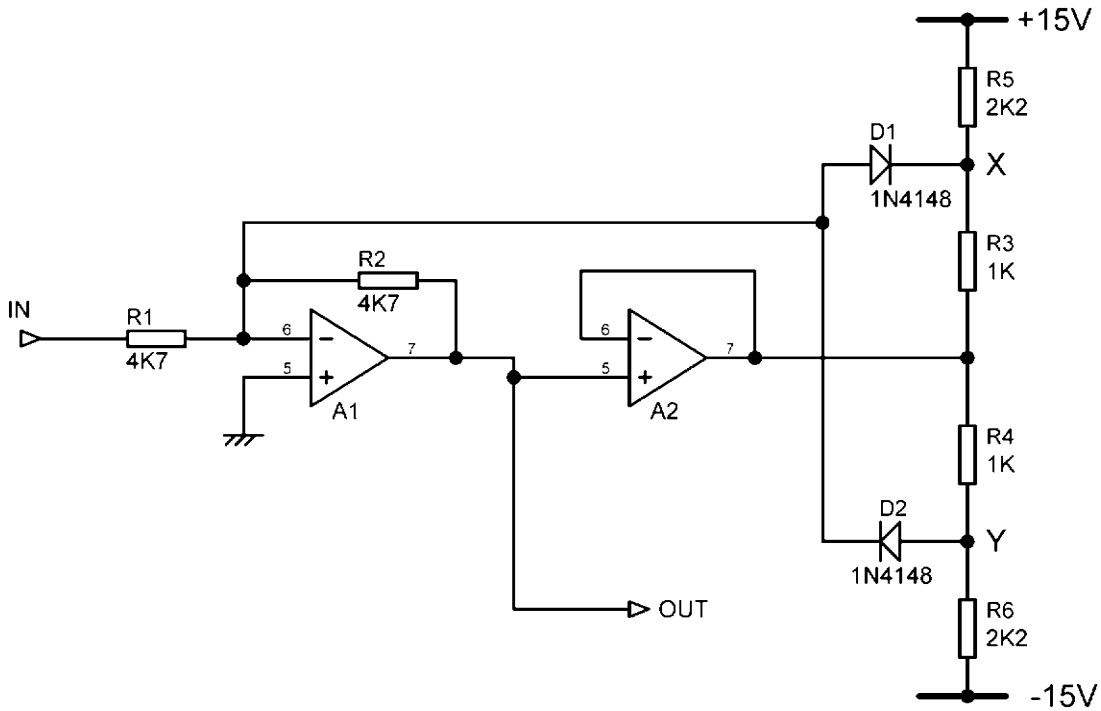


Figure 19.29: A modification of Figure 19.8 that removes the loading distortion from the signal path

www.raguvaran.puzl.com

Y do not move. The diodes are still connected to the summing point at the inverting input of A1, which also does not move, and we now get plot C in Figure 19.28, which is essentially the test-gear distortion plus a little circuit noise.

Putting extra op-amps in a negative-feedback loop is not something to be entered into lightly or inadvisedly, because of the danger that accumulating phase-shifts can make the loop unstable. However, it often works if the extra op-amp is configured as a voltage-follower because the 100% local feedback in the follower makes its bandwidth as great as possible, and significantly greater than that of an inverting stage, which works at a noise gain of 2. Here it works dependably.

While this circuit is some way short of perfection, it can be useful sometimes, especially if there happens to be a spare op-amp half left over to act as the buffer A2 in Figure 19.29. Don't forget that this circuit gives a phase inversion.

Attempt 3: feedforward clipping

The Attempt 2 clipper demonstrated one way to reduce sub-threshold distortion, but it also highlighted the difficulty of getting hard clipping and tight level control by using negative-feedback techniques. While feedback is a most powerful technique, it is not the only way – sometimes feedforward does it better.

The circuit of a feedforward clipper is shown in Figure 19.30. Below the clipping level it acts simply as a unity-gain inverting stage, with the forward path through R1. This configuration eliminates common-mode distortion in A2 during normal operation.

The clipping circuit consists of two shunt-feedback precision rectifiers A1, A3 that are biased by currents injected into their summing points via R2, R3 so they conduct only above the desired clipping threshold. A1 handles negative clipping; R2 injects 15 V/47k = 319 μA into the summing node of A1, and as long as this exceeds the current being pulled out of this node via R5 during negative inputs it will be counteracted by feedback through D2. When the current through R5 exceeds that injected by R2, the op-amp output must go positive, so it can maintain its inverting input at ground by feedback through D1 and R6. As before, the gain of the op-amp A1 is used to 'servo out' the non-linearity of the diodes D1, D2.

When D1 conducts, a clamped version of the input signal appears at the junction of R6 and D1, and a unity-gain but phase-inverted path is established through R4; the overall gain suddenly drops to zero as the two signals are canceled at the summing point of A3, so the output cannot move any further and is clipped. The precision rectifier A3 acts in the same way for positive inputs.

Since A1 and A3 are never active at the same time, they can share the resistor R4, which makes for a very elegant circuit. The downside of this economy is that the active precision rectifier has to also drive the feedback resistance (R6 or R8) of the inactive precision rectifier, its other

www.raguvaran.puzl.com

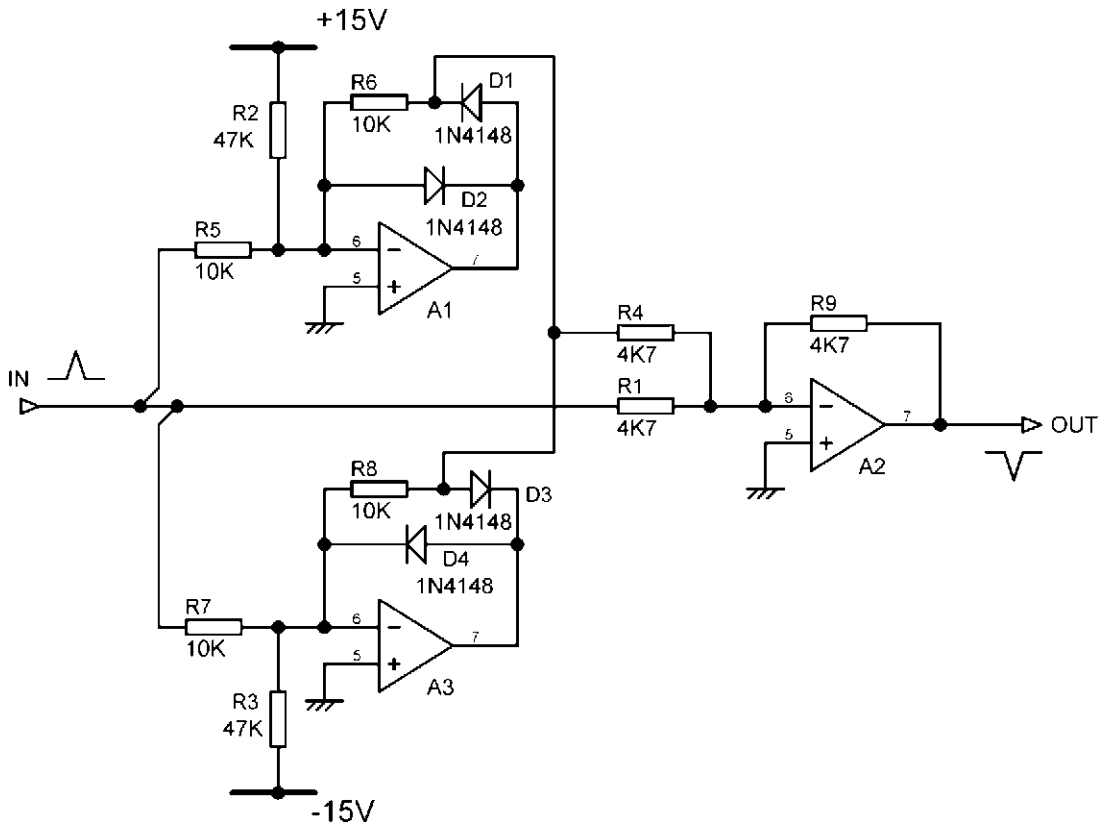


Figure 19.30: A feedforward clipping circuit. The clipping threshold is 2.25 Vrms with the values shown

end being connected to virtual ground, and this will use up slightly more power. From this point of view the saving of a resistor is not quite as elegant as it looks; if power economy is paramount then separate cancellation resistors to each precision rectifier remove the problem.

As Figure 19.31 demonstrates, this circuit is a success. There is now no measurable distortion, even just below the clipping threshold. (The step around 20 kHz is an artefact of the test-gear used.) The reason for this is that there are no op-amp loading issues and no reverse-biased diodes with signal on one side connected to the audio path on the other side. The two clamp stages have zero output until the threshold is reached, with R4 being connected only to a virtual ground.

With perfectly accurate resistors, and indeed perfect op-amps, this circuit would give complete cancellation and the top of the clipped waveform would be absolutely flat. Cancellation processes have something of a dubious reputation in engineering circles, because of the need for accurate matching of amplitude and phase to get a good cancellation. Where this depends

www.raguvaran.puzl.com

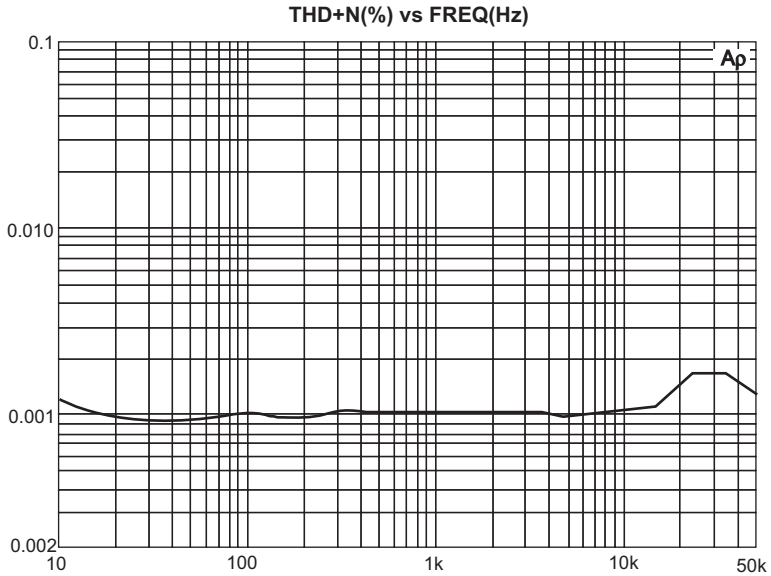


Figure 19.31: Distortion against frequency just below the clipping threshold. Input level 2 Vrms. This is once again the distortion of the test-gear plus a little added noise

on several factors – and particularly if semiconductor characteristics enter the equation – it can indeed be difficult to control. However, in this particular case the cancellation accuracy depends only on a few well-matched resistors. One percent tolerance is quite good enough, and most modern resistors are this accurate. The only consequence of mismatching is a very slight curvature of the clipped part of the waveform, and this is not normally enough to be troublesome.

As shown by the phase spikes in the diagram, this circuit gives a phase inversion, and this must be taken into account during the system design.

Noise Generators

As with the introduction of deliberate clipping, it may seem perverse to go to a lot of trouble to generate noise when we have spent great efforts so far to minimize it. However, noise sources have their uses, for example in loudspeaker testing, in room equalization, and in analog synthesizers. Noise comes, as is well known, in various colors. This is not due to synaesthesia ('Tuesdays are red!'), but is a convenient way of describing the spectral content of various particularly useful kinds of noise.

White noise has equal power in equal bandwidth, so there is the same power between 100 and 200 Hz as there is between 1100 and 1200 Hz. It is the type of noise produced by most electronic noise mechanisms.

www.raguvaran.puzl.com

Pink noise has equal power in equal ratios of bandwidth, so there is the same power between 100 and 200 Hz as there is between 200 and 400 Hz. The energy per Hz falls at 3 dB/octave as frequency increases. Pink noise is very important as it gives a flat response when viewed on a third-octave or other constant percentage bandwidth spectrum analyzer.

Red noise has its energy per Hz falling at 6 dB/octave; its main use is in synthesizers. There is more on noise of different colors in Chapter 1.

Most noise generators these days are digital, using maximal-length sequences to produce noise that is tightly defined in amplitude and spectral content. That, however, is rather outside the province of this analog book, so instead we will take a quick look at analog noise generation.

The most popular method is to use the white noise produced when a bipolar transistor is reverse-biased and allowed to work as a Zener diode, as shown in Figure 19.32. This does the transistor no harm, so long as the current through it is limited, and is surprisingly consistent, when you consider that this is not exactly the official way to use a transistor. I took a collection of 10 BC184C transistors of varying pedigree and provenance, and found that with an emitter current of 36 µA the noise output varied between −61.5 and −65.6 dBu, with one outlier at −56.9 dBu (bandwidth 22 Hz – 22 kHz). The Zener voltage was between 7.7 and 8.9 V. When I tried a real Zener diode, of 6V2 voltage, the noise output was much less, around −85 dBu. This is presumably because Zeners are designed to produce minimal noise, though they still generate much more than simple silicon diodes; noise is distinctly unwanted in voltage reference applications.

The transistor noise output is well below a millivolt, and a good deal of amplification is needed to get it up to a useful level. Figure 19.32 shows a humble TL072 performing this service; there is of course absolutely no point in using a low-noise op-amp such as the 5532, and the low bias currents of the TL072 mean that high-value resistors can be used without DC offset troubles.

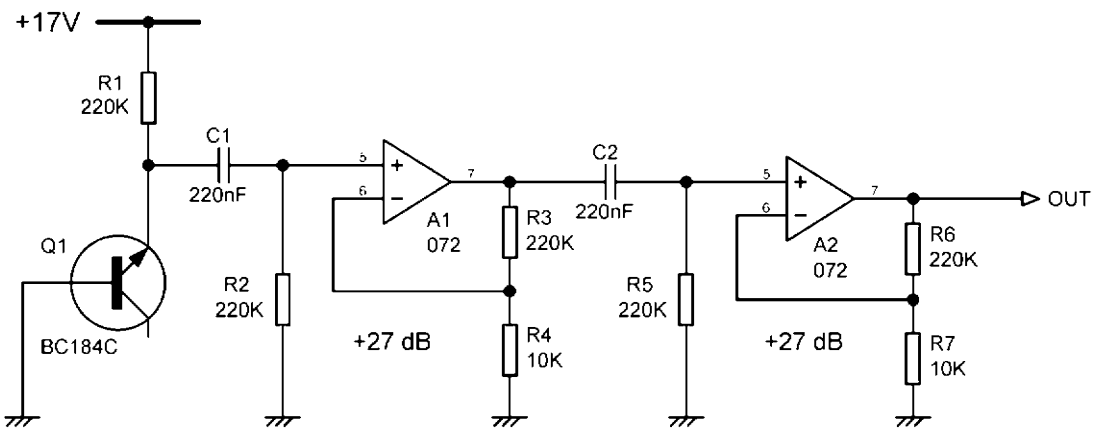


Figure 19.32: A white noise generator using a bipolar transistor as a Zener diode

www.raguvaran.puzl.com

Each stage has a gain of 27 dB. A point to remember is that the TL072 has limited open-loop bandwidth compared with more modern op-amps, and if you try to take too much gain in one stage this will lead to a high-frequency roll-off. The last amplifier stage must not be allowed to clip as this will modify the energy spectrum. The output here is about 400 mVrms so clipping will statistically be very rare. Remember there was a 4 dB variation in noise output for different transistor samples, and a preset gain trim may be needed for some applications.

Pinkening Filters

Converting white noise into the much more useful pink noise is proverbially difficult because you can't make a filter with a true −3 dB/octave slope, as filter slopes come in multiples of 6 dB/octave only. The only solution when you need a pinkening filter is a series of overlapping low-pass and high-pass time-constants (i.e. alternate poles and zeroes) that approximate to the required slope. The more pairs of poles and zeroes used, the more accurately the response approximates to a −3 dB/octave slope. Two possible versions are shown in Figure 19.33: that in (a) uses three pole-zero pairs, with a final unmatched pole introduced by C4; that in (b) has four pole-zero pairs, with a final unmatched pole introduced by C8. The closer the pole-zero spacing, the less the wobble on the frequency response.

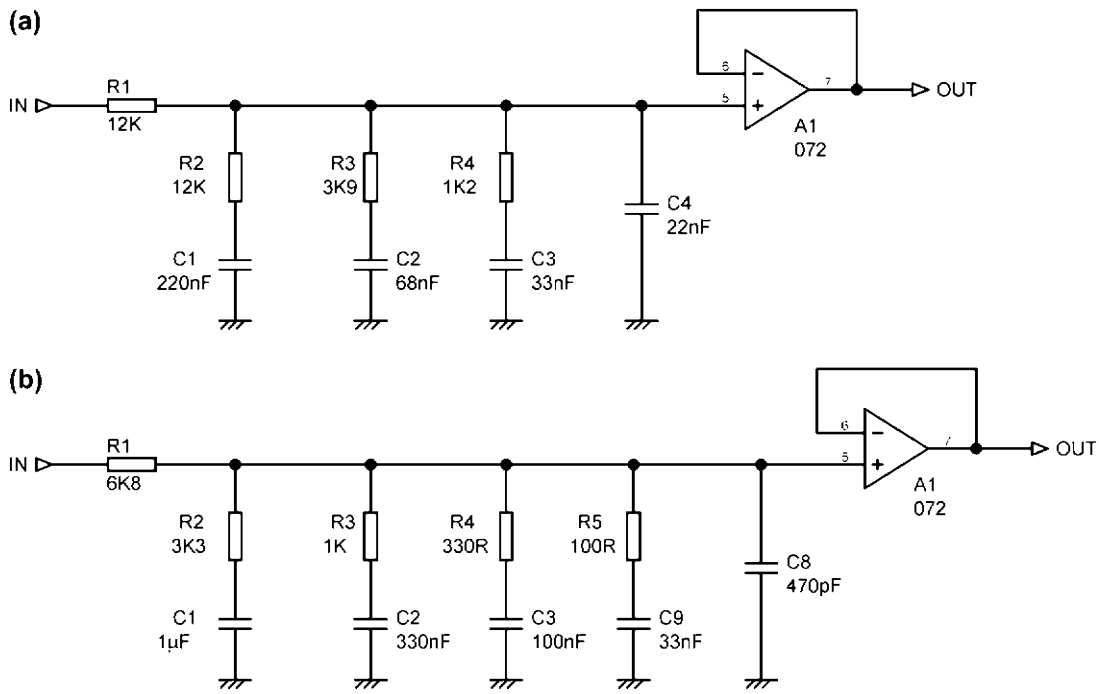


Figure 19.33: Two −3 dB/octave pinkening filters for turning white noise into pink noise

www.raguvaran.puzl.com

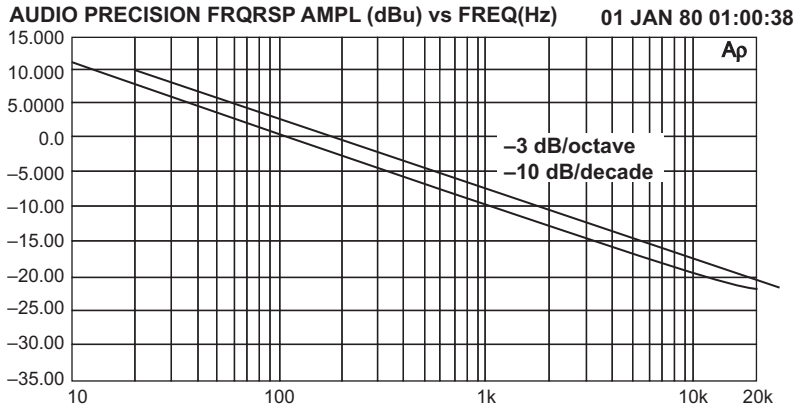


Figure 19.34: The frequency response of the pinkening filter in Figure 19.33(b)

The response of (b), the more expensive and more accurate version, is shown in Figure 19.34. The measured curve (the lower one) is seen to be accurately parallel to the −3 dB/octave line drawn above it on the graph.

Red noise is rarely required, but if needed it can be made from white noise by passing it through a simple integrator, which has the necessary −6 dB/octave slope.

References

[1] D. Self, A high-quality compressor/limiter, Wireless World (December 1975) p. 587.
[2] B. Duncan, VCAs investigated, Parts 1–3, Studio Sound (June/August 1989).
[3] P. Buff, (Allison Research), The new VCA technology, db (August 1980) p. 32.
[4] R. Bransbury, Automation and the VCA, Studio Sound (August 1985) p. 80.
[5] D. Frey, The ultimate VCA, AES preprint.
[6] M. Hawksford, Topological enhancements of translinear two-quadrant gain cells, JAES (June 1989) p. 465.
[7] Shorter, et al., The dynamic characteristics of limiters for sound programme circuits, BBC Engineering Monograph No. 70 (October 1967).
[8] D. Self, Audio Power Amplifier Design Handbook, fifth ed, Focal Press, 2009, p. 270.
[9] D. Self, Audio Power Amplifier Design Handbook, fifth ed, Focal Press, 2009, p. 39.

www.raguvaran.puzl.com

CHAPTER 20

Power Supplies

'We thought, because we had power, we had wisdom.'

Stephen Vincent Benet, *Litany for Dictatorships*, 1935

Op-Amp Supply-Rail Voltages

It has been mentioned several times in the earlier chapters of this book that running op-amps at the slightly higher voltage of ±17 V rather than ±15 V gives an increase in headroom and dynamic range of 1.1 dB for virtually no cost and with no reliability penalty. Soundcraft ran all the op-amps in their mixing consoles at ±17 V for at least two decades, and op-amp failures were almost unknown. This recommendation assumes that the op-amps concerned have a maximum supply voltage rating of ±18 V, which is the case for the Texas TL072, the new LM4562, and many other types.

The 5532 is (as usual) in a class of its own. Both the Texas and Fairchild versions of the NE5532 have an absolute maximum power-supply voltage rating of ±22 V (though Texas also gives a 'recommended supply voltage' of ±15 V), but I have never met any attempt to make use of this capability. The 5532 runs pretty warm on ±17 V when it is simply quiescent, and my view (and that of almost all the designers I have spoken to) is that running it at any higher voltage is simply asking for trouble. This is a particular concern in the design of mixing consoles, which may contain thousands of op-amps – anything that impairs their reliability is going to cause a *lot* of trouble. In any case, moving from ±17 V rails to ±18 V rails only gives 0.5 dB more headroom. Stretching things to ±20 V would give 1.4 dB more than ±17 V, and running on the ragged edge at ±22 V would yield a more significant 2.2 dB more than ±17 V, but you really wouldn't want to do it. Pushing the envelope like this is also going to cause difficulties if you want to run op-amps with maximum supply ratings of ±18 V from the same power supply.

We will therefore concentrate here on ±17 V supplies for op-amps, dealing first with what might be called 'small power supplies', i.e. those that can be conveniently built with TO-220 regulators. This usually means an output current capability that does not exceed 1.5 A, which is plenty for even complicated preamplifiers, electronic crossovers, etc., but will only run a rather small mixing console; the needs of large consoles are dealt with later in this chapter.

Small Signal Audio Design; ISBN: 9780240521770
Copyright © 2010 Elsevier Ltd; All rights of reproduction, in any form, reserved.

www.raguvaran.puzl.com

TABLE 20.1 PSRR specs for common op-amps

Op-amp type	PSRR minimum (dB)	PSRR typical (dB)
5532	80	100
LM4562	110	120
TL072	70	100

An important question is: how low does the noise and ripple on the supply output rails need to be? Op-amps in general have very good power-supply rejection ratios (PSRRs) and some manufacturers' specs are given in Table 20.1.

The PSRR performance is actually rather more complex than the bare figures given in the table imply; PSRR is typically frequency dependent (deteriorating as frequency rises) and different for the +V and −V supply pins. It is, however, rarely necessary to get involved in this degree of detail. Fortunately even the cheapest IC regulators (such as the venerable 78xx/79xx series) have low enough noise and ripple outputs that op-amp PSRR performance is rarely an issue.

There is, however, another point to ponder; if you have a number of electrolytic-sized decoupling capacitors between rail and ground, enough noise and ripple can be coupled into the non-zero ground resistance to degrade the noise floor. Intelligent placing of the decouplers can help – putting them near where the ground and supply rails come on to the PCB means that ripple will go straight back to the power supply without flowing through the ground tracks on the rest of the PCB. This is of limited effectiveness if you have a number of PCBs connected to the same IDC cable, as in many small mixing desks, and in such cases low-ripple power supplies may be essential.

Apart from the op-amp supply rails, audio electronics may require additional supplies, as shown in Table 20.2.

It is often convenient to power relays from a +9 V unregulated supply that also feeds the +5 V microcontroller regulator (see later in this chapter). The use of +24 V to power LED metering systems is dealt with in Chapter 18 on metering, and +48 V phantom supplies are examined at the end of this chapter.

TABLE 20.2 Typical additional supply rails for op-amp-based systems

Supply voltage (V)	Function
+5	Housekeeping microcontroller
+9	Relays
+24	LED bar-graph metering systems, discrete audio circuitry, relays
+48	Microphone phantom power

www.raguvaran.puzl.com

Designing a ±15 V Supply

Making a straightforward ±15 V, 1 A supply for an op-amp-based system is very simple, and has been ever since the LM7815/7915 IC regulators were introduced (which was a long time ago). The parts are robust and inexpensive with both over-current and over-temperature protection, and give low enough output noise for most purposes. We will look quickly at the basic circuit because it brings out a few design points that apply equally to more complex variations on the theme. Figure 20.1 shows the schematic, with typical component values; a center-tapped transformer, a bridge rectifier, and two reservoir capacitors C1, C2 provide the unregulated rails that feed the IC regulators. The secondary fuses must be of the slow-blow type. The small capacitors C7–C9 across the input to the bridge reduce RF emissions from the rectifier diodes; they are shown as X-cap types not because they have to withstand 230 Vrms, but to underline the need for them to be rated to withstand continuous AC stress. The capacitors C3, C4 are to ensure HF stability of the regulators, which like a low AC impedance at their input pins, but these are only required if the reservoir capacitors are not adjacent to the regulators, i.e. more than 10 cm away. C5, C6 are not required for regulator stability with the 78/79 series – they are there simply to reduce the supply output impedance at high audio frequencies.

There are really only two electrical design decisions to be made: the AC voltage of the transformer secondary and the size of the reservoir capacitors. As to the first, you must make sure that the unregulated supply is high enough to prevent the rails dropping out (i.e. letting hum through) when a low mains voltage is encountered, but not so high that either the maximum input voltage of the regulator is exceeded, or it suffers excessive heat dissipation. How low a mains voltage it is prudent to cater for depends somewhat on where you think your equipment is going to be used, as some parts of the world are more subject to brown-outs than others. You must consider both the minimum voltage drop across the regulators (typically 2 V)

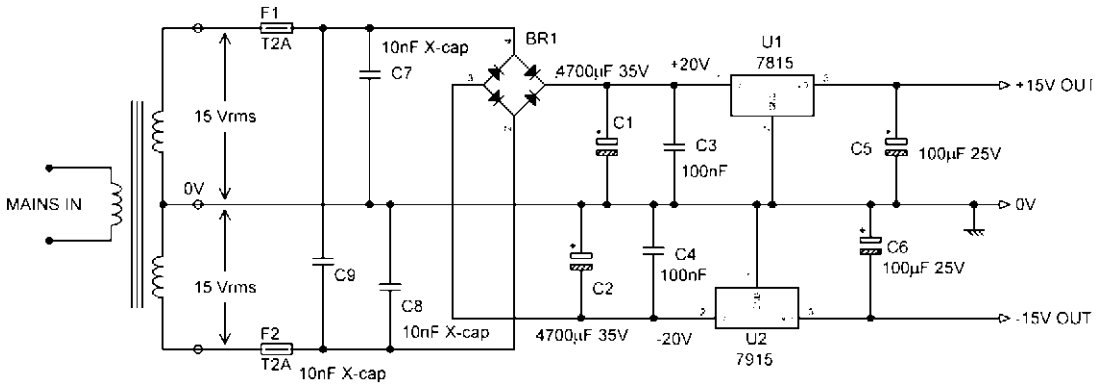


Figure 20.1: A straightforward ±15 V power supply using IC regulators

www.raguvaran.puzl.com

and the ripple amplitude on the reservoirs, as it is in the ripple troughs that the regulator will first 'drop out' and let through unpleasantness at 100 Hz.

In general, the rms value of the transformer secondary will be roughly equal to the DC output voltage.

The size of reservoir capacitor required depends on the amount of current that will be drawn from the supply. The peak-to-peak ripple amplitude is normally in the region of 1–2 volts; more ripple than this reduces efficiency as the unregulated voltage has to be increased to allow for unduly low ripple troughs, and less ripple is usually unnecessary and gives excessive reservoir capacitor size and cost. The amount of ripple can be estimated with adequate accuracy by using:

$$V_{\text{pk-pk}} = \frac{I \cdot \Delta t \cdot 1000}{C} \qquad \text{(Equation 20.1)}$$

where $V_{\text{pk-pk}}$ is the peak-to-peak ripple voltage on the reservoir capacitor, I is the maximum current drawn from that supply rail in amps, Δt is the length of the capacitor discharge time, taken as 7 milliseconds, C is the size of the reservoir capacitor in microfarads, and the '1000' factor simply gets the decimal point in the right place.

Note that the discharge time is strictly a rough estimate, and assumes that the reservoir is being charged via the bridge for 3 ms, and then discharged by the load for 7 ms. Rough estimate it may be, but I have always found it works very well.

The regulators must be given adequate heat-sinking. The maximum voltage drop across each regulator (assuming 10% high mains) is multiplied by the maximum output current to get the regulator dissipation in watts, and a heat-sink selected with a suitable thermal resistance to ambient (in °C per watt) to ensure that the regulator package temperature does not exceed, say, 90°C. Remember to include the temperature drop across the thermal washer between regulator and heat-sink.

Under some circumstances it is wise to add protective diodes to the regulator circuitry, as shown in Figure 20.2. The diodes D1, D3 across the regulators are reverse-biased in normal operation, but if the power supply is driving a load with a large amount of capacitance, it is possible for the output to remain higher in voltage than the regulator input as the reservoir voltage decays. D1, D3 prevent this effect from putting a reverse voltage across the regulators. Such diodes are not usually required with normal op-amp circuitry, as the amount of rail decoupling, shown as C7, C8 in Figure 20.2, is usually modest.

The shunt protection diodes D2, D4 are once again reverse-biased in normal operation. D2 prevents the +15 V supply rail from being dragged below 0 V if the −15 V rail starts up slightly faster, and likewise D4 protects the −15 V regulator from having its output pulled above 0 V. This can be an important issue if rail-to-rail decoupling such as C9 is in use; such

www.raguvaran.puzl.com

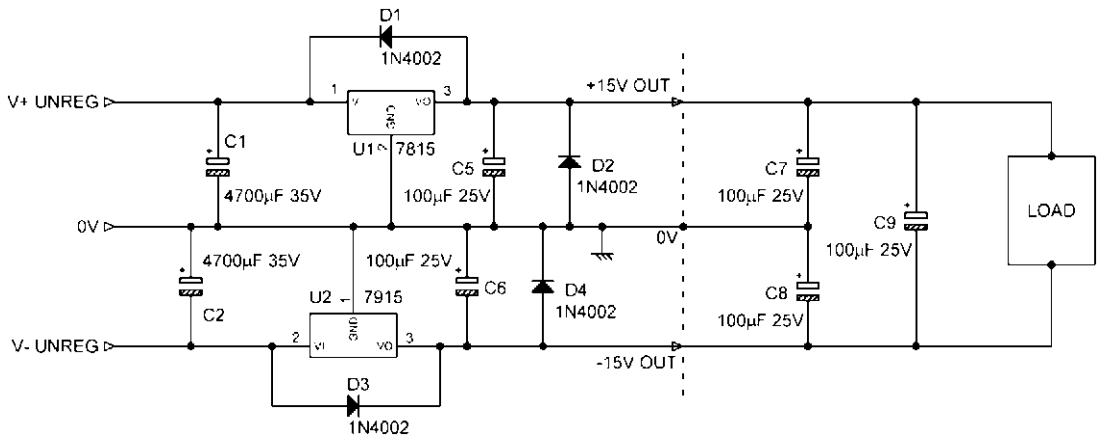


Figure 20.2: Adding protection diodes to a ±15V power supply. The load has decoupling capacitors to both ground (C7, C8) and between the rails (C9); the latter can cause start-up problems

decoupling can be useful because it establishes a low AC impedance across the supply rails without coupling supply-rail noise into the ground, as C7, C8 are prone to do. However, it also makes a low-impedance connection between the two regulators. D2, D4 will prevent damage in this case, but leave the power supply vulnerable to start-up problems; if its output is being pulled down by the −15 V regulator, the +15 V regulator may refuse to start. This is actually a very dangerous situation, because it is quite easy to come up with a circuit where start-up will only fail one time in 20 or more, the incidence being apparently completely random, but presumably controlled by the exact point in the AC mains cycle where the supply is switched on, and other variables such as temperature, the residual charge left on the reservoir capacitors, and the phase of the moon. If even one start-up failure event is overlooked or dismissed as unimportant. then there is likely to be serious grief further down the line. *Every power supply start-up failure must be taken seriously.*

Designing a ±17 V Supply

There are 15 V IC regulators (7815, 7915) and there are 18 V IC regulators (7818, 7918) but there are no 17 V IC regulators. This problem can be effectively solved by using 15 V regulators and adding 2 volts to their output by manipulating the voltage at the Ref pin. The simplest way to do this is with a pair of resistors that divide down the regulated output voltage and apply it to the REF pin, as shown in Figure 20.3(a). (The transformer and AC input components have been omitted in this and the following diagrams, except where they differ from those shown above.) Since the regulator maintains 15 V between the Out and Ref pin, with suitable resistor values the actual output with respect to 0 V is 17 V.

www.raguvaran.puzl.com

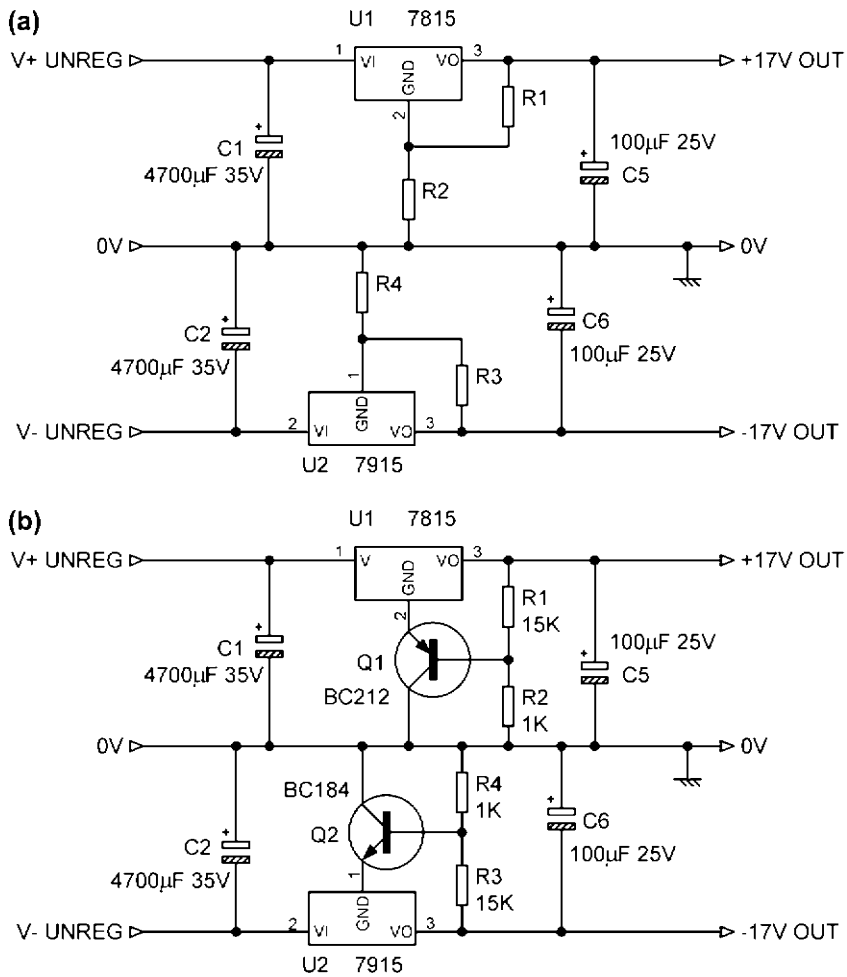


Figure 20.3: Making a ±17 V power supply with 15 V IC regulators. (a) Using resistors is inefficient and/or inaccurate. (b) Adding transistors to the voltage-determining resistor network makes the output voltage more predictable and reduces the power consumed in the resistors

The snag with this arrangement is that the quiescent current that flows out of the Ref pin to ground is not well controlled; it can vary between 5 and 8 mA, depending on both the input voltage and the device temperature. This means that R1 and R2 have to be fairly low in value so that this variable current does not cause excessive variation of the output voltage, and therefore power is wasted.

If a transistor is added to the circuit as in Figure 20.3(b), then the impedance seen by the Ref pin is much lower. This means that the values of R1 and R2 can be increased by an order of magnitude, reducing the waste of regulator output current and reducing the heat liberated. This sort of maneuver is also very useful if you find that you have a hundred thousand

www.raguvaran.puzl.com

15 V regulators in store, but what you actually need for the next project is an 18 V regulator, of which you have none.

What about the output ripple with this approach? I have just measured a power supply using the exact circuit of Figure 20.3(b), with 2200 μF reservoirs, and I found −79 dBu (87 μVrms) on the +17 V output rail, and −74 dBu (155 μVrms) on the 17 V rail, which is satisfyingly low for inexpensive regulators, and should be adequate for almost all purposes; note that these figures include regulator noise as well as ripple. The load current was 110 mA. If you *are* plagued by ripple troubles, the usual reason is a rail-decoupling capacitor that is belying its name by coupling rail ripple into a sensitive part of the ground system, and the cure is to correct the grounding rather than design an expensive ultra-low-ripple PSU. Note that doubling the reservoir capacitance to 4400 μF only improved the figures to −80 and −76 dBu respectively; just increasing reservoir size is not a cost-effective way to reduce the output ripple.

Using Variable-Voltage Regulators

It is of course also possible to make a ±17 V supply by using variable-output-voltage IC regulators such as the LM317/337. These maintain a small voltage (usually 1.2 V) between the Output and ADJ pins, and are used with a resistor divider to set the output voltage. The quiescent current flowing out of the ADJ pin is a couple of orders of magnitude lower than for the 78/79 series, at around 55 μA, and so a simple resistor divider gives adequate accuracy of the output voltage, and transistors are no longer needed to absorb the quiescent current. A disadvantage is that this more sophisticated kind of regulator is somewhat more expensive than the 78/79 series; at the time of writing they cost something like 50% more. The 78/79 series with transistor voltage setting remains the most cost-effective way to make a non-standard-voltage power supply at the time of writing.

It is clear from Figure 20.4 that the 1.2 V reference voltage between ADJ (shown in figures as GND) and Out is amplified by many times in the process of making a 17 or 18 V supply; this not only increases output ripple, but also output noise as the noise from the internal reference is being amplified. The noise and ripple can be considerably reduced by putting a capacitor C7 between the ADJ pin and ground. This makes a dramatic difference; in a test PSU with a 650 mA load the output noise and ripple was reduced from −63 dBu (worse than 78xx series) to −86 dBu (better than 78xx series) and so such a capacitor is usually fitted as standard. If it is fitted, it is then essential to add a protective diode D1 to discharge C7, C8 safely if the output is short-circuited, as shown in Figure 20.5.

The ripple performance of the aforementioned test PSU, with a 6800 μF reservoir capacitor and a 650 mA load, is summarized for both types of regulator in Table 20.3. Note that the exact ripple figures are subject to some variation between regulator specimens.

www.raguvaran.puzl.com

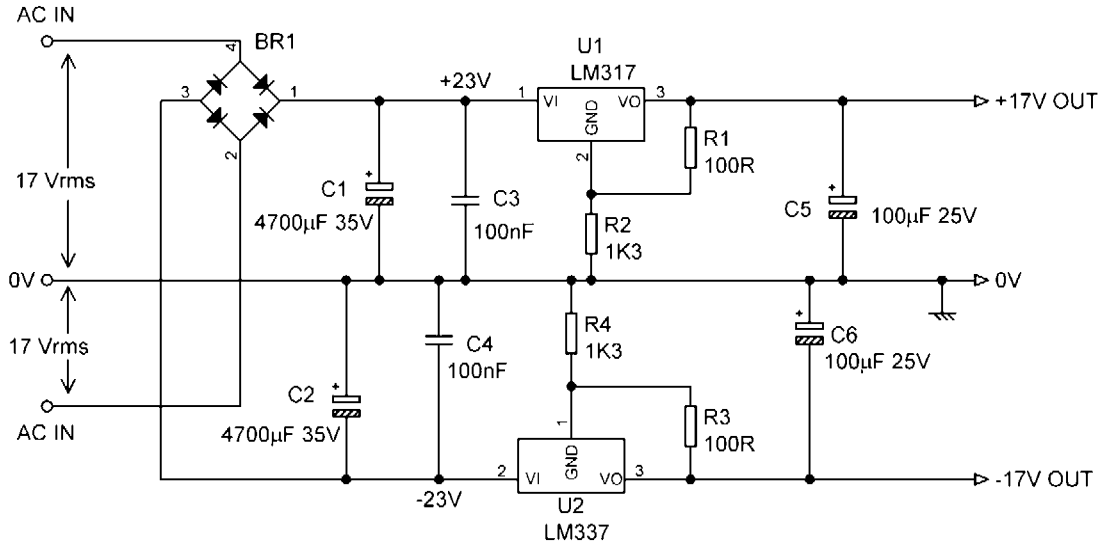


Figure 20.4: Making a ±17 V power supply with variable-voltage IC regulators

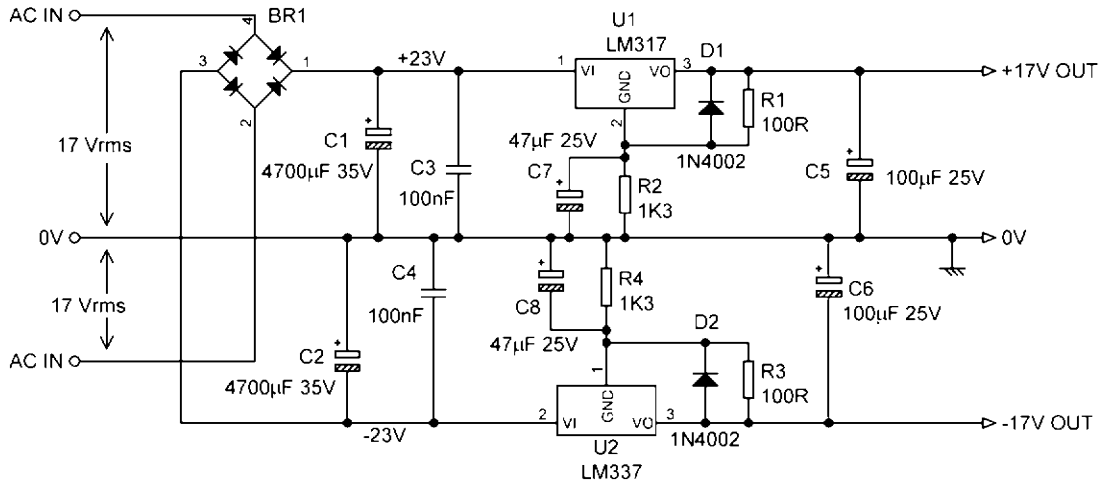


Figure 20.5: Ripple improvement and protective diodes for a variable-voltage IC regulator

Improving Ripple Performance

Table 20.3 shows that the best noise and ripple performance that can be expected from a simple LM317 regulator circuit is about −86 dBu (39 µVrms) and this still contains a substantial ripple component. The reservoir capacitors are already quite large at 4700 µF, so what is to be done if lower ripple levels are needed? The options are as follows:

1. Look for a higher-performance IC regulator. They will cost more and there are likely to be issues with single sourcing.

www.raguvaran.puzl.com

TABLE 20.3 Comparing the noise and ripple output of various regulator options

	7815 + transistor (dBu)	LM317 (dBu)
No C on LM317 ADJ pin	−73 (all ripple)	−63 (ripple and noise)
47 μF on LM317 ADJ pin	−73 (all ripple)	−86 (ripple and noise)
Input filter 2.2Ω and 2200 μF	−78 (ripple and noise)	−89 (mostly noise)
Input filter 2.2Ω and 4400 μF	−79 (mostly noise)	−90 (all noise)

2. Design your own high-performance regulator using discrete transistors or op-amps. This is not a straightforward business if all the protection that IC regulators have is to be included. There can also be distressing issues with HF stability.

3. Add an RC input filter between the reservoir capacitor and the regulator. This is simple and pretty much bulletproof, and preserves all the protection features of the IC regulator, though the extra components are a bit bulky and not that cheap. There is some loss of efficiency due to the voltage drop across the series resistor; this has to be kept low and the capacitance large.

The lower two rows of Table 20.3 show what happens. In the first case the filter values were 2.2 Ω and 2200 μF. This has a −3 dB frequency of 33 Hz and attenuates the 100 Hz ripple component by 10 dB. This has a fairly dramatic effect on the output ripple, but the dB figures do not change that much as the input filter does not affect the noise generated inside the regulator. Increasing the capacitance to 4400 μF sinks the ripple below the noise level for both types of regulator.

Dual Supplies From a Single Winding

It is extremely convenient to use third-party 'wall-wart' power supplies for small pieces of equipment, as they come with all the safety and EMC approvals already done for you, though admittedly they do not look appropriate with high-end equipment.

The problem is that the vast majority of these supplies give a single AC voltage on a two-pole connector, so a little thought is required to derive two supply rails. Figure 20.6 shows how it is done in a ±18 V power supply; note that these voltages are suitable only for a system that uses 5532s throughout. Two voltage-doublers of opposite polarity are used to generate the two unregulated voltages. When the incoming voltage goes negative, D3 conducts and the positive end of C1 takes up approximately 0 V. When the incoming voltage swings positive, D1 conducts instead and the charge on C1 is transferred to C3. Thus the whole peak-to-peak voltage of the AC supply appears across reservoir capacitor C3. In the same way, the peak-to-peak voltage, but with the opposite polarity, appears across reservoir C4.

www.raguvaran.puzl.com

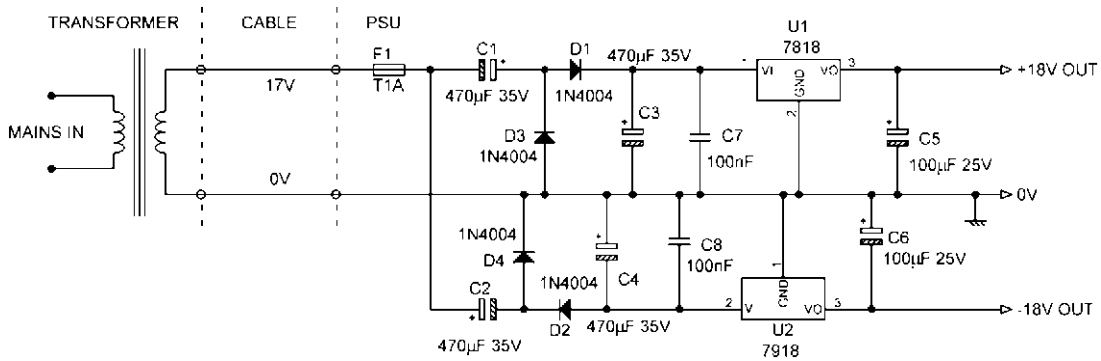


Figure 20.6: A ±18 V power supply powered by a single transformer winding

Since voltage-doublers use half-wave rectification, they are not suitable for high current supplies. When choosing the value of the reservoir capacitor values, bear in mind that the discharge time in Equation 20.1 above must be changed from 7 to 17 ms. The input capacitors C1, C2 should be the same size as the reservoirs.

Power Supplies for Discrete Circuitry

One of the main reasons for using discrete audio electronics is the possibility of handling larger signals than can be coped with by op-amps running off ±17 V rails. The use of ±24 V rails allows a 3 dB increase in headroom, which is probably about the minimum that justifies the extra complications of discrete circuitry. A ±24 V supply can be easily implemented with 7824/7924 IC regulators.

A slightly different approach was used in my first published preamplifier design [1]. This preamp in fact used two LM7824 +24 V regulators connected as shown in Figure 20.7 because at the time the LM7924 −24 V regulator had not yet reached the market. The use of a second positive regulator to produce the negative output rail looks a little strange at first sight but I can promise you it works. It can be very useful in the sort of situation described above; you have a hundred thousand +15 V regulators in store, but no −15 V regulators – I'm sure you see the point.

Note that this configuration requires two separate transformer windings; it cannot be used with a center-tapped secondary.

Larger Power Supplies

So, what if you need more than 1 amp of current? This will certainly be the case for all but the smallest mixing consoles. There are of course IC regulators with a greater current capability than 1 A. The LM338K variable-voltage 5 A regulators come in a TO-3 package – which it has

www.raguvaran.puzl.com

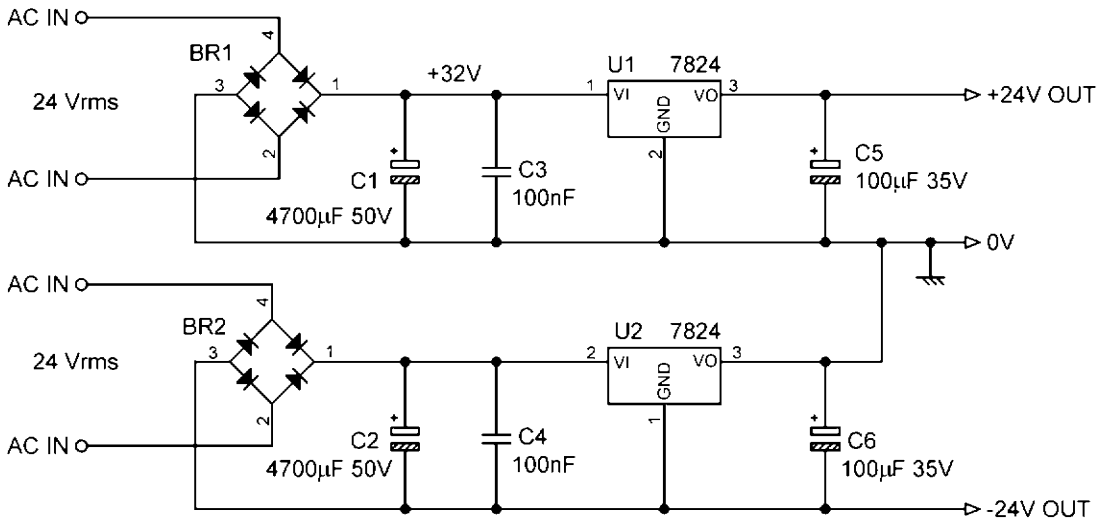


Figure 20.7: A ±24V power supply using only positive regulators

to be said is not the easiest format to mount. They were once widely used in mixing console power supplies but have been superseded by more modern devices such as the Linear Tech LT1083CK, which is a 7.5 A variable-voltage regulator that works very nicely, and can be obtained in a TO-3P package. I have used lots of them in mixing console power supplies. An example of their use, incorporating a mutual shutdown facility, is shown in Figure 20.8.

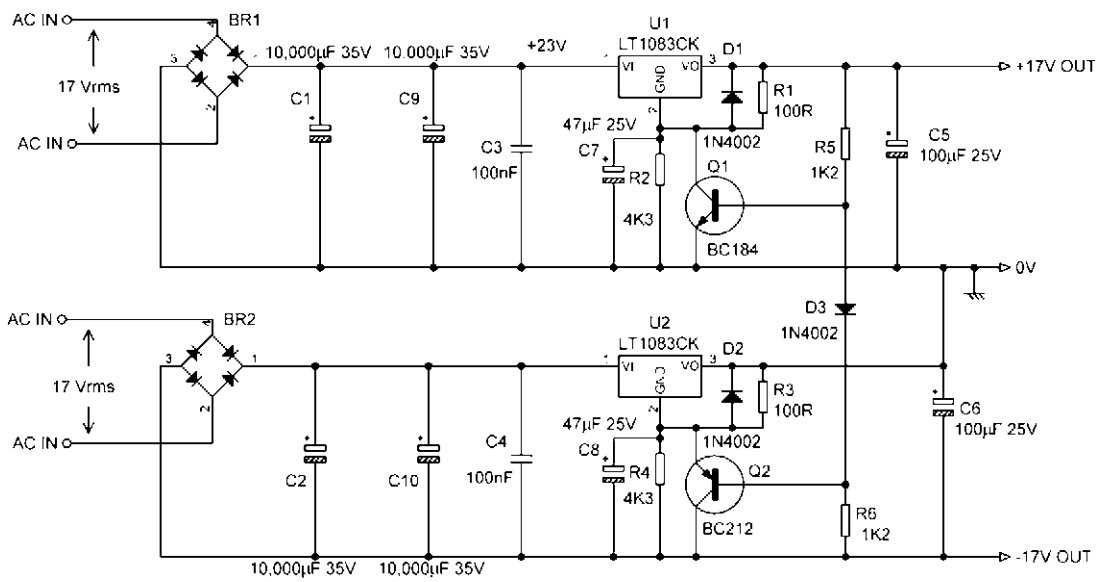


Figure 20.8: A high-current ±17 V power supply with mutual shutdown circuitry

www.raguvaran.puzl.com

There is no negative version of the LT1083CK, so a ± supply has to be made with two separate secondary windings, just as in Figure 20.7. When using any high-current regulator it pays to be a little cautious about the maximum rated current. Trying to use them too close to the maximum can cause start-up problems.

Mutual Shutdown Circuitry

It is an awkward quirk of 5532 op-amps that if one supply rail is lost and collapses to 0 V, while the other rail remains at the normal voltage, they can under some circumstances get into an anomalous mode of operation that draws large supply currents and ultimately destroys the op-amp by overheating. To prevent damage from this cause, which could be devastating to a large mixing console, the op-amp supplies are very often fitted with a mutual shutdown system. Mutual shutdown ensures that if one supply rail collapses, because of over-current, over-temperature or any other cause, the other rail will be promptly switched off. The extra circuitry required to implement this is shown in Figure 20.8, which is an example of a high-current supply using 7.5 A regulators.

The extra circuitry to implement mutual shutdown is very simple: R5, D3, R6, and Q1 and Q2. Because R5 is equal to R6, D3 normally sits at around 0 V in normal operation. If the +17 V rail collapses, Q2 is turned on by R6, and the Ref pin of U2 is pulled down to the bottom rail, reducing the output to the reference voltage (1.25 V). This is not completely off, but it is low enough to prevent any damage to op-amps.

If the −17 V rail collapses, Q1 is turned on by R5, pulling down the Ref pin of U1 in the same way. Q1 and Q2 do not operate exactly symmetrically, but it is close enough for our purposes.

Note that this circuit can only be used with variable-output-voltage regulators, because it relies on their low reference voltages.

Very Large Power Supplies

By very large I mean too big to be implemented with IC regulators – say 7 A and above. This presents a difficult problem, to which there are several possible answers:

1. Split up the system supply rails so that several IC regulators can be used. This is in my view the best approach. The amount of design work is relatively small; in particular, the short-circuit protection has been done for you.

2. Use power transistors as series-pass elements, controlled by op-amps. This can be a surprisingly tricky technology. The feedback system has to be reliably stable and the short-circuit protection has to be foolproof. Designing the latter is not too simple.

www.raguvaran.puzl.com

3. Use switch-mode power supplies. Those found in PCs seem to be a mature technology and are very reliable. Custom designs, of the sort required for this application, are in my experience quite another matter. There is also the issue of RF emissions.

The first two methods obviously involve increased heat dissipation in proportion to their output current. This usually means that fan cooling is required to keep the heat-sinks down to a reasonable size, which is fine for PA work but not welcome in a studio control room. There are no great technical difficulties in powering even a large console over a 20-meter cable, (this is the recommended maximum for Neve consoles) so long as remote sensing is used to compensate for voltage drops. This is usually long enough to allow the power supply to be placed in another room.

The technology of very large audio supplies is a specialized and complicated business, and it would not be appropriate to dig any further into it here.

Microcontroller and Relay Supplies

It is very often most economical to power relays from an unregulated supply. This is perfectly practical as relays have a wide operating voltage range. If 9 V relays are used then the same unregulated supply can feed a +5 V regulator to power a microcontroller, as shown in Figure 20.9.

Hum induced by electrostatic coupling from an unregulated relay supply rail can be sufficient to compromise the noise floor; the likelihood of this depends on the physical layout, but inevitably the signal paths and the relay supply come into proximity at the relay itself. It is therefore necessary to give this rail some degree of smoothing, without going to the expense of another regulator and heat-sink. (There must be no possibility of coupling between signal ground and relay power ground; these must only join right back at the power supply.) This

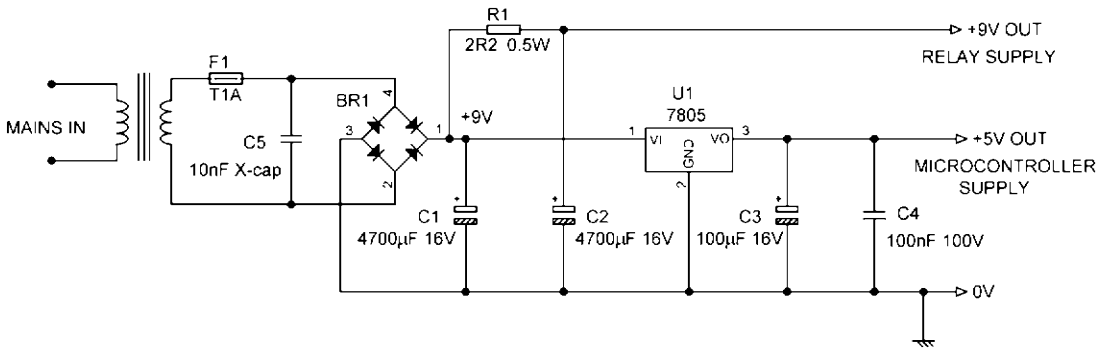


Figure 20.9: A +5 V PSU with an RC-smoothed +9 V relay supply

www.raguvaran.puzl.com

method of powering relays is more efficient than a regulated rail as it does not require a voltage drop across a regulator that must be sufficient to prevent drop-out and consequent rail ripple at low mains voltages.

Simple RC smoothing works perfectly well for this purpose. Relays draw relatively high currents, so a low R and a high value C are used to minimize voltage losses in R and changes in the relay supply voltage as different numbers of relays are energized.

The RC smoothing values shown in Figure 20.9 are typical, but are likely to need adjustment depending on how many relays are powered and how much current they draw. R1 is low at 2.2 Ω and C2 high at 4700 μF; fortunately the voltage is low so C2 need not be physically large.

+48 V Phantom Power Supplies

Making a discrete +48 V regulator with the necessary low amounts of noise and ripple is not too hard, but making one that is reliably short-circuit proof is a little more of a challenge, and by far the easiest way to make a +48 V phantom supply is to use a special high-voltage regulator called the TL783C, as shown in Figure 20.10. This extremely handy device can supply 700 mA, subject to power dissipation constraints.

It is a variable-voltage device, maintaining typically 1.27 V between the Out and ADJ pins. It combines BJT circuitry with high-voltage MOS devices on the same chip, allowing it to withstand much higher voltages than standard bipolar regulators. Since MOS devices are not subject to secondary breakdown or thermal runaway, the TL783 still gives full overload protection while operating with up to a 125 V voltage drop from the input to the output. The TL783 has current limiting, safe-operating-area (SOA) protection, and thermal shutdown.

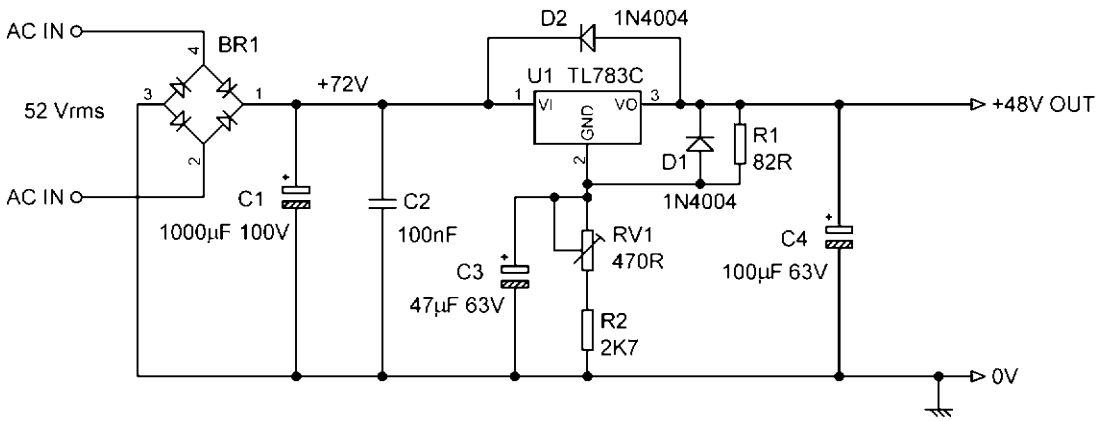


Figure 20.10: A +48 V phantom power supply using the TL783C regulator. Because of the large reference multiplication factor, a preset is required to set the output voltage exactly

www.raguvaran.puzl.com

Even if the ADJ pin is accidentally disconnected, the protection circuitry stays operational. It is a very useful and reliable IC, and I have deployed thousands of them.

As with other variable-voltage regulators, the low voltage maintained between the Out and ADJ pins needs to be amplified by a considerable ratio to get the desired output voltage, and so the reference voltage tolerances are also amplified. In this case the amplification factor is as high as 37 times, and so a preset is used to adjust the output voltage to exactly +48 V. The filter capacitor C3 is essential for the same reason – without it the ripple is amplified along with the reference voltage.

The unregulated supply can be derived from a completely separate transformer secondary as in Figure 20.10, or alternatively by means of a voltage-doubler. The latter is usually more economic, but obviously this depends on the cost of an extra transformer winding versus the cost of the extra capacitor in the doubler.

The arrangement of a voltage-doubler phantom supply is shown in Figure 20.11. Note that the familiar voltage-doubler circuit C13, C2, D5, D6 is actually working as a voltage-tripler,

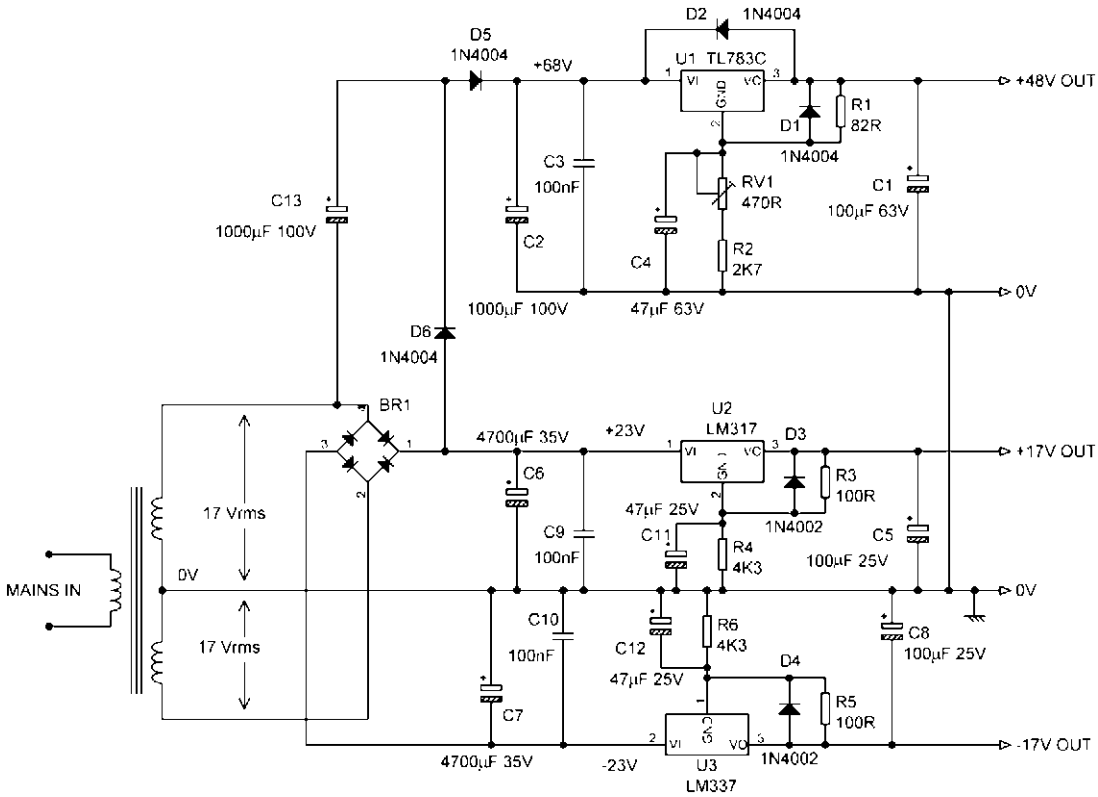


Figure 20.11: +48 V phantom power supply fed by a voltage-doubler

www.raguvaran.puzl.com

because it is perched on the unregulated +23 V supply to the +17 V regulator. If it worked as a true voltage-doubler, based on the 0 V rail, it would generate insufficient unregulated voltage for the phantom regulator. Because of their inherently half-wave operation and relatively poor regulation, voltage-doubler or voltage-tripler methods are not suitable for high-current phantom supplies.

Reference

[1] D.R.G. Self, An advanced preamplifier design, Wireless World (November 1976).

www.raguvaran.puzl.com

CHAPTER 21

Interfacing with the Digital Domain

The advance of digital audio has greatly improved the fidelity of audio storage media, and generally made wonderful things possible, but sound waves remain stubbornly analog, and so conversion from analog to digital and vice versa is very necessary. Today's analog-to-digital converters (ADCs) and digital-to-analog converters (DACs) have excellent performance, with 24-bit accuracy at a 192 kHz sampling rate commonplace, but to achieve this potential performance in an application there are a good number of factors that need to be appreciated. Some of them, such as the need for effective HF decoupling, are relatively straightforward provided you follow the manufacturer's recommendations, but others, involving the actual interfacing to the analog input and output pins, are a bit more subtle.

Having said that, contemporary ADCs and DACs are far easier to apply than their ancestors. Oversampling technology means that it is no longer necessary to put a ninth-order brickwall low-pass anti-aliasing filter in front of an ADC, or place a ninth-order brickwall low-pass reconstruction filter after a DAC. If you've ever tried to design a ninth-order filter to a price, you will know that this is a very significant freedom. It has had a major effect in reducing the price of digital equipment, particularly in applications like digital mixers, where a large number of ADCs and DACs are required.

ADC and DAC technology moves rapidly; the device examples I have chosen here (2009) will probably soon be out of date. The general principles I give here should be more enduring, and will be valid for the foreseeable development of the technologies.

PCB Layout Considerations

The PCB layout for both ADCs and DACs requires observance of certain precautions, which are basically the same for both functions. A double-sided PCB is necessary not only because of the large number of connections that have to be made in a small space, but also because it allows tracks that need to be isolated from each other to be put on opposite sides of the board. A most important consideration is to keep digital signals, particularly fast ones such as clocks, out of the analog inputs, so use as much physical spacing between these as possible. Critical tracks on opposite sides of the board should be run at right angles to each other to minimize coupling through the PCB. Do not run digital tracks topside under the IC as they may couple noise directly into the die from underneath.

Small Signal Audio Design; ISBN: 9780240521770
Copyright © 2010 Elsevier Ltd; All rights of reproduction, in any form, reserved.

www.raguvaran.puzl.com

Separate analog and digital ground planes should be used. Most conversion ICs have their analog and digital interfaces at opposite ends or opposite sides of the package, facilitating the use of separate ground planes. It is usually best to run the analog ground plane under the IC to minimize the coupling of digital noise. The two ground planes must of course be connected together at some point, and this should be implemented by a single junction close to the IC. Some manufacturers (e.g. Analog Devices) recommend that the junction should be made through a ferrite bead to filter out high RF frequencies. A maximum-copper (minimum-etch) PCB layout technique is generally the best for ground planes as it gives the most screening possible.

The power supply tracks to the IC should be as wide as possible to give low-impedance paths and reduce the voltage effects of current glitches on the power-supply lines. Ideally a four-layer PCB (and such boards are now cheaper than they have ever been) should be used so that two layers can be devoted to power-supply planes.

Thorough decoupling is always important when using high-speed devices such as ADCs and DACs. All analog and digital supplies should be decoupled to analog ground and digital ground respectively, using 0.1 μF ceramic capacitors in parallel with 10 μF electrolytic capacitors. Some manufacturers recommend using tantalum capacitors for this. To achieve the best possible decoupling, the capacitors should be placed as physically close to the IC as possible and solidly connected to the relevant ground plane.

When you are designing ADCs or DACs into a system, my experience is that significant time can be saved by doing preliminary testing on manufacturers' evaluation boards; this has particular force when you are using parts from a range you have not used before. Higher authority may urge you to go straight to a PCB layout, but unless you are very sure what you are doing – if, for example, you are cutting and pasting from an existing satisfactory design – it is a relatively high-risk approach. Evaluation boards are usually expensive, as they are produced in small quantities, but in my view it is money very well spent.

Nominal Levels and ADCs

The best use of the dynamic range of an ADC is only possible if it is presented with a signal of roughly the right amplitude. Too low a level degrades the signal-to-noise ratio as the top bits are not used, and too high a level will not only cause unpleasant-sounding digital clipping, but can cause damage to the ADC if current flows are not limited. Analog circuitry is therefore needed to scale the signal to the right amplitude.

A typical application of ADCs is in digital mixing consoles. These must accept both microphone and line-input levels. Since the signal level from a microphone may be very low (lute music) or quite high (microphone in the kick-drum), an input amplifier with a wide

www.raguvaran.puzl.com

variable gain range is required, typically 70 dB and sometimes as much as 80 dB. The signal level range of line signals is less but still requires a gain range of some 30 dB to cope with all conditions. It is therefore necessary for the operator to adjust the input gain, by reference to a level meter, so that good use is made of the available dynamic range without risking clipping. In live situations with unpredictable levels this is always something of a judgement call.

The signal level required at the ADC input to give maximum output, which is usually referred to as full scale (FS), varies from manufacturer to manufacturer; this important point is brought out in the next section.

Some Typical ADCs

There are a large number of ADCs on the market, and it is necessary to pick out just a few to look at. You will note that the various parts are actually very similar in their application. The inclusion of a device here does not mean that I am giving it any personal recommendation. All the devices mentioned are capable of 24-bit 192 kHz operation. In some cases the input voltage required for FS appears to exceed the supply voltage; this is not so, the quoted peak-to-peak voltage is the difference between two differential input pins. And now in alphabetical order:

The Analog Devices AD1871 is a stereo audio ADC with two 24-bit conversion channels each giving 105 dB of dynamic range, and each having a programmable gain amplifier (PGA) at the front end, a multi-bit sigma-delta modulator, and decimation filters. The digital details are rather outside our scope here and will not be alluded to further. The PGA has five gain settings ranging from 0 to 12 dB in 3-dB steps. The differential input required for FS is 2.828 $V_{pk\text{-}pk}$ and the input impedance is 8 kΩ. Like most of its kind, the AD1871 runs its analog section from +5 V, but the digital section from +3.3 V to save power. This IC is unusual in that it is permissible to run the digital section from +5 V, which can save you a regulator.

The Analog Devices AD1974 is a quad ADC with four differential analog inputs having a very useful CMRR of 55 dB (typical, at both 1 and 20 kHz). These inputs are not buffered and require special interfacing, which will be described later. A differential input of 5.4 $V_{pk\text{-}pk}$ is needed for FS; the input impedance is 8 kΩ. This IC runs from +3.3 V only.

The Burr-Brown PCM1802 is a stereo ADC with single-ended analog voltage inputs with input buffer amplifiers. It requires 3.0 $V_{pk\text{-}pk}$ to reach FS and has a resistive input impedance of 20 kΩ. The analog section is powered from +5 V, the digital section from +3.3 V.

The Wolfson WM8782 is a stereo ADC with two single-ended analog inputs with buffer amplifiers. It requires 2.82 $V_{pk\text{-}pk}$ (1.0 Vrms) to reach FS and the input impedance is 10 kΩ. The analog section is powered from +5 V and the digital core from +3.3 V.

www.raguvaran.puzl.com

Interfacing with ADC Inputs

The issues involved in interfacing with an ADC depend very much on how the ADC input is configured. As we saw in the previous section, some ADCs, such as the Burr-Brown PCM1802 and the Wolfson WM8782, have internal buffer amplifiers that present a relatively high impedance to the outside world (in this case 20 and 10 kΩ respectively). These inputs are very straightforward to drive. Such buffers are usually only found in ADCs made in a bipolar or bi-CMOS process, as making good low-noise, low-distortion amplifiers in a straight CMOS technology is very difficult.

Others, such as the AD1974, do not have buffering and must be driven from special circuitry. In the case of the AD1974 and similar devices, the differential inputs must be driven from a differential signal source to get the best performance. The basic principle is shown in Figure 21.1. The input pins connect to switched internal capacitors, and these generate glitches. Each input pin must be isolated from the op-amp driving it by an external series resistor R5, R6 together with a capacitor C5, C6 connected from input to ground. This capacitor must not generate non-linearity when the voltage across it changes, so ceramic NP0

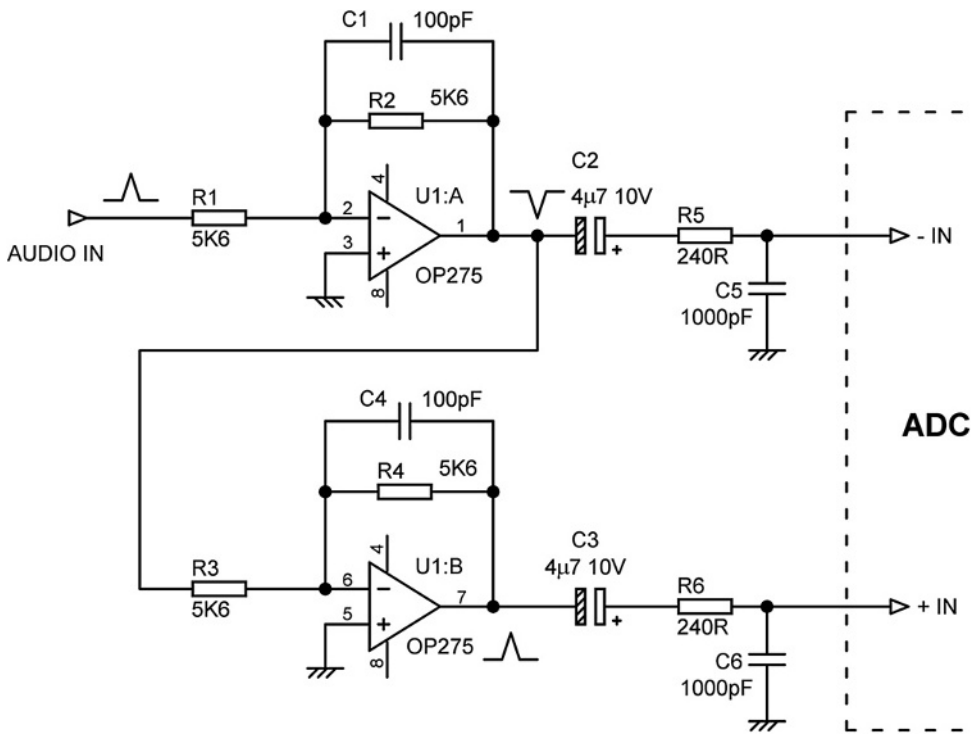


Figure 21.1: A typical drive circuit for an unbuffered differential ADC input

www.raguvaran.puzl.com

or polypropylene film types must be used; recommended values for the resistors and capacitors are usually given in the application notes. Note that since the external op-amps are referenced to ground, and the ADC internals are referenced to half the +5 V rail, blocking capacitors C2, C3 are needed.

Adding external resistance will slow down the charging of input sampling capacitors. These must be allowed to charge for many time-constants if they are to get close enough to the final value to avoid degrading the performance. External resistance increases the time-constant and can degrade accuracy; manufacturers usually provide guidance as to how much external resistance is permissible for a given number of bits of accuracy.

A point that is obvious but easily overlooked is that the inputs must not be driven to excessive levels. This usually means that the input voltage should not go outside the supply rails by more than 300 mV; for example, Wolfson specify this restriction for both the analog and digital inputs of the WM8782 stereo ADC, and other manufacturers quote similar ratings.

While ADC inputs invariably have clamp diodes for ESD protection that are intended to prevent the inputs moving outside the supply rails, these are small-dimensioned devices that may be destroyed by the output current capability of an op-amp. This is why the input voltage should not go outside the supply rails by more than 300 mV – this voltage will not cause a silicon diode to conduct significantly, even at elevated temperatures. The diodes can usually handle 5 mA, but to subject them to anything more is to live dangerously. Obviously the manufacturer's absolute maximum ratings should be followed on this point, but not all manufacturers give a current rating for their clamp diodes.

Bulletproof protection against input over-voltages is given by running the driving op-amp from the same supply rails as the analog section of the ADC, the op-amp saturation voltages ensuring that the input can never reach the supply rails, never mind exceed them. This does, however, restrict your choice of op-amp to one that is happy working on low supply voltages; these are likely to be more expensive than the popular audio op-amps such as the 5534/5532, which will not give good performance from such low rails.

If you want to stick with the usual audio op-amps, working from higher supply rails than the ADC, then an effective means of protection is the use of external clamping diodes, which in conjunction with a series resistance will limit the voltage swing at the ADC input. The principle is shown in Figure 21.2; if the op-amp output exceeds +5 V then D1 will conduct, while if it goes negative of 0 V D2 will conduct, safely clamping the ADC input.

A vital point here is that the clamp diodes must be of the Schottky type, so their forward voltage is substantially less than that of the conventional silicon diodes on-chip, for otherwise they will give little or no protection. The on-chip diodes will be warmer and would conduct before conventional external silicon diodes.

www.raguvaran.puzl.com

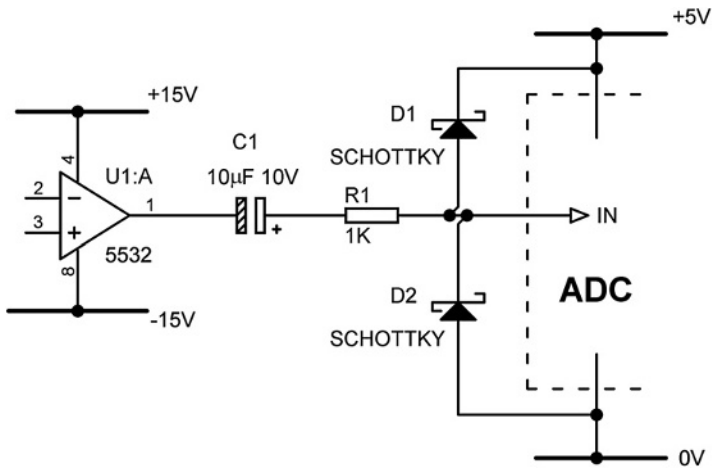


Figure 21.2: Diode clamping circuit to prevent overdriving an ADC input. Note that the diodes must be Schottky types

R1 must be large enough to limit the current in D1, D2 to safe levels, but not so large that it causes a roll-off with the ADC input capacitance. It must also not be so large that the non-linear capacitance of the diodes causes significant non-linearity; 1 kΩ should be safe in this respect. Note that R1 is also useful in isolating the op-amp output from the ADC input capacitance, which can otherwise erode stability margins.

Some Typical DACs

Unlike ADCs, DACs come in two different types – voltage output and current output. Both types of output require some kind of low-pass filtering, but current output DACs also need current-to-voltage (*I–V*) conversion stages. There are a large number of DACs on the market, and it is essential to be selective in examining a few typical devices. Once again, the inclusion of a device here does not mean that I am giving it my personal recommendation. All the devices mentioned here are capable of 24-bit operation. And now in alphabetical order:

The Analog Devices AD1854 is a stereo audio DAC delivering 113 dB dynamic range and 112 dB SNR (A-weighted) at a 48 kHz sample rate. Maximum sample rate is 96 kHz. Differential analog voltage outputs give a maximum output of 5.6 $V_{pk\text{-}pk}$ at FS and the output impedance is less than 200 Ω. It operates from a single +5 V supply rail, though there are separate supply pins for the analog and digital sections.

The Texas PCM1794A is a stereo audio DAC supporting sample rates up to 192 kHz. It has differential analog current outputs giving a maximum of 7.8 $mA_{pk\text{-}pk}$ at FS. The analog section is powered from +5 V, the digital section from +3.3 V.

www.raguvaran.puzl.com

The Wolfson WM8740 is a stereo audio DAC supporting word lengths from 16 to 24 bits and sample rates up to 192 kHz. Differential analog voltage outputs give a maximum output of 2.82 $V_{pk\text{-}pk}$ at FS. It can operate from a single +5 V supply rail, or the digital section can be run from +3.3 V to reduce power consumption.

Interfacing with DAC Outputs

Modern DACs use oversampling so that brickwall reconstruction filters are not necessary at the analog outputs. Nonetheless, some low-pass filtering is essential to remove high-frequency components from the output that could cause trouble downstream.

If you are using a DAC with a current output, the first thing you have to do is convert that current to a voltage. This is usually done with a shunt-feedback stage as shown in Figure 21.3, frequently called an *I–V* converter. The op-amp most popular for this job (and in fact explicitly recommended for the Texas PCM1794A) is no less than our old friend the 5534/5532. The filter capacitors C1, C2 keep down the slew rate required at the outputs of the *I–V* converters, and with their parallel resistors give a −3 dB roll-off at 88.2 kHz. The current output is simply scaled by the value of R1, R2 and results in a voltage output of 3.20 V_{pk} for 3.9 mA_{pk} out (half of the total 7.8 $mA_{pk\text{-}pk}$ FS output) and when the two anti-phase voltages are combined in the differential amplifier that follows the total output is 6.4 V_{pk} or 4.5 Vrms. The differential

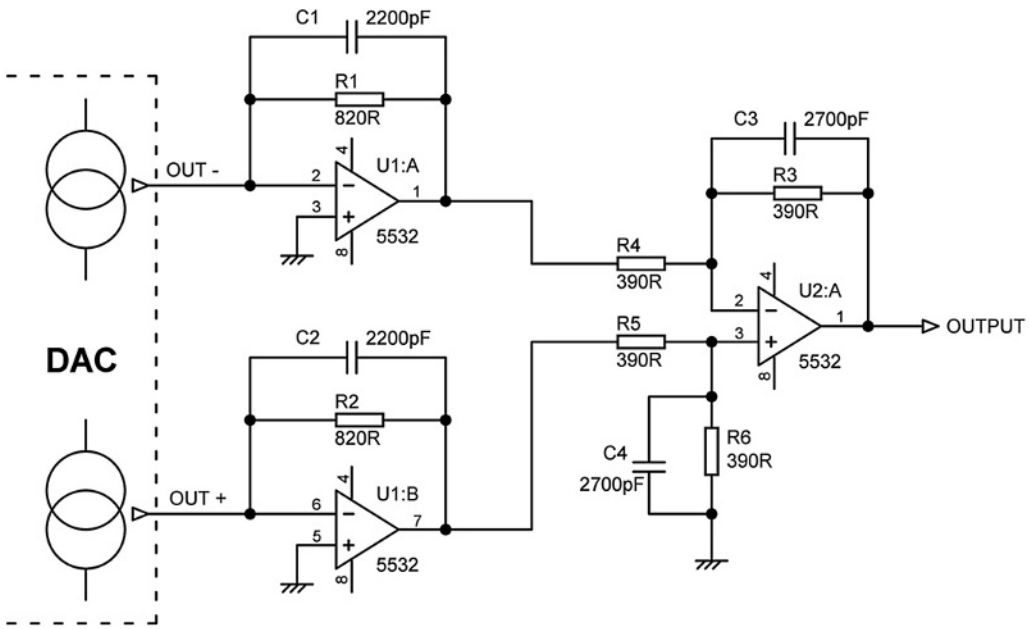


Figure 21.3: A typical output stage for a current-output DAC, with *I–V* converters and differential output filter

www.raguvaran.puzl.com

amplifier has its own HF roll-off at 151 kHz to give further filtering, implemented by C3 and C4. The capacitors used must be linear; NP0 ceramic, polystyrene, or polypropylene are the only types suitable.

Voltage output DACs are somewhat simpler to apply as there is no need for *I–V* converters and the outputs can drive an active low-pass filter directly. The output is usually differential to obtain enough voltage-swing capability within the limited supply voltage available, so differential to single-ended conversion is still required, and this is often cunningly implemented in the form of a differential low-pass filter.

Figure 21.4 shows a typical differential low-pass filter system; it has a third-order Bessel characteristic with a corner frequency of 92 kHz. The outputs are combined, and the first two poles are implemented by the differential multiple-feedback filter around U1:A and the third pole is produced by the passive network R7, C5. Note that the circuitry uses E96 resistor values in order to obtain the desired accuracy. Multiple-feedback filters are often preferred for this kind of application because they do not suffer from the failure of attenuation at very high frequencies that afflicts Sallen-and-Key filters, due to the inability of the op-amp to maintain a low impedance at its output when its open-loop gain, and hence its feedback factor, has fallen to a low value.

It must not be assumed from this that all DACs have differential outputs. For example, the Wolfson WM8726, described as a 'low-cost stereo DAC', has single-ended voltage outputs; it is recommended they are followed by a second-order low-pass filter.

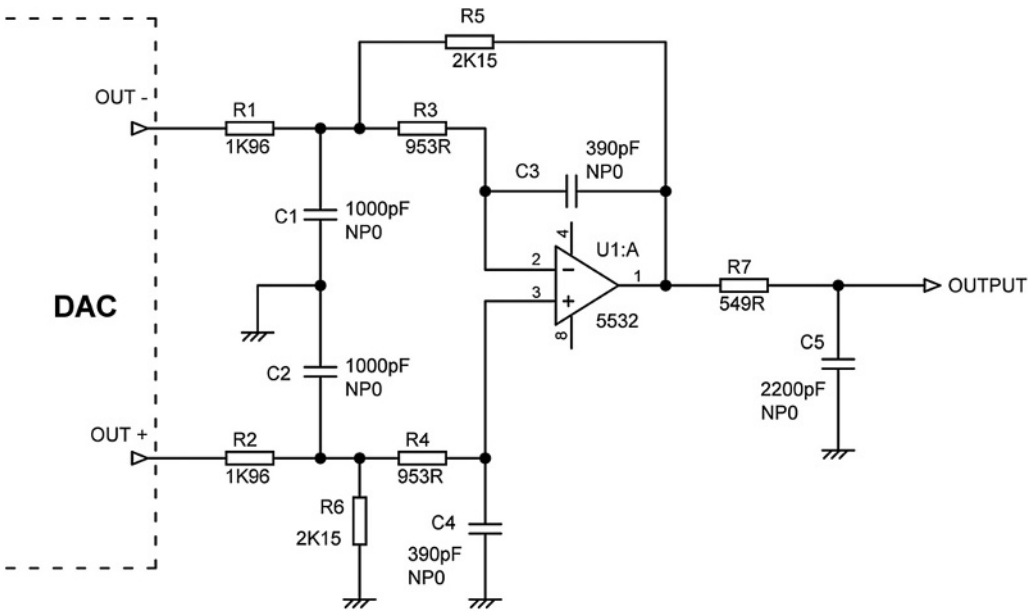


Figure 21.4: A typical output stage for a voltage-output DAC, with a differential output filter

www.raguvaran.puzl.com

Index

2N4403 transistor, 215, 216
2N5457 transistor, 494
2N5459 transistor, 494
2SB737 transistor, 20, 193, 196–8, 217, 459
2.0Db step switched attenuator, 246
2.5Db step switched attenuator, 247
3.0Db step switched attenuator, 248
±15V power supply, 527–9
±17V power supply, 529–31
+48V phantom power supplies, 538–40
4000 series analog gates, 398
4016 analog gates, 398, 398–400, 402, 405–7
4066 analog gates, 399–402, 404
5532 op-amp:
  active panpots, 439
  ADCs, 545, 547
  balanced inputs, 361
  bias current, 99
  bipolar input amps, 105–8, 198
  cost, 100, 126
  electronic cartridge loading, 200
  heavy loads, 389–90
  input offset voltage, 98
  LM 4562 op-amp, 121–3
  low-noise balanced inputs, 374
  MAGNOISE spreadsheet model, 193–5
  negative feedback, 120
  noise, 23, 96, 197
  OPA2134 comparison, 133
  postfade amplifiers, 428–30
  selection, 115–17
  slew rate, 97
  summing amplifiers, 456–7, 458
  supply-rail voltages, 525–6
  tone-control stage, 298
  virtual contact, 395
5534 op-amp, 194, 196, 198, 378, 428–30, 456–8, 545, 547
7824/7924 integrated circuits, 534
A-weighting and noise, 9, 196–7
Acronyms xx1–xx11
Active balance controls, 252–4
Active clipping:
  op-amps
    clamping, 514
    feedforward, 519–21
    negative feedback, 514–19
  transistors, 510–12
Active filters, 138
Active gain controls, 8, 158
Active gain stage (AGS) control system, 235, 237–8, 291, 299–301
Active input impedance, 201
Active panpots, 438–9
Active preamplifiers, 154–6
Active volume controls, 231–5
AD797 op-amp, 23, 96, 97, 100, 124, 378–80
AD1854 DAC, 546
AD1871 ADC, 543
AD1974 ADC, 543, 544
AFL systems, 471–2, 473
All-pass filters, 139–40, 504
ALPS SPUN switch, 393, 442
Ambler concept (tone), 269, 271
Amplification:
  moving-coil head amplifiers, 208–9
  preamplifier architectures, 156–7
  quiet, 25–6
  then attenuation, 6–7
Amplifiers:
  hybrid summing, 459–64
  CA3080E, 496
  Cambridge Audio 840A, 249
  components mismatch, 357–9
  current, 2, 3
  DBX2150 voltage-controlled, 498
  high input impedance, 92
  instrumentation, 368–9
  Johnson noise, 15–16
  line inputs, 340
  LM13600, 498
  LM13600/13700, 496
  Neumann SAL 74B, 168
  noise, 14–16
  Radford HD250, 157
  stereo power, 41
  SU900, 186
  summing, 456–65
  type, 2–3
    See also op-amps
Analog-to-digital converters (ADCs), 541–2, 542–3, 543–6
Attack artefacts, 503
Attenuation:
  quiet, 23–5
  then amplification, 7
Attenuator volume controls, 241–9
Audio Note Cartridge, 207
Audio Precision (AP) SYS-2702 measuring system, 71, 79
Augustadt, Herbert, 148
Auxiliary sends and panpots, 446

Bailey, Arthur, 185
Bainter filters, 148
Balance controls:
  active, 252–4
  mono/stereo switches, 254
  passive, 250–2
  width controls, 254–5
Balanced hybrid summing amplifiers, 462–5
Balanced inputs:
  cables and interference, 346–8
  combined unbalanced/balanced, 362–3
  connectors, 348
  high-impedance, 366–7
  instrumentation amplifier, 368–9
  interconnections, 344–6, 349–50
  inverting two op-amp, 367–8
  practical, 359–62, 362–9
  signal levels, 348
  switched-gain, 364–5
  variable-gain, 365–6
Balanced line outputs, 386
Balanced summing systems for panpots, 449–51
Balanced-feedback hybrid microphone preamplifier, 326–9
Bandpass filters, 138
Bar-graph metering, 486–8, 488–91
Bateman, Cyril, 52, 56
Baxandall configuration:
  active gain stage, 8, 237–8, 300
  active volume controls, 235–7
Baxandall, Peter, 235
Baxandall tone controls:
  middle, 271, 271–3
  one-capacitor LF control, 260–3
  switched frequency, 266–7
  three-banded EQ in one stage, 273–5
  two-capacitor HF control, 265–6
  two-capacitor tone control, 263–4
BBC, 241, 363, 503
Benet, Stephen Vincent, 525
Bessel filters, 141–2
Beta (h_{fe}) ratio, 66
Bias current in op-amps, 99–100
Bipolar junction transistors (BJTs), 8
  +48V phantom power supplies, 538
  op-amps, 97, 117–27
  OPA2134 op-amp, 134
Bipolar transistors:
  current noise, 17–20
  noise, 16–20
  voltage noise, 16–17
Blackmer, David, 498
Blue noise, 9
Boctor filter, 150
Bootstrapping:
  collector, 93
  junction field-effect transistors, 114–15
  series-feedback amplifiers, 116
  supply rail, 110–14
Braided screen cables, 346
Bridged-differentiator notch filter, 149
Bridging impedance, 312
Broadcast consoles, 304–5, 465, 506
Brownian motion, 9, 337
Bus:
  residual, 448, 458
  spelling history, 424
  trim control, 310, 319
Butterworth filters, 140–4, 168, 191, 201–4, 335

CA3080E amplifier, 496
Cables:
  braided screen, 346
  mercury, 35

www.raguvaran.puzl.com

Cables (*Continued*)
overlapping foil screen cables, 346–7
ribbon, 423, 425–6, 448, 451–2, 469, 471
wiring resistance, 36
Cambridge Audio:
840A amplifier, 249
840E preamplifier, 249
840W power amplifier, 377
P50 integrated amplifier, 235
Capacitance:
impedance and crosstalk, 41
Johnson noise, 10–11
PCB track-to-track crosstalk, 39
resistors, 45
Capacitors:
blocking, 52–3, 266, 268, 326, 329, 335, 340, 360, 362
ceramic, 52, 542
decoupling, 52, 129, 477, 488, 526, 539
electrolytic, 58–61, 458
head microphone amplifiers, 335–7
input sampling, 545
leakage, 51
non-linearity, 52–3
electrolytic, 58–61
non-electrolytic, 53–7
polyester, 53, 145–6
polyethylene terephthalate, 53
polypropylene, 51, 53, 136–7, 145–7, 203, 361, 545, 548
polystyrene, 53
properties, 51–3
reservoir, 527, 529, 532
tan-δ (tan-delta), 51
X2, 57
Capsule polarizing voltage, 135
Carbon composition resistors, 42, 47
Carbon-film resistors, 42, 46, 48
Cartridges:
DC/AC coupling, 183
impedances, 10, 181–2, 194, 207, 210
loading, 181–2, 194, 197
CAST *See* Current Audio Signal Transmission
CBS (broadcaster), 483
CD players, 155, 156, 163
Ceramic capacitors, 52, 542
Channel modules:
group module circuit blocks, 446
in-line, 319–21
routing, 443, 445
split-format, 310–13
Chebyshev (elliptical) filter, 140
Clamping diodes, 371, 545
Clipping:
biased diode, 510–11
by clamping, 514
diodes, 508–11, 514
feedforward, 519–21
negative-feedback, 514
op-amps, 512–21
soft, 474
transistors, 510–12
CMOS analog gates:
control voltage feedthrough, 407
current mode, 404–5
series-shunt current mode, 406–7
signal switching, 397–407
voltage mode, 399–404
CMOS technology, 544
Collector bootstrapping, 93
Colpitts oscillator, 74
Columbia (record label), 161
Combined LED indicators, 483
Combined low-pass/high pass filters, 139
Combined unbalanced/balanced inputs, 362–3
Common-mode (CM) range for op-amps, 98
Common-mode rejection:
balanced inputs, 354–5
line inputs, 349–52
op-amps, 354–5
Common-mode rejection ratio (CMRR):
amplifier component mismatch, 357–9
balanced inputs, 360
balanced interconnections, 349–51
electrostatic coupling, 346–54
high-input impedance balanced inputs, 366–7
instrumentation amplifier, 369
inverting two-op-amp, 368
low-noise balanced inputs, 372
magnetic coupling, 347
microphone pads, 329–31
microphone preamplifiers, 324
mixer inputs, 424
op-amp frequency response, 355–7
switched-gain balanced inputs, 364–5
transformer microphone inputs, 325
transformers, 346
unbalanced outputs, 381
variable-gain balanced inputs, 365–6
Complementary feedback pair (CFP):
configuration, 66, 73
constant-current, 77
emitter-follower circuits, 73–9
parametric middle EQs, 284
push–pull operation, 75, 78
two-transistor series feedback gain, 87
Component lifetimes in mixer subsystems, 478–9
Compressors, 499–506
Compressors/limiters:
attack artefacts, 503–4
combination, 500–3, 505
decay artefacts, 504–5
Conductors:
cables, 36
copper, 34–5
gold, 35–6
properties, 34
silver, 34
wire resistance, 36
Connor, Gareth, 465
Consoles:
broadcast, 304–5, 465, 506
cooling in mixer subsystems, 477–8
Constant-current emitter follower circuits, 70–1, 77
Constantan, 34
Control interaction, 259, 268, 272, 274, 276, 299
Control law, 219–23, 226, 228–36, 301, 324, 428, 430, 494–5, 499
Control-room monitor (CRM), 306–8
Cooling:
electronic, 201
fan, 537
Copper:
balanced cables and interference, 346
cable and wire resistance, 36
conductors, 33
metallurgy, 35
mixing bus systems, 424
moving-coil transformers, 209
PCB track resistance, 37–8
properties, 34–5
Correlated noise, 14, 195, 379
Cost of op-amps, 100
Crosstalk:
impedances, 40–2
inductors, 61
input-select switching, 394–5
inter-bus, 448
Current amplifiers, 2, 3
Current Audio Signal Transmission (CAST), 1
Current feedback amplifiers (CFAs), 3
Current source:
Johnson noise, 11
noise in amplifiers, 15–16
transconductance amplifiers, 2
See also individual circuits
Current-carrying capacity, 38
Current-mode connections, 1
Current-sharing, 211

DBX2150 voltage-controlled amplifier, 498
DC:
blocking capacitors, 52–3, 266, 268, 326, 329, 335, 340, 360, 362
discrete JFET switching, 420
drain resistors, 343, 361
potentiometers, 238–9
Decay artefacts, 504–5
Decoupling capacitors, 52, 129, 477, 488, 526, 539
'Destructive solo', 473
Differential amplifier array, 377, 378, 379
Differential filters, 150
Digital domain interface:
analog-to-digital converters, 541–2, 543–6
digital-to-analog converters, 541–2, 546–8
nominal levels and ADCs, 542–3
printed-circuit boards, 541–2
Digital-to-analog converters (DACs), 3, 138, 541–2, 546–8
DIN interconnection standard, 2
Dinsdale, Jack, 184
Diode clipping, 508–11, 514
Direct outputs in mixer subsystems, 430–1
Discrete circuitry and power supplies, 534
Discrete JFETs switching:
advantages, 407–8
changeover circuit, 420–1
control-voltage feedthrough in JFETs, 421
current mode, 414–17
DC, 420
JFET drive circuitry, 418–21
physical layout/offness, 419–20
reducing distortion by biasing, 417–18

www.raguvaran.puzl.com

voltage mode
series switch, 408–13
shunt JFET switch, 413–14
Discrete transistor circuitry:
bipolars/FETs, 64–5
gain stages, 80–8
high-input-impedance bipolar stages, 92–4
op-amps, 88–92
reasons for use, 63–4
transistor equation, 65–6
unity-gain buffer stages, 66–80
Distortion:
op-amps, 101–15
operational transconductance amplifiers, 496
reduction by biasing, 417–18
Sallen-and-Key filters, 145–7
ways of quoting, 102
Distributed gain control, 157, 224
Distributed peak detection, 481–3
Distributed summing systems for panpots, 452–6
Dominant-pole compensation, 183, 186
Dual supplies from a single winding (power), 533–4
Dual-action volume controls, 223–6
Dynamic range, 5–6, 24, 161–2, 345, 499, 502, 525, 542–3, 546
Dynamics section, 312, 426, 508

E3 capacitor values, 171
E12 resistor values, 242
E24 resistor values, 145, 168, 242–4, 365, 475
E96 resistor values, 168, 548
Early Effect, 65, 93, 108
Early, Jim, 65
Ebers–Moll transistor equation, 65–6
Effects return modules, 313
Effects return (split-format), 313–14
Efficient LED bargraph architecture, 488–9
Einstein, Albert, 12
Electret microphones, 317, 473
Electrolytic capacitors, 58–61, 458
Electrolytic tough pitch (ETP) copper, 35
Electronic balanced line inputs, 352–4
Electronic cooling, 201
Electronic switching, 397
Electronic v. transformer balanced inputs, 349
Electrostatic coupling, 346, 346–7, 479, 537
Emitter-follower (EF) circuits:
CFP, 73–9
constant-current, 70–1, 77
push–pull, 71–2, 78
simple, 67–70
stability, 72–3
End-stop resistors, 260, 263, 270
Equalizers (EQs), 61, 139, 267–9, 273–6, 280–3, 283–6
Equivalent input noise (EIN), 16, 87, 191, 193–8, 216–17, 326
Equivalent series resistance (ESR), 51
Evaluation boards, 542
Excess resistor noise, 46–8
External signal levels, 339

Fader flip facility, 310
Fader-start switches, 230
Faders:
linear, 240
short, 310, 319
slide, 229
Faders (mixer subsystems), 428
Fairchild uA709 op-amp, 95
Fan cooling, 537
Feedforward:
clipping, 519–21
compressors/limiters, 502–4
Field-effect transistors (FETs):
balanced summing systems, 449–51
CMOS analog gates, 398, 407
compressors/limiters, 502, 505
control-voltage feedthrough, 421
distributed peak detection, 483
gain-control, 494
JFETs in current mode, 414–15
line-up oscillators, 477
noise, 8
noise gates, 506–8
op-amps, 100, 108–10, 343, 402
prefade listen, 471
shunt JFET switch, 413
Filters:
active, 138
all-pass, 139–40, 504
Bainter, 148
bandpass, 139
Bessel, 141–2
Boctor, 150
Butterworth, 137, 140–4, 167–8, 191, 201–4, 335
characteristics, 140–5
Chebyshev (elliptical), 140
combined low-pass/high pass, 139, 204
differential, 150, 548
high-pass, 138–9, 146, 201–3, 335
low-pass, 138, 141, 204
multiple-feedback, 548
multiple-feedback bandpass, 147–8
notch, 139, 148–50
passive, 137
pinkening, 523–4
radio frequency, 341, 345
roofing, 137
Sallen-and-Key, 140–5, 201, 203
distortion, 145–7
state-variable, 139, 150
subsonic, 162, 201–4
subsonic/ultrasonic combination, 204
third-order, 202
Filters low pass, 138, 141, 143
First Age of Vinyl, 138
First Law of Thermodynamics, 11
Flicker noise, 8, 12, 17
Foldback, 305–6, 308, 313–14, 317
Four-layer summing, 456
Frequency response:
all-pass filters, 139
Bessel filters, 142–3
Butterworth filters, 141–3, 202–3
cartridge-preamplifier interaction, 183
electrolytic capacitors, 58
equalization, 166
high-pass filters, 146
moving-coil transformers, 209–10
negative feedback, 4
non-electrolytic capacitors, 53
notch filters, 148
op-amps, 355–7
RIAA amplifiers, 52
Sallen-and-Key filters, 145
ultrasonic filters, 143
vinyl, 163
violet noise, 9
Front of house (FOH) mixer, 305

Gain law, 237, 326, 331, 366, 501, 506
Gain stages, 120–1, 235
one-transistor shunt-feedback, 80–2
op-amps, 101, 120–1
two-transistor series-feedback gain stages, 87–8
two-transistor shunt-feedback, 82–7
Gain structures, 6–8
Gain-control:
history, 493
junction field-effect transistors, 493–5
operational transconductance amplifiers, 496–7
voltage controlled amplifiers, 495, 497–9, 502, 505
Gain-distribution and preamplifier architectures, 156–7
Galvanic isolation, 61, 346–7, 349, 388, 397
Gaumont, Leon, 431
Gershwin's Law ('it ain't necessarily so'), 33, 68
Gilbert, Barrie, 66
Gold, 13, 34–6, 241
Gold-plating, 35, 230
Graphic equalizers:
constant-Q, 289
description, 286–9
LCR filters, 61
proportional-Q, 289
Gray noise, 9
Green noise, 9
Groove curvature, 165
Ground plane, 40, 301, 423, 542
Ground summing systems for panpots, 451–2
Ground voltages, 348, 354, 383, 451
Ground-canceling outputs, 314, 383–4
Group modules:
circuit blocks and panpots, 446–7
split-format, 314–15

Heat sinks, 537
Heavy loads line outputs, 389–92
Hi-fi preamplifiers, 432
High frequency (HF) correction pole, 170
High input impedance amplifiers, 92
High-impedance inputs:
balanced, 366–7
bipolar stages, 92–4
preamplifiers, 293–4

www.raguvaran.puzl.com

High-pass filters, 138–9, 141, 144–8, 146, 193, 201–3, 296, 312, 335
Hitachi Cable Company, 35
Hofer, Bruce, 388
Holman, Tomlinson, 180
Hybrid microphone preamplifier, 325–6
Hybrid summing amplifiers, 459–64
Hysteresis, 192

I-V converters, 547, 548
IEC 98 standard, 166
IEC 60098 standard, 166
'IEC Amendment', 166–71, 175, 180–1, 191–2, 201, 291
Impedances:
  bridging, 312
  crosstalk, 40–2
In-line channel modules (mixer architectures), 319–21
In-line mixer architectures, 308–10
Inductance, 10, 15, 43–5, 61, 73, 171, 181–3, 189, 191–7, 207, 368, 370
Inductors, 61
Input:
  bias current, 92, 95, 108, 132, 238, 343, 428, 430
  offset voltage, 98, 368, 371–2
  op-amps, 97
  overvoltage protection, 371
  sampling capacitors, 545
  selector switch, 219, 291, 341
  signal and nominal level, 7–8
Input-select switching, 394–5
Insert:
  point, 312–13, 316, 319, 382, 423–6
  return, 314
  send, 314, 316, 425–6, 430, 456–7
Instrumentation amplifier, 368–9
Insulation-displacement connectors (IDCs), 423
Integrated circuits (ICs):
  transconductance amplifiers, 2
  volume controls, 250
Inter-bus crosstalk, 448
Interfacing with digital domain, 541–8
Internal signal levels, 339–40
Intrinsic emitter resistance, 9, 16
Inverse RIAA equalization, 178, 179
Inversion, quiet, 26–8
Inverting two op-amp balanced input, 367–8

Johnson, John B., 9
Johnson noise:
  amplifiers, 15–16
  description, 9–11
  electronic cartridges, 197, 199
  emitter-followers, 73
  field-effect transistors, 108–9
  low-noise balanced inputs, 375
  moving-coil transformers, 209
  noise, 12
  noise measurements, 23
  op-amps, 22
  RIAA preamplifiers, 191, 194
  switched attenuators, 241
  tone-control, 296
  virtual-earth summing, 448
Junction field-effect transistors (JFETs):
  5532 op-amp, 120–1
  bootstrapping, 114–15
  capacitor microphones, 335
  gain-control, 493–5
  input op-amps, 97, 117
  noise, 20–1
  op-amps, 110, 114–15, 128–36
  switching
    cascaded FET switches, 412–13
    changeover, 412
    control voltage feedthrough, 421
    current mode, 414–17
    DC blocking, 420
    drive circuitry, 418–19
    electronic, 397, 407
    physical layout/offness, 419–20
    reduced distortion by biasing, 417–18
    series switch (voltage mode), 408–13
    shunt switch (voltage mode), 413–14
    soft changeover circuit, 420–1

Larger power supplies, 534–6
Late Effect, 65
Laws:
  sine/sine squared, 431–2, 439–1
  thermodynamics, 11
  volume controls, 220–1
LCD meter displays, 491
Leakage from capacitors, 51
Left-Center-Right (LCR):
  filters, 137, 504
  graphic equalizers, 61
  panpots, 439–42
Level control and gain-control elements, 493–9
Level indication:
  distributed peak detection, 481–3
  peak, 479–81
  signal-present, 479
Light emitting diodes (LEDs):
  bar-graph metering, 486–8, 488–91
  combined, 483
  distributed peak detection, 481
  metering, 526
  moving circuit blocks, 427
  panpot offness, 437
  signal-present indication, 479–80
  virtual-earth PFL detection, 469
Limiters, 499–506
Line inputs:
  amplifier component mismatch, 357–9
  amplifier functions, 340
  balanced, 359–62, 362–9
  balanced cables and interference, 346–8
  balanced interconnections, 344–6
  balanced signal levels, 334
  common-mode rejection, 349–52, 354–5
  electronic balanced, 352–4
  electronic v. transformer balanced inputs, 349
  external signal levels, 339
  input overvoltage protection, 371
  internal signal levels, 339–40
  low-noise balanced, 372–7, 377
  op-amp frequency response, 355–7
  pads and microphone preamplifiers, 329–31
  preamplifiers, 293
  transformer balanced, 370–1
  ultra-low-noise balanced, 378–80
  unbalanced, 340–3
Line outputs:
  balanced, 386
  ground-canceling, 383–5
  heavy loads, 389–92
  quasi floating, 386–8
  transformer balanced, 388–9
  unbalanced, 381–2
  zero-impedance, 382–3
Line-up oscillators, 474–7
Linear faders, 240
Linearity in circuits, 68–9
Linkwitz–Riley crossover, 141
Lipshitz, Stanley, 175
'Listen' facility, 318
LM 833 op-amp, 96, 97, 100
LM317/337 variable-voltage IC regulators, 531–3
LM338K variable-voltage regulator, 534–5
LM741 op-amp, 96, 97, 103, 118
LM833 op-amp, 97, 100
LM3914 integrated circuit, 488
LM3915 integrated circuit, 488
LM4562 op-amp, 23, 96–8, 100, 121–4, 195, 198, 356, 378–80, 526
LM7815 fixed IC regulator, 527
LM7824 fixed IC regulator, 534
LM7915 fixed IC regulator, 527
LM7924 fixed IC regulator, 534
LM13600 amplifier, 496
LM13700 amplifier, 496
Load synthesis, 199–201
Loaded linear pots (volume controls), 221–3
'Low-impedance design', 219
Low-noise:
  balanced line inputs, 372–7, 377
  operational amplifier circuitry, 22–3
Low-pass filters, 138, 141, 143, 204
LP records (33⅓ rpm), 161
LT1028 op-amp, 23, 96, 97, 100, 195, 356, 378–9
LT1083CK variable-voltage regulator, 535–6

Magnetic conductors, 33
Magnetic coupling, 345, 347–8
MAGNOISE model, 192–3, 198
Manganin, 34
Master module (split-format), 316–17
Master section, 306, 308, 310, 316–17, 423, 465–6, 471, 473
MathCAD software, 192
MC33078 op-amp, 96, 97, 100
Mechanical switches, 393

www.raguvaran.puzl.com

Mercury:
cables, 35
properties, 34–5
Metal Electrode Face-Bonded (MELF) resistors, 43
Metal oxide semiconductor field-effect gates (MOSFETs), 398
Metal-film resistors, 42
Metal-oxide resistors, 42–3, 46
Metering:
LCD meter displays, 491
LED bar-graph, 486–8, 488–91
PPM meters, 485–6
vacuum fluorescent displays, 491
VU meters, 483–4
Metering and light emitting diodes (LEDs), 526
Mic/line switch, 312, 329, 333–4
Microcontroller and relay supplies, (power), 537–8
Microgroove technology, 161
Microphone capsules, 115, 324, 335
Microphone preamplifiers:
balanced-feedback hybrid, 326–9
capacitor head, 335–7
line-input pads, 329–31
padless, 331–5
requirements, 313–14
transformer microphone inputs, 324–5
Microphone transformer, 197, 325
Middle controls (tone), 271
Miller compensation, 120
Mixer architectures:
description, 305–6
in-line, 308–10
in-line channel, 319–21
internal levels, 304
performance, 303–4
split mixing, 306–8
split-format modules, 310–19
Mixer bus systems, 423–4
Mixer subsystems:
AFL systems, 471–3
component lifetimes, 477–8
console cooling, 477–8
direct outputs, 430–1
equalization, 424
faders, 428
group modules circuit blocks, 446–7
inputs, 424
insert points, 424–6
line-up oscillators, 474–7
mixer bus, 423–4
moving circuit blocks, 426–8
panpots, 431–42
PFL systems, 465–71
postfade amplifiers, 428–30
routing systems, 442–6
solo-in-place systems, 473
summing amplifiers, 456–65
summing systems, 447–56
talkback microphone amplifiers, 473–4
Molly-guard, 473
Monitor mix, 305, 308, 310, 314, 419
Monitor mixer, 305
Mono/stereo switches, 254
Motorized potentiometers, 239–40
Moving-coil (MC) head amplifiers:
amplification strategies, 208–9
cartridges, 207–8
circuit, 216
configuration, 213–15
input amplifiers, 210–13
performance, 216–17
transformers, 209–10
Moving-magnet (MM) disc inputs:
cartridges
DC/AC coupling, 183
impedances, 181
loading, 181–3
pre-amplifier interaction, 183
'Neumann Pole', 168
noise
A-weighting, 196
electronic cartridge loading, 197–201
measurements, 196–7
preamplifiers, 191–6
op-amps, 187–91
RIAA equalization
components, 175–8
discontents, 166–8
inverse, 178–9
passive/semi passive, 171–5
physical inverse, 178
spurious signals, 162–3
stages, 183–91
subsonic filters, 201–4
vinyl medium, 161–3, 163–5
Mr Spock (Star Trek), 40
Mu-metal, 139, 197, 209, 324
Multipath amplifiers and ultra-low-noise, 28–32
Multiple voltage controlled amplifiers (VCAs), 500
Multiple-feedback bandpass filters, 147–8
Multiple-feedback filters, 548
Mute bloc, 312, 414–17, 419–20
Mutual shutdown circuitry and power supplies, 536
Mutual-shutdown facility, 535

NBC (broadcaster), 483
NE5532 op-amp, 120–1, 356, 525
NE5532/5334 op-amp, 118–20
Negative feedback, op-amps, 519
Negative feedback (NFB):
5532 op-amp, 120
cartridge preamplifier interaction, 183
clipping, 514–19
discrete op-amps, 90
noise in RIAA amplifiers, 194
properties, 4–5
simple emitter-follower, 67
two-transistor shunt-feedback gain, 82–4
Negative-impedance-converter (NIC), 438
Neumann pole, 168
Neumann SAL 74B amplifier, 168
Nickel, 36, 44, 46
NJM4556 op-amp, 96, 97
NJM4556A op-amp, 389
Noise:
A-weighting, 196–7
active gain control, 238
amplifiers, 14–16
bipolar transistors, 16–20
blue, 9
correlated, 14, 195, 379
description, 8–9
electronic cartridge loading, 197–201
excess resistor, 46–8
flicker, 8, 12, 17
gain
5532, op-amp, 298
active volume controls, 232
balanced input, 361
balanced output, 386
bipolar input op-amps, 105
CMOS gates in current mode, 405
common-mode distortion, 104
DC, 420
ground-canceling summing systems, 452–3
JFETs in current mode, 417
LM4562 op-amp, 122
negative feedback clipping, 519
op-amps, 22, 27
OP275 op-amp, 129
OPA627 op-amp, 136
routing, 442–5
summing amplifiers, 458
virtual-earth summing, 448
gates, 312, 485, 506–8
generators, 26, 521–2, 521–3
gray, 9
green, 9
junction field-effect transistors, 20–1
measurements, 23
op-amps, 21–2, 197
pink, 9
red, 9
resistors, 46–8
RIAA
measurements, 196–7
preamplifiers, 191–6
routing systems, 444
simple balanced amplifiers, 361
sources, 13–14
summing systems, 454–5
violet, 9
voltage, 10, 12, 15–19, 21–3, 96–7, 195–6, 343, 374, 378–9
white, 9, 11
See also Johnson noise
Noise figure (NF), 16, 191
Nominal levels and ADCs, 542–3
Nominal signal levels, 5, 7–8, 339, 407
Non-conductors, 34
Non-electrolytic capacitors non-linearity, 53–7
Non-linearity:
capacitors, 52–3
electrolytic capacitors, 58–61
non-electrolytic capacitors, 53–7
resistors, 48–51
Normalling contacts, 312, 319, 423
Notch filters, 138, 148–50

Offness:
CMOS gates in voltage mode, 399–404
mute bloc, 416
panpots, 437, 444
physical layout, 419–20
routing systems, 444
Ohm, Georg Simon, 33
Ohm's Law, 33, 48
One-capacitor LF control (Baxandall tone controls), 260–3
'One-ounce copper', 37
One-transistor:
hybrid summing amplifiers, 460

www.raguvaran.puzl.com

One-transistor (*Continued*)
series-feedback gain stages, 82
shunt-feedback gain stages, 80–2
Op-amps (operational amplifiers):
active clipping, 512–21
bias current, 99–100
BJT input, 117–27
common-mode range, 98
common-mode rejection, 354–5
common-mode rejection ratio, 357
cost, 100
discrete, 88–92
distortion
bipolar input, 104–8
bootstrapping series-feedback JFET op-amps, 114–16
common-mode, 104, 106, 110–13
internal, 101
loading, 103–4
rail bootstrapping to reduce CM distortion, 110–13
simple rail bootstrapping, 113–14
slew-rate limiting, 101–3
thermal, 104
field-effect transistors, 100, 108–10, 343, 402
frequency response, 355–7
gain stages, 101
gain-distribution dilemma, 156
history, 95–6
input, 97
input offset voltage, 98
integrated circuits, 96, 374
JFET input, 128–36
Johnson noise, 22
line inputs, 341
loads, 103
microphone preamplifiers, 325
MM disc input stages, 187–91
negative feedback, 519
noise, 21–2
Sallen-and-Key filters, 141
selection, 115–17
supply-rail voltages, 525–6
See also individual names
OP27 op-amp, 96, 97, 100, 124–5, 195, 356
OP270 op-amp, 96, 97, 125–6
OP275 op-amp, 96, 97, 100, 117, 126–7
OPA604 op-amp, 96, 97, 132–4
OPA627 op-amp, 96, 97, 100, 356
OPA2134 op-amp, 96, 100, 110, 113, 117, 131–2, 134, 135–6, 198, 356, 403, 420
Operational transconductance amplifiers (OTAs), 2–3, 496–7
Oscillators, 9, 34, 48, 50, 73–4, 276, 317–18, 337, 474–7
Overlapping foil screen cables, 346–7
Overload margin, 163, 169–70, 172, 179–81, 186, 251

Padless microphone preamplifiers, 331–5
Palladium, 44
Panpots (panoramic potentiometers):
active, 438–9
auxiliary sends, 446
full-width LCR, 440–1
group module circuit blocks, 446–7
half-width LCR, 440–2
law bending, 433–4, 436–7
laws, 431–7, 438–9
Left-Center-Right, 439–42
mixer subsystems, 431–42
negative-impedance-converter, 438
offness, 437, 444
passive, 432–7
routing systems, 442–6
sine/sine squared laws, 431–2
summing systems, 447–56
Parametric equalizers, 285
Parametric middle EQs tone controls, 283–6
Passive balance controls, 250–2
Passive filters, 137, 504
Passive panpots, 432–7
Passive preamplifiers, 153–5, 250
Passive tone controls, 258–9
Passive/semi passive RIAA equalization, 171–5
PCM1794A stereo audio DAC, 546, 547
PCM1802 stereo ADC, 543, 544
Peak detection, distributed, 481–3
Peak indication, 479–81
Peak program meters (PPMs), 485–6
Perpetual motion machine of the second kind, 11
PGA2310 integrated circuit, 250
Phantom power, 312, 323–6, 336–7, 526, 538–9
Phono connectors, 348
Physical inverse RIAA equalization, 179
Pink noise, 9
Pinkening filters, 523–4
Platinum, 34–5
Polyester capacitors, 53, 145–6
Polyethylene terephthalate capacitors, 51, 53
Polypropylene capacitors, 51, 53, 136–7, 145–7, 203, 361, 545, 548
Polystyrene capacitors, 53
Popcorn noise, 13
Postfade amplifier, 312, 314, 428–9, 481
Postfade sends, 306, 313, 430
Potentiometers:
DC in, 238–9
motorized, 239–40
Power supplies, 'wall-wart', 533
Power supplies:
±15V, 527–9
±17V, 529–31
+48V phantom, 538–40
discrete circuitry, 534
dual supplies from a single winding, 533–4
larger, 534–6
microcontroller and relays, 537–8
mutual shutdown circuitry, 536
op-amps supply-rail voltages, 525–6
ripple performance, 531, 532–3
variable voltage regulators, 531–2
very large, 536–7
Power supply rejection ratio (PSRR), 526
Practical balanced inputs, 359
'Pre-echo', 164
Preamplifiers:
active, 154–6
active gain control, 158, 291, 299, 300–1
amplification, 156–7
architecture, 153–9, 291–3
description, 291–3
evolution, 154
gain-distribution, 156–7
Hi-fi, 432
high-impedance input buffer, 293–4
hybrid microphone, 325–6
line inputs, 293
LM7824, 534
microphone, 323–37
passive, 153–5, 250
performance, 301–2
philosophy, 291–3
postfade, 312, 314, 428–9, 481
precision, 181, 238, 291
recording, 158
tone controls, 158–9
tone-control stage, 294–300
Precision preamplifiers, 181, 238, 291
Prefade listen (PFL):
AFL, 471–2
channel module (split format), 312
detection, 468–9
in-line channel module, 319
master module, 317
peak indication, 479
simple, 468
split mixing, 306, 308
summing, 466
switching, 466–8
systems, 465–6
virtual-earth detection, 469–71
Prefade sends, 313
Printed-circuit boards (PCBs):
digital domain interface, 541–2
layout, 405, 541–2
mixer bus systems, 423–4
Signal Transfer Company, 191
track current capacity, 38
track resistance, 37–8
track-to-track crosstalk, 39–40
Progressive Electronics, 281
Protection diodes, 400, 528–9
Psychoacoustic weighting, 196
Public address (PA), 138, 305, 325, 388, 478
Push–pull:
CFP emitter-follower, 73–5, 76, 78
CFP follower, 78
emitter-followers, 71–3

Quasi-floating line outputs, 386–8
Quiet amplification, 25–6
Quiet attenuation, 23–5
Quiet balance, 28
Quiet inversion, 26–8

Radford HD250 amplifier, 157
Radio frequency (RF):
field strengths, 388
filters, 341, 345
Rail bootstrapping, 110, 112–15
RC smoothing, 538
RCA connectors, 348
RCA Victor (record label), 161

www.raguvaran.puzl.com

Recording Industry Association of America (RIAA):
amplifiers, 52, 375
differential amplifiers, 375
equalization, 178, 179
amplification, 208–9, 217
Audio Note cartridge, 207
cartridge-–preamplifier interaction, 183
components, 175
tolerances, 175–8
curve, 164, 166–8, 168–71
input signals, 7
inverse, 178, 179
moving-magnets disc input, 185
noise in preamplifiers, 191–3
non-electrolytic capacitors, 53
passive/semi passive, 171–4
physical inverse, 179
subsonic filters, 201
moving-magnet input stages, 184–6
noise in preamplifiers, 191–6
overload margins, 179–81
preamplifiers, 125
signal levels on vinyl, 163–5
Recording and preamplifiers, 158
Red noise, 9
Relay switching, 397
Relay-switched volume controls, 249, 249–50
Reservoir capacitors, 527–529, 532
Resistances and Johnson noise, 10
Resistors:
carbon composition, 42, 47
carbon-film resistors, 42, 46, 48
common-mode rejection ratio, 351
end-stop, 260, 263, 270
imperfections, 45–6
MELF, 43
metal-film, 42
metal-oxide, 42–3
noise, 46–8
non-linearity, 48–51
surface-mount, 43–5, 48–50
switched attenuators, 246–8
temperature coefficient, 34, 42–5
thick-film, 44, 46
thin-film, 44, 46
through-hole, 43–5
tolerances, 242, 331, 351, 358, 372, 375, 388
voltage coefficients, 48–9
wirewound, 43, 46
Return-to-flat (RTF) tone-control, 295, 298
RF filtering, 138, 340–1, 345, 360
Ribbon cable, 423, 425–6, 448, 451–2, 469, 471
Ripple performance (power supplies), 526, 528, 531, 532–3
Roofing filters, 137
Routing systems:
back grounding, 442
matrix, 430, 446, 451
offness, 444
panpots, 442–6
Rumble and record decks, 162
Ruthenium oxide, 44, 46

Safe-operating-area (SOA), 538
Sallen-and-Key filters:
capacitor microphones, 335
description, 140–5, 203
distortion, 145–7
subsonic filters, 201, 203
Schlotzaur configuration, 79, 327, 507
Screening, 39–40, 197, 209–10, 301, 324, 337, 346–7, 426, 542
'Sea of micro-diodes', 33
Second Law of Thermodynamics, 11
Series-feedback, 4, 22, 82–3
See also individual circuits
Servo integrator, 327, 507
Short fader, 310, 319
Shot noise, 8, 12
Shunt-feedback, 3–4, 22
See also individual circuits
Shure M75ED 2 cartridge, 193, 200
Sidechain, 295, 501
Signal switching:
CMOS analog gates, 397–407
discrete JFET switching, 407–21
input-select switching, 394–5
mechanical switches, 393
offness, 399–404, 416, 418–19
relays, 397
virtual contact, 395–7
Signal Transfer Company, 191, 217, 377
Signal-present indication, 479
Signal-to-noise, 2, 4–6, 24, 28, 187, 193, 211, 217, 304, 337, 386, 542
Signals, warp, 163
Signals on vinyl (RIAA), 163–5
Silver properties, 34, 209
Simple emitter-follower (EF) discrete circuit-block, 67–70
Sine laws, 431–2, 439–41
Sine-squared laws, 431–2, 439–41
Sine/cosine characteristic, 431–2
Single jack insert, 425
Single transistor series-feedback gain stage, 82–3
Single-gang sweep middle tone controls, 278–80
'Slate facility', 317
Slide faders, 229
Slide faders (volume controls), 229–30
Soft clipping, 474
Solo-in-place systems (SIP), 473
Soundcraft 3200 recording console, 465
SPICE simulation, 23, 48, 68–9, 76, 79, 178, 192, 202, 350, 355
Split mixing (mixer architecture), 306–8
Split-format modules:
channel, 310–13
effects return, 313–14
group, 314–15
master, 316–17
mixer architectures, 310–19
oscillators, 317–19
talkback, 317–19
Splitter box, 305, 388
Spurious signals and moving-magnet disc inputs, 162–3
Star Trek, 40
State-variable filters, 139, 150
Stepped volume controls, 240
Stereo image, 219, 221, 254, 294, 306
Stereo power amplifiers, 41
Stylus acceleration, 165
SU900 amplifier, 186
Subsonic filters, 162, 201–4
Subsonic/ultrasonic filters combination, 204
Subtractive VCA control, 505–6
Summing amplifiers:
5534/2 op-amp, 456–9
balanced hybrid, 462–5
hybrid, 459–62
one-transistor hybrid, 460
two-transistor hybrid, 461, 462, 463
Summing noise sources, 13–14
Summing systems:
amplifiers, 456–65
balanced, 449–51
bus residual, 448
distributed, 452–6
four-layer, 456
ground-canceling, 451–2
noise, 454–5
three-layer, 456
two-layer, 453–6
virtual-earth summing, 445, 448–9
voltage summing, 447–8
Superbal input, 363–4
Supply-rail voltages and op-amps, 525–6
Surface-mount (SM) resistors, 43–5, 48–50
Sweep middles tone controls, 276–8
Switched attenuator:
resistors, 246–8
volume controls, 241, 243, 250
Switched frequency Baxandall tone controls, 266–7
Switched peak/shelving LF/HF EQs tone controls, 281–3
Switched-gain balanced inputs, 364–5
Switched-Q variable-frequency Wien middle EQs tone controls, 280–1
Symmetrical clipping, 474
Szilaki configuration, 74

Talkback:
microphone amplifiers, 473–4
split-format modules, 317–19
Tan-δ (tan-delta in capacitors), 51
Tapped volume controls, 226, 226–9
Temperature coefficient, 34, 42–5
Thermal distortion, 104
Thermodynamics laws, 11
Thevenin–Norton transformation, 11
Thick-film resistors, 44, 46
Thin-film resistors, 44, 46
Third-order filters, 202
Three-layer summing, 456
Through-hole (TH) resistors, 43–5

www.raguvaran.puzl.com

Time-constants, 53, 166–7, 169–72, 174–5, 178–9, 295–6, 504–6, 523, 545
TL052 op-amp, 129–31
TL072 op-amp:
  Baxandall tone controls, 265
  cost, 100
  DC conditions, 420
  distortion, 115
  distortion due to loading, 103
  field-effect transistors, 109, 195, 238
  frequency response, 356–7
  junction field-effect transistors, 128–9
  low-cost mixing consoles, 117
  noise, 96, 193, 198
  noise generators, 522–3
  notch filters, 148
  offness, 402–3
  OPA2134 op-amp, 132
  postfade amplifiers, 428
  rail bootstrapping, 112
  RIAA networks, 180
  slew-rate, 97
  supply-rail voltages, 526
  volume controls, 216
TL783C high-voltage regulator, 538
Tolerances of resistors, 242, 331, 351, 358, 372, 375, 388
Tone controls:
  Ambler concept, 269, 271
  Baxandall, 259–67, 271–3, 273–5, 300
  description, 257–8
  graphic equalizers, 286–9
  middle, 271
  parametric middle EQs, 283–6
  passive, 258–9
  preamplifier architectures, 158–9
  return-to-flat, 295, 298
  single-gang sweep middle, 278–80
  sweep middles, 276–8
  switched peak/shelving LF/HF EQs, 281–3
  switched-Q variable-frequency Wien middle EQs, 280–1
  tone-balance, 269–71
  variable frequency LF/HF EQs, 267–9
  Wien fixed middle EQ, 276
Tone-balance controls, 269–71
Tone-cancel switch, 296, 300
Tone-control stage and preamplifiers, 294–300
Total harmonic distortion (THD):
  bipolar input amplifiers, 107
  bootstrapping series-feedback amplifiers, 116
  capacitors, 52–3
  distortion
    presentation, 101–2
    Sallen-and-Key filters, 145–6
  emitter follower circuits, 68, 77
  input overvoltage protection, 371
  LM741 op-amp, 118
  LM4562 op-amp, 122
  non-linearity
    capacitors, 53
    electrolytic capacitors, 58
    non-electrolytic capacitors, 53–5, 57
    resistors, 48, 51
  OP270 op-amp, 127
  resistor voltage coefficients, 48–9
  TL072 op-amp, 128–9
  two-transistor shunt-feedback, 83, 86
  unity gain buffers, 79–80
Track current capacity, 38
Track resistance in printed-circuit boards, 37–8
Track-return/group switch, 319
Track-to-track crosstalk, 39–40
Transconductance amplifiers, 2–3
Transformer balanced line inputs, 370–1
Transformer balanced line outputs, 388–9
Transformer-tap volume controls, 250
'Transient intermodulation distortion' (TID), 155
Transimpedance amplifiers, 3
Transistor equation, 65–6
Transmission gates, 397
Transmission of signals, 1–2
Twin-T notch filter, 148–9
Two layer summing, 453–6
'Two-ounce copper', 37
Two-transistor hybrid summing amplifier, 461, 462, 463
Two-transistor series-feedback gain stages, 80–2
Two-transistor shunt-feedback gain stages, 82–7

UA709 op-amp, 95
Ultra-low-noise:
  amplifiers, 30–2
  balanced line inputs, 378–80
  multipath amplifiers, 28–32
  voltage buffers, 29–30
Unbalanced line inputs, 340–3
Unbalanced line outputs, 381–2
Uncorrelated noise, 14, 211
Unipolar peak detector, 480
Unity-gain buffer stages, 64, 66–80, 90, 129, 140, 143, 145, 155–6

Vacuum fluorescent displays, 491
Van de Gavel, Marcel, 195, 197, 199–200
Variable frequency LF/HF EQs tone controls, 267–9
Variable voltage regulators, 531–2
Variable-gain balanced inputs, 365–6
Varigroove technology, 161, 162
Very large power supplies, 536–7
Vinyl disc, 161, 163–4, 303
Vinyl medium:
  description, 161–2
  maximum signals, 163–5
Violet noise, 9
Virtual contact (signal switching), 395–7
Virtual-earth PFL detection, 469–71
Vogel, Burkhard, 181
Voltage coefficient, 42–5, 48–50
Voltage controlled amplifiers (VCAs), 2, 495, 497–9, 500, 502, 505
Voltage doubler, 533–4, 539–40
Voltage noise, 10, 12, 15–19, 21–3, 96–7, 195–6, 343, 374, 378–9
Voltage-followers, 30, 110, 114, 372, 395
Volume controls:
  active, 231–8
  attenuator volume controls, 241–9
  Baxandall active, 235–8
    gain stage, 237–8
  dual-action, 223–6
  integrated circuit, 250
  laws, 220–1
  loaded linear pots, 221–3
  potentiometers
    and DC, 238–9
    motorized, 239–40
  relay-switched, 249–50
  slide faders, 229–30
  stepped, 240
  tapped, 226–9
  transformer-tap, 250
VU meters, 81, 483–5

Walker, H.P., 185
'Wall-wart' power supplies, 553
Warp signals, 163
Wet monitoring, 308, 321
White cathode-follower, 71
White, Eric, 71
White noise, 9, 11
Widlar, Bob, 95
Width controls (balance), 254
Wien bridge configuration, 475
Wien fixed middle EQ tone controls, 276
Wirewound resistors, 43, 46
Wiring resistance and cables, 36
WM8726 stereo DAC, 548
WM8740 stereo audio DAC, 547
WM8782 stereo ADC, 543, 544–5

X2 capacitors, 57
XLR connectors, 326, 345, 347–8, 372, 381–2

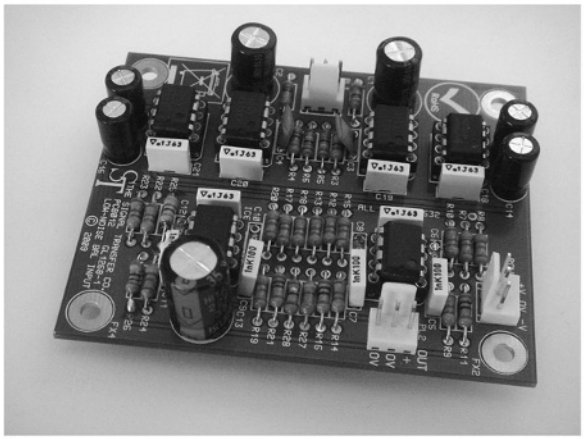
YAC523 integrated circuit, 250

Zero-impedance, 178, 382–3, 430, 4426
Zobel networks, 210, 324, 349, 370–1

www.raguvaran.puzl.com

The Signal Transfer Company

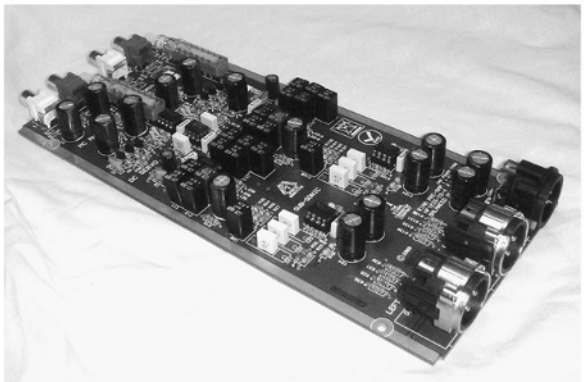
The Signal Transfer Company supplies PCBs, kits of parts, and fully built and tested modules based on the design principles in this book. All designs are approved by Douglas Self.



Low Noise Balanced Input PCB

A conventional balanced input stage built with 10K resistors and a 5532 op-amp has an output noise level of approximately –104 dBu.

The Signal Transfer low-noise balanced input card uses a multiple amplifier array that causes the noise from each amplifier to partially cancel, and in a similar way improves the common-mode rejection ratio; this array is driven by a multiple-buffer structure that allows the input impedance to be much higher than usual, preventing loading of external equipment and also further improving the CMRR. This elegant design does not require selected or exotic components. The noise output is less than –115 dBu.



Balanced Output Stereo Phono Preamp

This offers both moving-magnet and moving-coil inputs, the latter with two gain options for optimal performance. It gives extremely accurate RIAA equalization, a noise performance that approaches the theoretical limits, and superb linearity. It has fully balanced XLR outputs.

- Exceptional RIAA accuracy of +/–0.05 dB
- THD better than 0.002% at 6 Vrms. (Forty times the normal operating level)
- 3rd-order Butterworth subsonic filter: –3 dB at 20 Hz
- Separate input connectors for MM and MC inputs, switch selected
- Power indicator LED.

We supply the finest quality double-sided, plated-through hole, fiberglass PCBs. All boards have a full solder mask, gold-plated pads, and a silk-screen component layout. Each PCB is supplied with extensive constructional notes, previously unpublished information about the design, and a detailed parts list to make ordering components simple.

Signal Transfer products include several types of power amplifier, RIAA phono preamplifiers, and an enhanced version of the Precision Preamplifier described in Chapter 11. The specialized semiconductors required for some designs are also available.

For prices and more information go to: **http://www.signaltransfer.freeuk.com/**

Or contact:

The Signal Transfer Company
35 Hirst Grove
Dodd Naze
Hebden Bridge
West Yorkshire, HX7 8DN
United Kingdom

Tel: 01422-885196

www.raguvaran.puzl.com

This page intentionally left blank

www.raguvaran.puzl.com